

TIAGO GOMES DE ANDRADE

MECANISMOS REGULADORES DA SÍNTESE DE GLOBINAS:

Avaliação funcional da região R/PYR e análise da expressão gênica
diferencial na persistência hereditária de hemoglobina fetal
e na delta-beta talassemia

CAMPINAS

2006

TIAGO GOMES DE ANDRADE

MECANISMOS REGULADORES DA SÍNTESE DE GLOBINAS:

Avaliação funcional da região R/PYR e análise da expressão gênica
diferencial na persistência hereditária de hemoglobina fetal
e na delta-beta talassemia

*Tese de Doutorado apresentada à Pós-Graduação
da Faculdade de Ciências Médicas da Universidade
Estadual de Campinas, para obtenção do título de
Doutor em Fisiopatologia Médica, área de
concentração Biologia Estrutural, Celular,
Molecular e do Desenvolvimento*

ORIENTADOR: *Prof. Dr. Fernando F. Costa*

CAMPINAS

2006

*Dedico esta Tese,
fruto dos anos de estudo em Biologia Molecular,
que se iniciaram na aurora da Iniciação Científica e
se refletem mais maduros no término desta pós-graduação,
àqueles que sempre estiveram próximos, nem sempre em corpo,
mas presentes no Espírito de quem busca a Luz do conhecimento:
À minha família (Francisco, Tânia, Lucas e Mateus),
meus amigos e
minha companheira de sempre, Luciana.*

É grande o número daqueles que, de algum modo, contribuíram para o desenvolvimento deste trabalho. Elas formam ou formaram uma rede de possibilidades e acontecimentos únicos, de recomposição diria impossível, e imprimem nesta Tese sua colaboração particular, nem sempre conscientes disto. Sendo assim, cito apenas uma pequena porção de pessoas e Instituições que compõe esta rede e que tiveram um envolvimento mais direto ou óbvio com o Projeto.

Ao Dr. Fernando Ferreira Costa, pela receptividade e orientação ao longo da pós-graduação. Por ter me dado as condições possíveis para pensar e executar nossos projetos.

Ao Dr. Kenneth Peterson, da *University of Kansas Medical Center* (USA), pela sua receptividade, cuidado e por ter permitido a realização de parte desta Tese em seu laboratório nos seis meses de Doutorado sanduíche. A todos os membros de seu grupo de pesquisa: Renee, Adrian, Lesya, Taraz, Susanna, Halya, Anna e Flávia, pelo carinho e auxílio. Pelas possibilidades presentes e futuras de colaboração.

À FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), pelo auxílio financeiro.

Ao CNPq, Grupo Gerdau, Petrobrás/Procel, Fundação Roberto Marinho e à Comissão Julgadora do *XXI Prêmio Jovem Cientista 2005*, por terem me concedido *Diploma de Reconhecimento*, ao avaliarem o conjunto de trabalhos desenvolvidos ao longo de minha pós-graduação.

À Banca Examinadora desta Tese por dedicarem seu tempo na avaliação crítica do trabalho.

Ao Dr. Anderson Cunha pelos auxílios com clonagem e construção das bibliotecas.

A Tiago Ferraz e Tarcísio Perez pela ajuda na análise Bioinformática dos dados. Também, ao laboratório de Genômica e Expressão Gênica do Instituto de Biologia, pela utilização do BLAST local.

À Adriana Duarte e Carolina Lanaro pela cultura eritróide de duas fases.

À Dulcinéia Martins e Luciana Moreira pelos qRT-PCRs.

A todos os colegas, alunos, funcionários e docentes, do Hemocentro, Faculdade de Ciências Médicas e Instituto de Biologia-Unicamp, pela convivência e colaboração.

Aos portadores dos distúrbios talassêmicos, aos portadores das deleções estudadas e àqueles que doaram sangue para esta pesquisa.

Aos meus amigos distantes. Obrigado pela força e inspiração.

Ao meu Mentor dos primórdios científicos, Luiz Antônio, da Universidade Federal de Alagoas, por sua perspicácia intelectual. Certamente uma influência marcante para um jovem neófito da Ciência.

À Luciana novamente, pelo amor, carinho, paciência, e por ter me ensinado um pensar crítico, tão importante para a prática científica.

À minha família, pelo apoio, e, sem a qual, nada disto seria possível. Considero este trabalho uma extensão de seus esforços.

À Onipresente Natureza, causa e substrato de todas as realizações.

“O desenvolvimento da ciência é o desenvolvimento das idéias dos cientistas”.

Ernst Mayr (1904-2005)

*“Não há ciência pronta, acabada, dona da verdade.
Ela é uma tarefa perene em busca das quase-verdades que pode alcançar.
Por isso é modesta e é humilde. Não canta vitória antes do tempo e
sabe que suas “verdades” podem ser mudadas. Ela é histórica (está se fazendo) e
dialética (está sempre questionando suas próprias “verdades”).
Duvida de tudo, mas duvida principalmente de si mesma.”.*

Newton Freire-Maia (1918-2003)

*“Cada nova descoberta, cada novo pensamento pode
imprimir uma nova fisionomia ao mundo”.*

Carl G. Jung (1875-1961)

	<i>Pág.</i>
RESUMO	xxxiii
ABSTRACT	xxxvii
1- INTRODUÇÃO	41
1.1- Regulação da expressão gênica no <i>cluster</i> da beta-globina	43
1.1.2- Elementos <i>cis</i> regulatórios.....	44
1.1.2.1- Promotores.....	44
1.1.2.2- Acentuadores (<i>enhancers</i>).....	44
1.1.2.3- Silenciadores.....	45
1.1.2.4- Elementos “demarcadores” (<i>Boundary elements</i>).....	45
1.1.2.5- Elementos “isoladores” (<i>Insulators</i>).....	45
1.1.2.6- LCR (<i>Locus Control Region</i>).....	46
1.1.3- Fatores de transcrição.....	46
1.1.4- Remodelamento de cromatina.....	47
1.1.4.1- PYR.....	47
1.2- Alterações das hemoglobinas	48
1.2.1- Persistência Hereditária de Hemoglobina Fetal.....	48
1.2.2- Delta-beta Talassemias.....	49
1.3- Origem das deficiências gênicas	51
1.4- Efeito das deficiências gênicas e avaliação de seqüências regulatórias <i>in cis</i>	51
1.5- Avaliação de Fatores de Transcrição	53
1.6- Importância do estudo para a pesquisa básica e clínica	53

2- OBJETIVOS.....	55
3- CASUÍSTICA E MÉTODOS.....	59
3.1- Casuística.....	61
3.2- Métodos.....	61
3.2.1- Dados hematológicos.....	61
3.2.2- Diagnóstico molecular.....	62
3.2.3- Análise funcional da região R/PYR.....	63
3.2.3.1- Obtenção da construção.....	63
3.2.3.2- Transfecção em linhagens eritróides.....	66
3.2.3.3- RT-PCR.....	69
3.2.4- Identificação de genes diferencialmente expressos em PHHF-2.....	69
3.2.4.1- Biblioteca de cDNA Subtrativa Supressiva (<i>Subtractive Suppression Hybridization -SSH</i>).....	69
3.2.4.2- Extração de RNA de reticulócitos.....	71
3.2.4.3- Síntese e amplificação de cDNA para construção das bibliotecas.....	71
3.2.4.4- Purificação do cDNA e digestão com RsaI.....	74
3.2.4.5- Ligação dos adaptadores.....	76
3.2.4.6- Hibridizações.....	77
3.2.4.7- Amplificações por PCR.....	77
3.2.4.8- Clonagem e Sequenciamento.....	79
3.2.4.9- Análise das seqüências geradas.....	81
3.2.4.10- Diferenciação de progenitores eritróides humanos em cultura líquida de duas fases.....	82
3.2.4.11- PCR quantitativo em tempo real (qRT-PCR).....	83

4- RESULTADOS	87
4.1- Análise funcional da região R/PYR	89
4.2- Bibliotecas de cDNA Subtrativas Supressivas	93
5- DISCUSSÃO	109
5.1- Avaliação de seqüências regulatórias <i>in cis</i>	111
5.2- Identificação de fatores de transcrição potencialmente envolvidos com a regulação gênica das globinas	114
5.2.1- Globinas.....	115
5.2.2- Fatores de transcrição e remodeladores de cromatina.....	117
5.2.3- Genes associados a outras funções.....	120
5.3- Considerações finais	122
6- CONCLUSÕES	125
7- REFERÊNCIAS BIBLIOGRÁFICAS	129
8- ANEXOS	147
Anexo I (PCR Multiplex Fluorescente)	149
Anexo II (Manuscrito 1)	161
Anexo III (Manuscrito 2)	183

LISTA DE ABREVIATURAS

DNA	Deoxyribonucleic acid
dNTP	deoxinucleotídeo trifosfato
Hb F	Hemoglobina Fetal
HS	<i>Hypersensitivity sites</i>
Kb	Kilobase
LCR	<i>Locus Control Region</i>
L	Litros
MEL	<i>Murine Erytroleukemia</i>
µl	Microlitros
min	Minutos
ml	Mililitros
mmol	Milimol
ng	Nanogramas
pb	Pares de base
PCR	<i>Polymerase Chain reaction</i>
PHHF	Persistência Hereditária de Hemoglobina Fetal
pmol	Picomol
PYR	<i>Polypyrimidine-rich</i>
qRT-PCR	PCR quantitativo em tempo real.
RDB	Reticulócitos δβ talassemia Siciliana
RNA	<i>Ribonucleic acid</i>

RN	Reticulócitos normais
RP	Reticulócitos PHHF-2
r.p.m	Rotações por minuto
Seg.	Segundos
SSH	<i>Subtractive Suppression Hybridization</i>
U	Unidades

	<i>Pág.</i>
Tabela 1 Dados hematológicos dos indivíduos estudados.....	61
Tabela 2 Iniciadores utilizados para análise e validação das Biblioteca subtrativas.....	85
Tabela 3 Transcritos encontrados na biblioteca subtrativa normal (RN), correspondendo a genes com expressão diminuída em PHHF-2.....	94
Tabela 4 Transcritos encontrados na biblioteca subtrativa para PHHF-2 (RP), correspondendo a genes com expressão aumentada em PHHF-2.....	96
Tabela 5 Transcritos encontrados nas duas bibliotecas.....	100
Tabela 6 Distribuição dos transcritos encontrados nos cromossomos 11, 16 e 17 nas bibliotecas de reticulócito Normal (RN) e de PHHF-2 (RP).....	101
Tabela 7 Primers marcados com 6-FAM (*) para rastreamento de deleções no cluster da β -globina.....	149
Tabela 8 Iniciadores que formam Multiplex I.....	151
Tabela 9 Iniciadores que formam Multiplex II.....	151
Tabela 10 Iniciadores que formam Multiplex III.....	152
Tabela 11 Iniciadores utilizados nas reações com Elongase.....	153

		<i>Pág.</i>
Figura 1	Esquema da disposição dos genes e alguns elementos regulatórios no cluster da β -globina.....	43
Figura 2	Deficiências gênicas no <i>cluster</i> da beta-globina.....	50
Figura 3	Esquema das construções utilizadas para o estudo da região R/PYR.....	63
Figura 4	Padrão de restrição confirmando obtenção do clone intermediário R/PYR- β	65
Figura 5	Seleção do clone 7 μ LCR- β^m -R/PYR- β	65
Figura 6	Padrão de restrição do clone 7 com diferentes enzimas de restrição.....	66
Figura 7	PCR diagnóstico para presença das construções em célula MEL após transfecção estável.....	68
Figura 8	Padronização do número de ciclos de PCR para amplificação do cDNA a ser utilizado na construção da biblioteca subtrativa.....	73
Figura 9	Padronização do número de ciclos de PCR para amplificação do cDNA a ser utilizado na construção da biblioteca subtrativa.....	74
Figura 10	cDNA purificado em colunas GFX.....	75
Figura 11	Avaliação da digestão do cDNA com RsaI.....	75
Figura 12	cDNA digerido com RsaI após purificação.....	76
Figura 13	1 ^a amplificação do cDNA subtraído.....	78
Figura 14	2 ^a amplificação do cDNA subtraído (PCR <i>Nested</i>).....	79

Figura 15	Amplificação dos insertos das bibliotecas subtrativas com iniciadores M13.....	81
Figura 16	Avaliação do padrão de expressão das duas cópias do gene β por RT-PCR e digestão, após transfecção transiente.....	90
Figura 17	Avaliação do padrão de expressão das duas cópias do gene β por RT-PCR e digestão (NcoI) após transfecção estável em células MEL.....	91
Figura 18	Avaliação do padrão de expressão das duas cópias do gene β por RT-PCR e digestão (ClaI) após transfecção estável em células MEL.....	92
Figura 19	Transcritos diferencialmente expressos distribuídos por localização genômica.....	101
Figura 20	Padrão de expressão gênica avaliado por PCR em tempo real dos genes das globinas alfa (HBA), beta (HBB) e gama (HBG).....	103
Figura 21	Padrão de expressão gênica avaliado por PCR em tempo real dos genes TSPYL1, ARID1B, ZHX2, SH3BGRL2, NAP1L4 e CD47.....	104
Figura 22	<i>Citospin</i> (A) e citometria de fluxo com anticorpo para HbF (B e C) das amostras colhidas de cultura primária de sangue periférico de um indivíduo controle.....	106
Figura 23	Padrão de expressão gênica avaliado por PCR em tempo real dos genes TSPYL1, ARID1B, ZHX2, SH3BGRL2 e das globinas beta (HBB) e gama (HBG) em cultura primária de II fases, coletadas em 0 h, sem EPO e após 11 dias (216 hs) da adição de EPO.....	107
Figura 24	Correlação inversa entre o padrão de expressão de ZHX2 (barras) e gama globina (linha) nas amostras de PHHF estudadas.....	119
Figura 25	Modelo para explicar o efeito de diferentes formas de deleções dentro do <i>cluster</i> da β -globina.....	124

Figura 26	Determinação da fase exponencial da reação de PCR para o gene beta.....	150
Figura 27	Posição aproximada dos pares de primer utilizados no rastreamento de deleções novas, por dosagem gênica baseada em Multiplex fluorescente.....	152
Figura 28	Eletroesferogramas do Multiplex I.....	157
Figura 29	Gráfico comparando os dados do Coeficiente de Dosagem Gênica.....	158
Figura 30	Imagem invertida do gel de agarose 0,9 %, corado com brometo de etídio, com fragmento contendo ponto de quebra.....	159
Figura 31	A- representação da seqüência correspondente ao ponto de quebra da PHHF-2; B- Eletroesferograma com a seqüência do fragmento de 10Kb amplificado por PCR.....	160

RESUMO



Persistência Hereditária de Hemoglobina Fetal (PHHF) consiste num grupo heterogêneo de alterações hereditárias, sem manifestações clínicas significativas, onde ocorrem falhas na mudança perinatal normal de hemoglobina fetal para hemoglobina adulta, resultando em altos níveis de Hb F durante a vida adulta. Nos tipos deletoriais ocorre normalmente aumento de ambas as cadeias, γ^A e γ^G , até 30%, estando associadas a deleções de seqüências de DNA dentro do grupo de genes β . As $(\delta\beta)^\circ$ - talassemias também se originam a partir de deleções, em muitos casos bastante similares às que originam as formas deletoriais de PHHF, entretanto com níveis de Hb F aumentados em menor proporção (5-20%), sendo acompanhadas nos heterozigotos por hipocromia e microcitose nas hemácias. Três hipóteses principais foram propostas para explicar a relação destas deleções com a ausência de supressão normal dos genes γ na fase adulta: competição entre promotores pelo LCR (*Locus Control Region*); justaposição de elementos acentuadores, normalmente encontrados a 3' do cluster, nas proximidades dos genes γ ; remoção de elementos silenciadores, localizados entre γ^A e δ . Neste trabalho, realizamos estudos funcionais com uma região intergênica, aqui denominada R/PYR, potencialmente envolvida com o silenciamento de γ e a troca na síntese das cadeias de globinas. Os resultados obtidos com transfecção estável em células MEL apontam uma possível participação deste elemento na regulação dos genes de globina. Também investigamos o possível envolvimento de fatores de transcrição, construindo Bibliotecas de cDNA Subtrativas Supressivas para identificar transcritos diferencialmente expressos em reticulócitos de um sujeito com PHHF-2. A análise dos perfis diferenciais de expressão gênica em reticulócitos de PHHF-2 e normais mostram alterações em larga escala e que podem estar associadas com o fenótipo de PHHF. Estas modificações podem ser decorrentes de perturbações nas interações cromossômicas e/ou nos níveis de RNAs regulatórios não codificantes e ajudarem a melhor compreender a disposição e organização dos genes transcricionalmente ativos no genoma, e de que maneira são regulados. A presença de fatores de transcrição e remodeladores de cromatina, como ARID1B, TSPYL1 e ZHX2, dentre outros, com expressão alterada indicam fortemente que outros mecanismos genéticos estão envolvidos com o aumento da síntese de gama globina no estágio adulto, além dos modelos previamente descritos na literatura. Além disto, são candidatos potenciais a reguladores de globina que ainda não haviam sido descritos,

e podem ajudar a compreender as diferenças fenotípicas encontradas entre $\delta\beta$ -talassemia e PHHF deletional. Identificamos ainda outros genes com expressão diferencial, além de seqüências sem similaridade a genes categorizados, como ORFs, ESTs e seqüências genômicas, que podem ter papel importante na fisiologia normal das células vermelhas ou nas hemoglobinopatias hereditárias. Além disto, a observação de que os níveis de RNAm de alfa globina estão diminuídos em PHHF-2 e $\delta\beta$ -Talassemia Siciliana sugere um mecanismo de regulação pós-transcricional compensatório para as cadeias de globina ou de regulação gênica por interação física direta entre os cromossomos 11 e 16. Ao nosso conhecimento, esta é a primeira descrição demonstrando as implicações na expressão gênica em larga escala em células contendo alterações no DNA como as deleções apresentadas.

ABSTRACT



The genetic mechanisms underlying the continued expression of the γ -globin genes during the adult stage in deletional hereditary persistence of fetal hemoglobin (HPFH) and $\delta\beta$ -thalassemias are not completely understood. For deletional HPFH, three main hypotheses were proposed to explain the relationship between these deletions and the non-suppression of γ -genes in the adult; 1- the removal of competitive regions that interact with the LCR; 2- the juxtaposition of enhancer elements located downstream from the breakpoint region; and 3- the removal of gene silencer elements. Herein, we studied a region, called R/PYR, that lies 1-3 Kb upstream from the δ gene. This region has been implicated in γ globin repression and globin gene switching. Stable transfections in MEL cells suggest a possible involvement of R/PYR in globin regulation. We also investigated the possible involvement of transcription factors, using the SSH method to identify differentially expressed transcripts in reticulocytes from a normal and a HPFH-2 subject. Some of the detectable transcripts may participate in globin gene regulation. Quantitative RT-PCR experiments confirmed the downregulation of ZHX2, a transcriptional repressor, in two HPFH-2 subjects and in a carrier of the Sicilian $\delta\beta$ -thalassemia trait. The chromatin remodeling factors ARID1B and TSPYL1 had a very similar pattern of expression with an incremental increase in HPFH and decreased expression in $\delta\beta$ -thalassemia. These differences suggest a mechanism to explain the heterocellular and pancellular distribution of F cells in $\delta\beta$ -thalassemia and deletional HPFH, respectively. Interestingly, α -globin mRNA levels were decreased, similar to β -globin in all reticulocyte samples analyzed, which may be the result of a common regulatory process affecting α -like and β -like globin synthesis. Part of the altered gene expression detected within the subtracted libraries may be an indirect effect due to the increased concentration of HbF in the cells. However, genetic alterations similar to those presented in this work may be implicated in the perturbation of chromosome interactions and/or disruption of the expression of regulatory non-coding transcripts with functional consequences for cellular gene expression. Additionally, other genes may have modifications of their expression, but were not detected in our experiments. To our knowledge, this is the first description of wide-ranging gene expression alterations in cells containing deletional HPFH and thalassemia. Finally, the dissection of the

differential expression data presented in this study may help elucidate important pathways of erythroid differentiation and maintenance of red cells in peripheral circulation, as well as aid in the understanding of genetic mechanisms involved in erythrocyte pathologies such as sickle cell disease and thalassemias.

1- INTRODUÇÃO

Durante o desenvolvimento ontogenético normal ocorrem mudanças (*switching*) na expressão dos genes responsáveis pelos diferentes tipos de hemoglobinas humanas. Assim, durante o período embrionário, são encontradas as hemoglobinas Gower I ($\zeta_2\varepsilon_2$), Gower II ($\alpha_2\varepsilon_2$) e Portland ($\zeta_2\gamma_2$), no período fetal a Hb F ($\alpha_2\gamma_2$) e durante a vida adulta as Hb A1 ($\alpha_2\beta_2$) e Hb A2 ($\alpha_2\delta_2$) (Stamatoyannopoulos & Nienhuis, 1994). Estas mudanças estão relacionadas com a organização dos genes de globinas, que constituem uma família multigênica separada em dois grupos (*clusters*), no cromossomo 16, para os genes α , e cromossomo 11, para os genes β . O primeiro grupo inclui os genes para as cadeias zeta (ζ), alfa (α) e teta (θ), enquanto o segundo é constituído pelos genes que codificam as cadeias épsilon (ε), gama (γ^A e γ^G), delta (δ) e beta (β) (Figura 1). Estes genes são ativados durante o desenvolvimento na mesma ordem linear em que estão distribuídos no cromossomo, o que fornece a base para o desenvolvimento de modelos de regulação e expressão destes genes (Bunn & Forget, 1986).

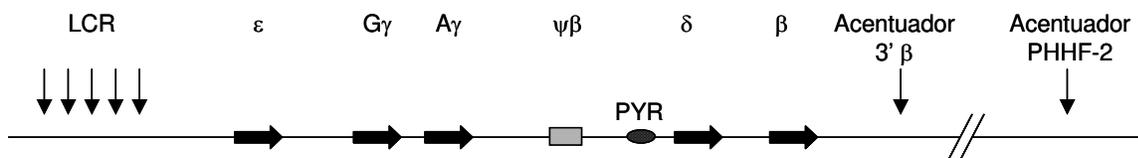


Figura 1- Esquema da disposição dos genes e alguns elementos regulatórios no cluster da β -globina. Algumas destas regiões serão discutidas mais adiante no texto. As setas indicam a posição de sítios hipersensíveis à DNase.

1.1- Regulação da expressão gênica no *cluster* da β -globina

A regulação dos genes dispostos no *cluster* da β -globina humana envolve duas classes de elementos: 1- elementos regulatórios *cis*, que compreendem seqüências dentro ou ao redor dos próprios genes (como promotores, silenciadores e acentuadores) ou distantes como o *Locus Control Region* (LCR); e 2- fatores de transcrição (elementos *trans*) ubíquos

ou eritróide específicos, que interagem entre si e com aquelas seqüências, formando uma rede complexa de interações proteína-proteína e DNA-proteína, e podem exercer seu papel diretamente na maquinaria de transcrição ou através de modificações na estrutura local da cromatina (Harju *et al*, 2002; Bank 2006). Acredita-se que a combinação de vários fatores seja necessária para taxas máximas ou adequadas de transcrição, e a habilidade para interagir em múltiplos padrões permite um potencial para uma enorme diversidade de funções, o que destaca a complexidade da regulação de sua expressão gênica (Jane, 1998; Stamatoyannopoulos *et al*, 2001). Em seguida, são apresentados alguns elementos conhecidos que participam na regulação dos genes das globinas do tipo beta. Apesar de sua descrição, ainda existem lacunas na compreensão da regulação gênica no *cluster* que dependem da caracterização de novos componentes da rede e de que maneira interagem nos diferentes estágio de desenvolvimento.

1.1.2- Elementos *cis* regulatórios

1.1.2.1- Promotores

A transcrição gênica depende da ligação da maquinaria basal de transcrição, o complexo da RNA Polimerase, em regiões específicas à montante dos genes, denominadas promotores. Seqüências TATA e CCAAT são encontradas a aproximadamente 30 e 70 pb à montante do início de transcrição e são fundamentais para o direcionamento correto da transcrição dos genes. Em adição, muitos genes eritróides possuem uma seqüência consenso CACCC, encontrada em duplicata no promotor de beta-globina, além de outros motivos de ligação para fatores de transcrição ubíquos (como Sp1, YY1 e Oct-1) e eritróide específicos (como GATA-1 e NF-E2) (Forget *et al*, 2001).

1.1.2.2- Acentuadores (*enhancers*)

Acentuadores são capazes de aumentar a expressão de um determinado gene, acima do que poderia ser direcionado apenas pelos promotores gênicos, independentemente da distância e orientação. Estes elementos podem estar localizados nas proximidades

montante ou jusante dos genes (Wall *et al*, 1988; Bodine & Ley, 1987), dentro dos próprios genes (seqüências *intrônicas*), ou a centenas ou milhares de pares de base distantes dos promotores, como será descrito mais adiante para o LCR e acentuadores localizados à jusante dos pontos de quebra das deleções (Figura 1).

1.1.2.3- Silenciadores

Elementos silenciadores se ligam a proteínas repressoras de transcrição que interferem na atividade dos promotores gênicos, diminuindo assim sua expressão. Um silenciador, por exemplo, localizado no promotor distal do gene ϵ controla sua repressão autônoma nos estágios fetal e adulto de desenvolvimento (Raich *et al*, 1992). Outros silenciadores potenciais foram localizados à montante dos genes γ e β (Berg *et al*, 1989; Stamatoyannopoulos *et al*, 1993). Ainda, como será discutido mais adiante, acredita-se que seqüências localizadas entre γA e δ podem ter efeito silenciador.

1.1.2.4- Elementos “demarcadores” (*Boundary elements*)

A principal característica destes elementos consiste em delimitar regiões de cromatina aberta e fechada, permitindo que estes subdomínios permaneçam nas conformações estáveis correspondentes (Gribnou *et al*, 2000). Elementos “demarcadores” podem estar presentes nas porções montante e jusante do *cluster*, definindo o domínio de cromatina do lócus da beta-globina, além de poder ser encontrados dentro do próprio lócus, demarcando subdomínios de cromatina associados a expressão gênica específica para cada estágio de desenvolvimento, a saber, embrionário, fetal e adulto (Gribnou *et al*, 2000).

1.1.2.5- Elementos “isoladores” (*Insulators*)

Estes elementos protegem uma região de efeitos negativos de heterocromatina localizada em sua proximidade e podem servir como elementos “demarcadores” que flanqueiam ou delimitam um domínio ativo de cromatina (Kellum & Schedl, 1991). O LCR possui propriedades de “isoladores”, já que transgenes ligados a este elemento são expressos em células eritróides independente do sítio de integração (Grosveld *et al*, 1987).

1.1.2.6- LCR (*Locus Control Region*)

A expressão normal dos genes dispostos no cluster da β globina depende da presença de um importante elemento regulador, denominado *Locus Control Region* (LCR), que é caracterizado por sua habilidade em dirigir uma alta expressão de genes ligados, de maneira tecido específica, em todos os sítios de integração examinados em experimentos com transgênicos (Bulger & Groudine, 1999; Li *et al*, 2002). O LCR consiste de cinco sítios hipersensíveis à DNase I (HSs) dispostos ao longo de 20-30 Kb, na região montante (5') do cluster (Figura 1). O modelo corrente para explicar o funcionamento do LCR propõe um mecanismo de abertura de cromatina, tornando-a ativa para transcrição, e interação do LCR com as regiões promotoras de cada gene, ao longo do desenvolvimento, através de formação de alças (*looping*) (Wijgerde *et al*, 1995). Outra proposta propõe a ligação em série de complexos protéicos (*linking*) (Bulger & Groudine, 1999). Entretanto, estudos recentes têm demonstrado a interação física entre sítios ativos do LCR e elementos regulatórios dos genes, o que reforça a hipótese de formação de alças (*looping*) no *cluster* (Patrinos *et al*, 2004).

1.1.3- Fatores de transcrição

Fatores de transcrição importantes para a regulação das globinas incluem NF-E2, um heterodímero composto por proteínas contendo *zipper de leucina*, uma subunidade de 45 kD (p45 NF-E2) e uma subunidade menor de 18 kD (p18 Maf) e que interage com sítios hipersensíveis no LCR (Andrews *et al*, 1993; Gong *et al*, 1996; Shavit *et al*, 1998; Fosberg *et al*, 2000); o *zinc finger* eritróide específico GATA-1, que possui múltiplos sítios de ligação nos promotores dos genes e no LCR, podendo ter papel ativador ou repressor de transcrição, dependendo do contexto da região de ligação e de sua interação com outras proteínas (Merika *et al*, 1993; Raich *et al*, 1995; Li *et al*, 1998); FOG (*friend of GATA-1*) é co-expresso com GATA-1 e interage com este em células eritróides e megacariocíticas durante o estágio embrionário de desenvolvimento (Tsang *et al*, 1997; Tsang *et al*, 1998); KLF1 ou EKLF (*Erythroid Krüppel-like factor*), um *zinc finger* ativador da transcrição de β -globina, via ligação com sítios CACCC

(Miller *et al*, 1993). Também tem participação na estrutura cromatínica do lócus e tem sido implicado no processo de troca de globinas (Donze *et al*, 1995; Lee *et al*, 1999); KLF11 ou FKLf (*Fetal Krüppel-like factor*) também regula a expressão de β -globina, porém em menor grau. Sua participação na expressão dos genes ϵ e γ parece ser mais importante (Asano *et al*, 2000).

1.1.4- Remodelamento de cromatina

Dados crescentes têm demonstrado uma participação importante de modificações de cromatina associadas ao padrão de expressão dos genes das globinas (Harju *et al*, 2002; Bank *et al*, 2005). Estas alterações incluem fosforilação, acetilação e metilação de histonas, que tornam a cromatina acessível em maior ou menor grau. Vários fatores modeladores de cromatina, como proteínas do complexo SWI/SNF (*switch/sucrose non-fermenting*) e CBP (*CREB-binding protein*)/p300 foram descritos em associação com fatores de transcrição eritróide específicos. O remodelamento de cromatina no lócus da β -globina pode determinar qual gene será alvo para ativação da transcrição. A análise de transcrição intergênica e da sensibilidade à DNase em camundongos transgênicos contendo todo o *cluster* da β -globina humana sugere a existência de pelo menos dois subdomínios ativos durante o desenvolvimento: um subdomínio embrionário/fetal e outro adulto (Jimenez *et al*, 1992; Gribnou *et al*, 2000; Plant *et al*, 2001). A cromatina no LCR, por sua vez, é constantemente ativa.

1.1.4.1- PYR

O complexo remodelador de cromatina PYR se liga a uma seqüência de DNA de 250 pb rica em pirimidinas (*polypyrimidine-rich*-PYR) a aproximadamente 1 Kb à montante do gene δ (Figura 1). Esta região pode compreender um elemento “demarcador” dos subdomínios fetal e adulto, e a ligação deste complexo pode influenciar na conformação de cromatina local, modificando a expressão dos genes de globinas. A ligação ao DNA é dependente da proteína *zinc finger* Ikaros, expressa em tecido

hematopoético adulto, e cuja ausência inviabiliza a formação do complexo (Lopez *et al*, 2002). Animais nulos para *Ikaros* possuem um atraso na troca dos genes das globinas, o que sugere que PYR pode ter participação neste processo (Lopez *et al*, 2002). PYR foi isolado de células MEL (*Murine Erythroleukaemia*), que possuem um ambiente protéico favorável à expressão de genes adultos (O'Neill *et al*, 1991). Sua caracterização bioquímica revelou a presença de subunidades ativadoras de transcrição, do complexo SWI/SNF e subunidades repressoras do complexo NuRD (O'Neill *et al*, 1999; O'Neill *et al*, 2000). Estas características tornam o complexo PYR um potencial regulador da regulação da transcrição dos genes das globinas do tipo β ao longo do desenvolvimento (Bank, 2006).

1.2- Alterações das hemoglobinas

As alterações das hemoglobinas podem ser classificadas em quatro grandes grupos: as variantes estruturais, sendo a Hb S, responsável pela anemia falciforme, a mais importante e freqüente, as síndromes talassêmicas, as anomalias do padrão de desenvolvimento ontogenético normal e algumas alterações adquiridas de sua produção (Weatheral & Clegg, 2001; Bradley & Ranney, 1973; Bunn & Forget, 1986; Serjeant & Serjeant, 2001).

1.2.1- Persistência Hereditária de Hemoglobina Fetal

Persistência Hereditária de Hemoglobina Fetal (PHHF) consiste num grupo heterogêneo de alterações hereditárias, sem manifestações clínicas significativas, onde ocorrem falhas na mudança perinatal normal de hemoglobina fetal para hemoglobina adulta, resultando em altos níveis de Hb F durante a vida adulta (Forget, 1998; Jane, 1998; Bollekens & Forget, 1991). As alterações moleculares ocorrem no grupo de genes β , sendo divididas em duas categorias: as formas delecionais e não delecionais de PHHF.

Os tipos não delecionais caracterizam-se, de forma geral, por aumentos significativos de um dos tipos específicos de cadeias γ , seja γ^A ou γ^G , associados a mutações de ponto nos promotores dos genes responsáveis pela síntese destas cadeias.

Propõe-se que estas alterações interfiram na ligação específica de elementos em *trans* com as seqüências promotoras, seja prevenindo a ligação de elementos supressores ou facilitando a ligação de fatores de regulação positiva (Forget, 1990).

Nos tipos deletoriais ocorre normalmente aumento de ambas as cadeias, γ^A e γ^G , até 30%, estando associadas a deleções de seqüências de DNA dentro do grupo de genes β , e uma distribuição homogênea (pancelular) dos altos níveis HbF nas células vermelhas. Um total de seis deleções relacionadas a PHHF já foi descrito na literatura (Forget, 1998) (Figura 2). PHHF-1 (Collins *et al*, 1987; Feingold & Forget, 1989) e PHHF-2 (Collins *et al*, 1987; Henthorn *et al*, 1990), caracterizam-se por grandes deleções de aproximadamente 105 Kb (Kilobases) com pontos de quebra deslocados apenas 5-6 Kb. PHHF-3 (Kutlar *et al*, 1984; Henthorn *et al.*, 1986) e PHHF-4 (Saglio *et al*, 1986) são menos extensas, envolvendo aproximadamente 50Kb e 40Kb respectivamente. PHHF-5 (Camaschella *et al*, 1990) é relativamente pequena e estende-se do ponto aproximadamente 3Kb à montante (5') do gene δ até 0,7 Kb à jusante (3') do gene β . PHHF-6 (Winichagoon *et al.*, 1990; Kosteas *et al*, 1997) tem tamanho similar a PHHF dos tipos 1 e 2 (aproximadamente 101 Kb), mas desloca-se na direção 5' e envolve a deleção do gene γ^A . As deleções responsáveis por PHHF são representadas na Figura 2.

PHHF-1 e PHHF-2 já foram descritas na população brasileira (Gonçalves, 1993; Gonçalves *et al*, 1995, De Andrade *et al*, 2002), e, segundo os estudos realizados pelos programas de triagem do Hemocentro e do Hospital das clínicas da Unicamp para alterações das hemoglobinas, a freqüência de Persistência Hereditária de Hemoglobina Fetal, ou seja, indivíduos com Hb F elevada, nesta população, é de aproximadamente 0.1%. Além disto, a análise regular das hemoglobinas em doadores de sangue e pacientes do Hemocentro da Unicamp possibilitam a identificação de numerosos casos, com alterações moleculares ainda não definidas.

1.2.2- Delta-beta Talassemias

As $(\delta\beta)^0$ talassemias também se originam a partir de deleções, em muitos casos bastante similares às que originam as formas deletoriais de PHHF, entretanto com níveis de HB F aumentados em menor proporção (5-20%), sendo acompanhadas nos heterozigotos

por hipocromia e microcitose nas hemácias (Weatheral *et al*, 2001). Diferentemente das formas deletoriais de PHHF, a expressão elevada de Hb F se dá em apenas algumas células (células F), de distribuição heterogênea (heterocelular).

A forma mais comum ocorre na região do Mediterrâneo (siciliana) (Henthorn *et al*, 1990), mas outras formas foram descritas em famílias do sudeste asiático (Zhang *et al*, 1988; Trent *et al*, 1988), leste europeu (Palena *et al*, 1992), afro-americanas (Negra) (Anagnou *et al*, 1985; Waye *et al*, 1994), turcas (Craig *et al*, 1995; Öner *et al*, 1996), indiana (Gilman *et al*, 1992), espanholas (Feingold & Forget, 1989) e do Japão (Shiokawa *et al*, 1988). A Figura 2 mostra a disposição de algumas deleções, ao longo do cluster da β globina, que originam Persistência Hereditária de Hemoglobina Fetal e as $(\delta\beta)^{\circ}$ talassemias siciliana e afro-americana.

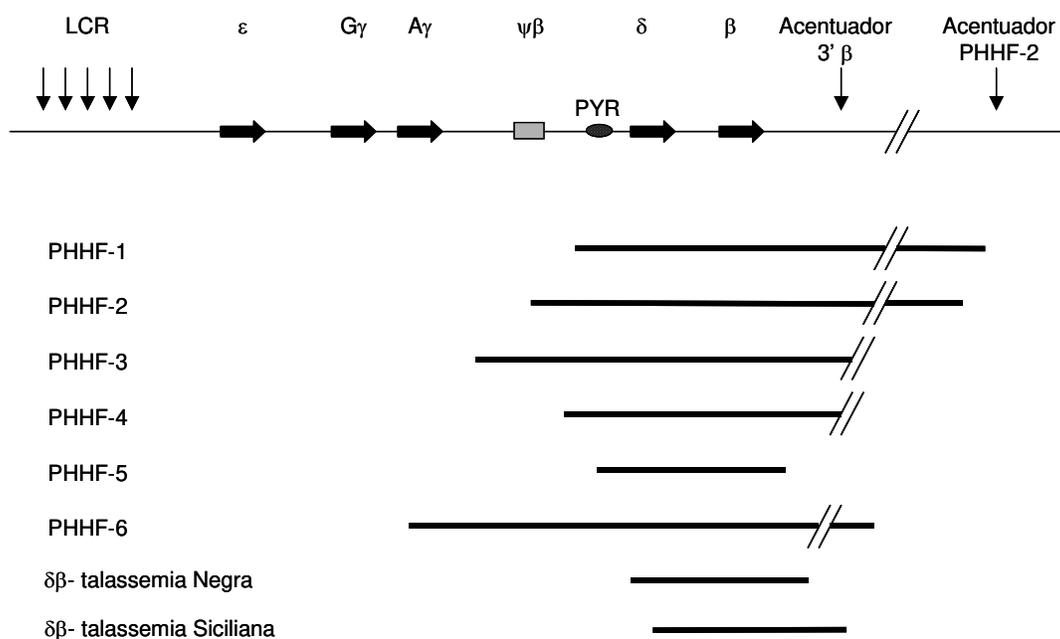


Figura 2- Deficiências gênicas no cluster da beta-globina. Barras representam a extensão das deleções correspondentes.

1.3- Origem das deficiências gênicas

Um dos mecanismos bem caracterizados, que podem originar deficiências gênicas, é o *crossing over* desigual, onde ocorre um pareamento irregular durante a meiose, com permuta entre as cromátides. Este processo é facilitado no complexo da β globina, por que este conjunto de genes é constituído de uma família gênica, com várias regiões de homologia. A hemoglobina Lepore, por exemplo, é originada por este mecanismo e se apresenta híbrida, com partes dos genes δ e β (Weatheral *et al*, 2001).

Um segundo mecanismo proposto surgiu da observação de que algumas deleções possuem extensão similar. É o caso dos tipos de PHHF-1 e PHHF-2, além de outras deleções que englobam todo o conjunto, como as $(\epsilon\gamma^G\gamma^A\delta\beta)^\circ$ talassemias. Neste caso, uma possível explicação seria a perda de alças inteiras de cromatina durante a replicação do DNA (Collins *et al*, 1987). Os pontos de quebra seriam deslocados, de uma deleção para outra, a depender de onde o DNA se encontra ancorado à matriz nuclear (MAR- *Matrix Attachment Region*), e de onde ocorrem as quebras nestes pontos de ancoramento.

Apesar de se identificar ao longo do cluster regiões potenciais para recombinação não homóloga, como repetições *Alu* e *LI*, ao se observar os pontos de quebra das deleções, estas regiões não são encontradas de forma significativa. Parece, assim, que elas não desempenham um papel marcante na origem de deficiências gênicas. Este fato também foi observado na análise de várias deleções no gene da distrofina humana (MacNaughton *et al*, 1998).

1.4- Efeito das deficiências gênicas e avaliação de seqüências regulatórias *in cis*

Três hipóteses principais foram propostas para explicar a relação destas deleções com a ausência de supressão normal dos genes γ na fase adulta. A primeira sugere a competição entre promotores pelo LCR. Em condições normais, esta competição favoreceria os genes δ e β na fase adulta, como resultado da conformação da cromatina e

atuação de fatores de transcrição. Com a remoção destes genes ou de seus promotores, o gene γ estaria livre para interagir com o LCR, elevando, desta forma, os níveis de Hb F (Weatheral *et al.*, 2001). Como as formas de β -talassemia, onde há remoção apenas do gene β , são mais graves, postula-se que deva haver a remoção tanto de δ quanto de β , para que haja um aumento expressivo de Hb F, como ocorre com PHHF e $(\delta\beta)^\circ$ talassemias.

Outra hipótese considera que a expressão aumentada dos genes γ reflete a justaposição de elementos acentuadores (*enhancers*), normalmente encontrados a 3' do cluster, nas proximidades dos genes γ (Tuan *et al.*, 1983; Anagnou *et al.*, 1995; Kansatoni *et al.*, 2003). Diversos estudos, inclusive em camundongos transgênicos, mostram que as seqüências à jusante destas deleções possuem atividade acentuadora, aumentando a expressão gênica (Feingold & Forget, 1989; Anagnou *et al.*, 1995; Arcasoy *et al.*, 1997; Calzolari *et al.*, 1999). Na PHHF do tipo 1, por exemplo, a seqüência a 3' do ponto de quebra possui atividade acentuadora em ensaios de laboratório (Feingold & Forget, 1989), é especificamente hipometilada (Feingold & Forget, 1989) e contém sítio hipersensível à DNase I nos eritrócitos (Elder *et al.*, 1990), além de codificar uma proteína da família de receptores olfativos (Feingold *et al.*, 1999).

Uma terceira hipótese sugere que elementos silenciadores, localizados entre γ^A e δ são removidos pelas deleções (Huisman *et al.*, 1974; Vitale *et al.*, 1994). Neste caso, a região seria removida nas formas delecionais de PHHF, mas mantida na maior parte dos tipos de $(\delta\beta)^\circ$ talassemias, onde os níveis de Hb F são mais baixos.

Neste trabalho, realizamos estudos funcionais com uma região localizada 1-3 kb à montante do gene para δ -globina, aqui denominada R/PYR, que praticamente define e distingue a PHHF-5 da $(\delta\beta)^\circ$ talassemia negra (Figura 2). Este elemento engloba o sítio de ligação do complexo PYR e uma região à montante (R), e tem sido freqüentemente descrito como um potencial candidato a silenciador de γ -globina, ou tendo participação importante no processo de troca na síntese das cadeias de globinas (Vitale *et al.*, 2001; Calzolari *et al.*, 1999; O'Neill *et al.*, 1991). Para investigar o efeito da disposição desta seqüência na regulação da expressão dos genes das globinas, realizamos ensaios de

expressão gênica *in vitro*, em linhagem celular eritróide, com a região R/PYR situada entre duas cópias do gene da β globina, além de um *micro* LCR (seqüência contendo as regiões centrais de quatro sítios hipersensíveis à DNase do LCR) na posição 5' da construção. Estudos demonstram que a cópia do gene β mais próxima do LCR é favorecida na competição, tendo um nível maior de expressão (Peterson & Stamatoyannopoulos, 1993). Neste trabalho, avaliamos se a segunda cópia é favorecida na competição pelo LCR, na presença da região R/PYR, ou se há um efeito inibitório na expressão da primeira cópia.

1.5- Avaliação de Fatores de Transcrição

Apesar das evidências a favor da participação de seqüências *in cis* no fenótipo de PHHF, elas ainda não são conclusivas. Em trabalho recente, Xiang e cols (Xiang *et al*, 2005) desenvolveram uma construção de YAC (*Yeast Artificial Chromosome*) com o locus completo de β -globina humana contendo a deleção responsável pela PHHF-2 e seqüências de acentuadores na região jusante. Inesperadamente, o gene da γ -globina foi totalmente silenciado no estágio adulto em camundongos transgênicos. Este dado sugere que outros mecanismos poderiam estar envolvidos nos níveis aumentados de HbF nesta condição. Os autores especularam que outras regiões localizadas à montante do *cluster* e ausentes na construção poderiam estar envolvidas nesta atividade. Neste trabalho, investigamos o possível envolvimento de fatores de transcrição, construindo Bibliotecas de cDNA subtrativas supressivas (*SSH- Subtractive Supression Hybridization*) (Diatchenko *et al*, 1996) para identificar transcritos diferencialmente expressos em reticulócitos de um sujeito com PHHF-2.

1.6- Importância do estudo para a pesquisa básica e clínica

Os mecanismos gênicos que provocam PHHF ou alternativamente $\delta\beta$ talassemias não são ainda totalmente compreendidos. De fato, a descoberta dos fatores que produzem incremento na síntese de globina gama sem produzir fenótipo talassêmico

são hoje intensamente pesquisados em todo o mundo. O estudo estrutural e funcional dessas deleções fornece importantes informações a respeito dos mecanismos que originam a deleções (Henthorn *et al.*, 1990; Roth *et al.*, 1985; Roth & Wilson, 1986), além do mecanismo molecular envolvido na expressão diferencial dos genes da globina durante o desenvolvimento (Bunn & Forget, 1986; Stamatoyannopoulos & Nienhuis, 1994). Também, pelo fato do nível aumentado de Hb F concomitante a certas hemoglobinopatias, como β -talassemias e anemia da célula falciforme, amenizar estas condições, é de fundamental importância a compreensão dos mecanismos envolvidos na regulação e expressão dos genes da globina γ , para que possam ser desenvolvidos métodos de terapia gênica ou intervenções farmacológicas eficazes para estas desordens (Bunn, 1993; Bunn, 1997; Jane, 1998).

2- OBJETIVOS

2.1- Objetivo geral

Investigar mecanismos genéticos associados com a regulação da síntese de globinas e aos níveis elevados de hemoglobina fetal nos tipos deletoriais de PHHF e em $(\delta\beta)^{\circ}$ talassemias.

2.2- Objetivos específicos

2.2.1- Estudar o efeito da disposição de uma seqüência que engloba o fragmento R e o sítio de ligação a PYR (região a 5' de δ - R/PYR), na regulação da expressão dos genes das globinas, realizando ensaios de expressão gênica *in vitro*, em linhagens de células eritróides.

2.2.2- Identificar genes diferencialmente expressos em PHHF com participação potencial na regulação da globina γ , através da construção de Bibliotecas de cDNA subtrativas supressivas (*SSH – Subtractive Suppression Hybridization*) a partir de amostras obtidas de reticulócitos de sangue periférico de um indivíduo portador da deleção responsável pela PHHF-2 e um sujeito normal.

2.2.3- Estudar a expressão de alguns genes identificados pelo método de SSH utilizando PCR quantitativo em tempo real (*Real-time PCR*) em amostras de reticulócitos de PHHF-2 e $(\delta\beta)^{\circ}$ talassemia Siciliana e em um modelo de diferenciação eritróide *in vitro*.

3- CASUÍSTICA E MÉTODOS

3.1- Casuística

Os dados hematológicos do indivíduo portador da PHHF-2 estão descritos na Tabela 1, bem como os dados de outros indivíduos cujos RNAs foram utilizado para análise de expressão dos genes encontrados nas bibliotecas. RP1 e RP2 correspondem a doadores de sangue e RDB uma paciente heterozigota para $\delta\beta$ -talassemia Siciliana, todos atendidos no Hospital das Clínicas e Hemocentro - Unicamp. Como controle, utilizamos amostras de um indivíduo com índices hematológicos normais. O sangue periférico foi colhido com aprovação do Comitê de Ética local e após consentimento livre e esclarecido dos sujeitos da pesquisa (Parecer CONEP nº 466/2003; Processo nº 25000.013493/2003-20).

Tabela 1- Dados hematológicos dos indivíduos estudados.

Sujeito	GV (106mm³)	Hb (g/dl)	VCM (fl)	HCM (pg)	HbF (%)	HbA₂ (%)	Diagnóstico Molecular
RP1	4.94	13.4	81.9	27.2	28,5	1,7	PHHF-2
RP2	5.03	15.5	91.5	30.8	30.0	2,5	PHHF-2
RDB	4.89	11.4	70.0	23.3	18.9	2.5	$\delta\beta$ -talassemia Siciliana

GV- glóbulos vermelhos; VCM- valor corpuscular médio; HCM- hemoglobina corpuscular média; Hb- hemoglobina; RP1- Sujeito cujo RNA foi utilizado na construção das bibliotecas; RP2 e RDB- Sujeitos cujo RNA foi utilizado nos experimentos de PCR em tempo real.

3.2- Métodos

3.2.1- Dados hematológicos

Os dados hematológicos foram determinados eletronicamente. A quantificação dos níveis de Hb A₂ e Hb F foi realizada utilizando o equipamento *Variant II Hemoglobin Testing System* (Bio-Rad, CA-USA). A eletroforese de hemoglobina foi realizada em acetato de celulose com tampão Tris-EDTA-ácido bórico em pH 8.9. A análise das globinas foi realizada em gel de urea-triton-acrilamida (Betke *et al.*, 1959). A proporção G γ /A γ foi estimada por análise densitométrica (Alter *et al.*, 1980).

3.2.2- Diagnóstico Molecular

A identificação molecular das deleções estudadas neste trabalho foi realizada utilizando uma combinação de PCR *GAP*, dosagem gênica fluorescente e sequenciamento, como previamente descrito (Craig *et al*, 1994; De Andrade *et al*, 2002; De Andrade *et al*, 2003) (ANEXO I). Basicamente, no PCR *GAP*, para cada reação, são utilizados três oligonucleotídeos iniciadores, desenhados para amplificar um produto específico na presença da deleção e um outro, de diferente tamanho, representando o alelo normal, ou controle. Cada deleção possui uma combinação específica de iniciadores (Craig *et al*, 1994). O método de Dosagem gênica consiste em uma modificação da técnica descrita previamente para Distrofia Muscular Duchenne/ Becker (Yau *et al*, 1996), baseada na possibilidade de se identificar diferenças quantitativas na amplificação de amostras por PCR fluorescente, em presença de deleções ou duplicações, onde o produto final de amplificação é diretamente proporcional à quantidade inicial de cópias da seqüência alvo. Desenhamos reações de *multiplexes* com conjuntos de pares de iniciadores marcados com fluorescência, que cobriam a extensão do *cluster* da β -globina, além do gene da beta-actina como controle diplóide (De Andrade *et al*, 2003) (ANEXO I). Os dados foram analisados em sequenciador automático de capilaridade (*310 Genetic Analyser* - ABI Prism - Applied Biosystems) e em software apropriado (*GeneScan*-Applied Biosystems). Os dados da área de cada pico foram transferidos para uma planilha de Excel, onde foi calculado o Coeficiente de Dosagem Gênica [CDG= (área do pico x da amostra/área do pico y da amostra)/(área do pico x do controle/área do pico y do controle)], tendo valores teóricos de 1 para região diplóide normal, 0,5 ou 2 para deleções e 1,5 ou 0,75 (a depender da posição dos valores na fração) para duplicações. Para amplificarmos o ponto de quebra das deleções, realizamos uma reação contendo enzima para fragmentos longos (*ELONGASE* - Life Technologies), e iniciadores que pudessem flanquear a deleção em estudo. Este fragmento foi sequenciado em equipamento automático 377 (Perkin Elmer-ABI Prism). Para confirmação do ponto de quebra, foi realizado alinhamento da seqüência do fragmento com a seqüência do *cluster* da β -globina (GI-455025/ GENE BANK), e de outras seqüências de pontos de quebra já caracterizados, utilizando os programas “*BLAST 2 sequence*” e “*Blastn*”, disponíveis em

<http://www.ncbi.nlm.nih.gov:80>, do “Nacional Center for Biothechnology Information” (ANEXO I).

3.2.3- Análise funcional da região R/PYR

3.2.3.1- Obtenção da construção

A construção $\mu\text{LCR}-\beta^m\text{-R/PYR}-\beta$ (Figura 3) foi obtida da seguinte maneira: um fragmento de PCR de 3.6 Kb contendo a região R/PYR foi clonado em vetor PUC 18 defosforilado e digerido com Sma I, seguindo protocolo do *Kit Sureclone* (Amersham). A segunda cópia do gene β , de 4 Kb, foi obtida por PCR e também clonada em PUC 18. Em seguida, o fragmento contendo o gene β , obtido a partir deste clone por digestão com as enzimas Hinc II e Kpn I, foi transferido para o vetor PUC 18-R/PYR digerido com Hpa I e Kpn I. O fragmento contendo R/PYR ligado à montante de β (R/PYR- β) foi obtido a partir deste clone com as enzimas Hind III e Kpn I e clonado no vetor *pBluescript* contendo um micro LCR ligado a um gene β marcado ($\mu\text{LCR}-\beta^m$ - Peterson *et al*, 1993), também digerido com Hind III e Kpn I. Como controle, foi utilizada a construção $\mu\text{LCR}-\beta^m-\beta$ (Peterson *et al*, 1993) (Figura 3).

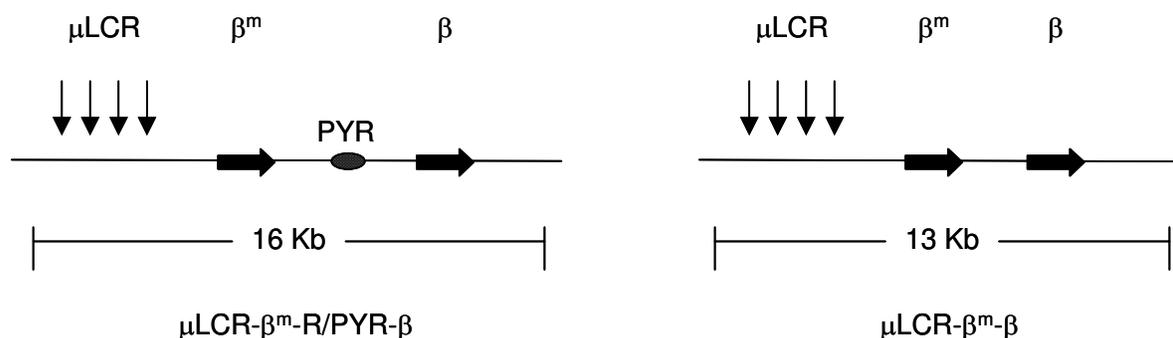


Figura 3- Esquema das construções utilizadas para o estudo da região R/PYR.

As reações de amplificação por PCR foram realizadas com enzima *Platinum Taq High Fidelity* (Invitrogen), para evitar erros durante o processo de duplicação das fitas de DNA. Para a reação, foram desenhados em programa *Generunner* os iniciadores R/PYR sense (5'-CGTACAAGCTTACTGCAACATCTGCCTC-3'), que contém um sítio para enzima Hind III (em negrito), R/PYR antisense (5'-TGACTCAGAAACCCTGTTCTCCTCTG-3'), β sense (5'-ACTGCATTAAGAGGTCTCTAG-3') e β antisense (5'-AACTACAGTCACGATG-3'). As reações consistiram em 2mM de MgSO₄, tampão High Fidelity 1X, 0.2mM de cada dNTP, 10 pM de cada iniciador, 1 U da enzima, 100 ng de DNA (cosmídeo pH C 79 contendo cluster da beta globina), para um volume final de 50 μ L. Os parâmetros de amplificação foram desnaturação inicial a 94°C por 1min, seguida por 32 ciclos de desnaturação a 94°C por 30 seg, anelamento a 60°C (para R/PYR) ou 58°C (para gene β) por 40 seg e extensão a 68°C por 3min e 40 seg (para R/PYR) ou 4 min (para gene β).

As reações de digestão foram realizadas segundo as recomendações específicas para cada enzima (New England Biolabs). Os fragmentos das digestões foram obtidos por extração do DNA de gel de agarose *low melting* (SeaPlaque – FMC Bioproducts) 1 % em TAE 0.5 X, utilizando o *Kit Qiaquick gel Extraction Kit* (QiaGen).

As clonagens foram realizadas utilizando T4 DNA ligase (Promega) e linhagens de *E. coli* DH5 α MCR, por transformação química (choque térmico). Os plasmídeos foram extraídos por lise alcalina utilizando soluções preparadas no laboratório ou com Kits comerciais de Miniprep (Qiagen, Gibco). Os clones foram selecionados por padrão de restrição. A Figura 4 mostra a seleção dos clones para a construção intermediária R/PYR- β (PUC 18), e as Figuras 5 e 6 para a construção final μ LCR- β^m -R/PYR- β .

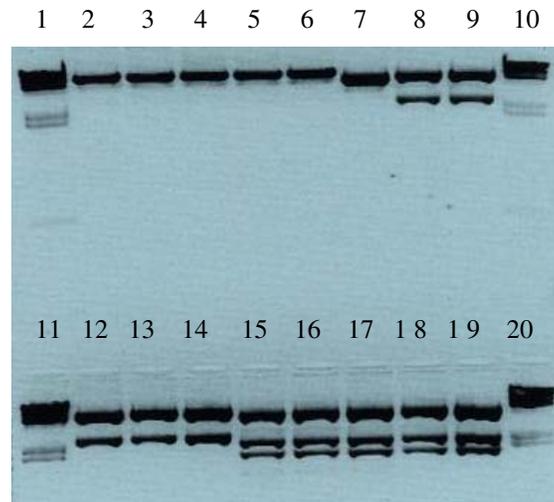


Figura 4- Padrão de restrição confirmando obtenção do clone intermediário R/PYR- β . 1, 10, 11 e 20- Marcador de peso molecular λ Hind III; 2-6- DNA plasmidial digerido com Hind III ; 7- DNA plasmidial de β PUC18 digerido com Hind III; 8-14- DNA plasmidial digerido com Hind III + Kpn I; 15-19- DNA plasmidial digerido com EcoRI.

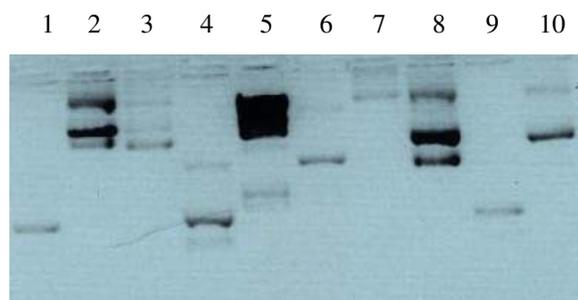


Figura 5- Seleção do clone 7 μ LCR- β^m -R/PYR- β . 1-4; 6-10. Plasmídios digeridos com Hind III; 5- Marcador de peso molecular λ Hind III. Apenas a amostra 7 apresenta o padrão específico (plasmídio linearizado com 16 Kb).

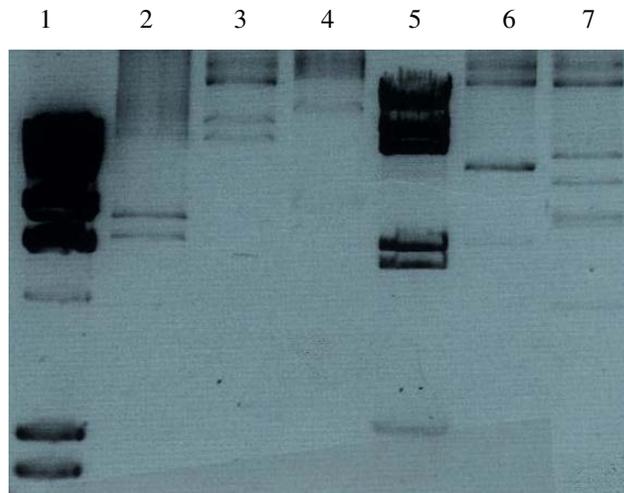


Figura 6- Padrão de restrição do clone 7 com diferentes enzimas de restrição. 1 e 5- Marcador de peso molecular λ Hind III; 2- Plasmídeo digerido com Hind III + Kpn I; 3- Sma I + Hind III; 4- Sma I + Kpn I ; 6- Eco RI; 7- Xba I. Todas as reações apresentaram os padrões esperados para a construção. (As amostras 1 e 2 foram submetidas a eletroforese em momento diferente das demais).

3.2.3.2- Transfecção em Linhagens Eritróides: Para avaliação funcional da região R/PYR e seu possível papel da regulação da transcrição de globinas, a construção μ LCR- β^m -R/PYR- β foi transfectada de maneira transiente e estável em células MEL (*Murine Erytroleukemia*), linhagem eritróide que possui um ambiente proteico favorável a expressão de genes de globina adultos. Esta linhagem celular também expressa o complexo PYR (O'Neill *et al*, 1991). Como controle, foi utilizada a construção μ LCR- β^m - β (Peterson *et al*, 1993). Nos estudos com transfecção transiente, as amostras foram avaliadas 48 horas após transfecção, para expressão dos transcritos de β globina.

Na transfecção estável, as amostras são inseridas na célula como moléculas lineares, permitindo a integração das construções no DNA da célula transfectada. Diferentemente das transfecções transientes, onde os plasmídios são inseridos em

conformação circular, a transfecção estável permite a análise da região regulatória numa disposição semelhante à naturalmente encontrada. Após Midipreparação (Kit Qiagen para Midiprep), as amostras foram quantificadas em fluorômetro e digeridas com as enzimas SmaI e KpnI para obtenção dos fragmentos contendo os genes e região de estudo. Além disto, foi utilizado vetor contendo o gene de higromicina como marcador de seleção dos clones. Este vetor (pGK-Hyg) foi digerido com as enzimas ClaI e HindIII para obtenção do fragmento desejado. Os fragmentos de digestão foram então purificados do gel utilizando kit de extração de gel da Qiagen (QiaexII) e submetidas a eletroporação. 1×10^7 células foram centrifugadas a 200g por 5 minutos em temperatura ambiente, lavadas com PBS 1X, ressuspensas em 0.8ml de HEPES/Sucrose gelado (272 mM Sucrose, 8 mM Hepes, esterilizado com filtro) e mantidas em gelo. Aproximadamente 1 μ g de DNA do gene da higromicina foi misturado a 1 μ g de DNA da amostra μ LCR- β^m -R/PYR- β e, separadamente, a aproximadamente 1 μ g de DNA do controle μ LCR- β^m - β , para co-transfecção. As amostras foram adicionadas a cuvetas de 4mm (Biorad) e submetidas a eletroporação com 0.5 K e 25 μ F, em aparelho *GenePulser II* (Biorad).

Após eletroporação, 20 μ L foram submetidos à contagem e teste de viabilidade, utilizando azul de Tripán e hemocitômetro e o restante do conteúdo de cada cuveta foi transferido para placa contendo 3mL de meio RPMI 1640 contendo L-glutamina, aminoácidos não essenciais, piruvato de sódio e 10% de Soro Bovino Fetal inativado. As células foram incubadas por 2 dias em estufa a 37°C e 5% de CO₂, centrifugadas e adicionadas em meio fresco, contendo o antibiótico para seleção (50 μ g de higromicina por mL de meio). Após crescimento das células na presença do antibiótico, foi feita diluição para 1 célula por 50 μ L de meio, e distribuídas em placas de 96 poços, 50 μ L cada poço. Neste caso, cada “clone” poderia conter as construções transfectadas inseridas no DNA de maneiras diferentes, o que poderia gerar expressões diferenciadas. Foi adicionado 50 μ L por poço de meio fresco, duas vezes por semana, e após duas semanas foi feita checagem por PCR para detecção dos fragmentos transfectados.

Para extração do DNA das células foi utilizado tampão de lise (100mM NaCl, 10mM Tris-Cl pH8.0, 25mM EDTA pH8.0, SDS 0.5% e Proteinase K 0.1mg/mL), extração com fenol:clorofórmio e precipitação com etanol. Para a reação de PCR foram

utilizados iniciadores específicos para o gene da β -globina (Sense: 5'-CTGTGTTCACTAGCAACCTC-3'; Antisense: 5'-AAGAACCTCTGGGTCCAAG-3') Aquelas colônias de células que apresentaram-se positivas para β -globina (Figura 7) foram submetidas à análise de expressão gênica por RT PCR e digestão.

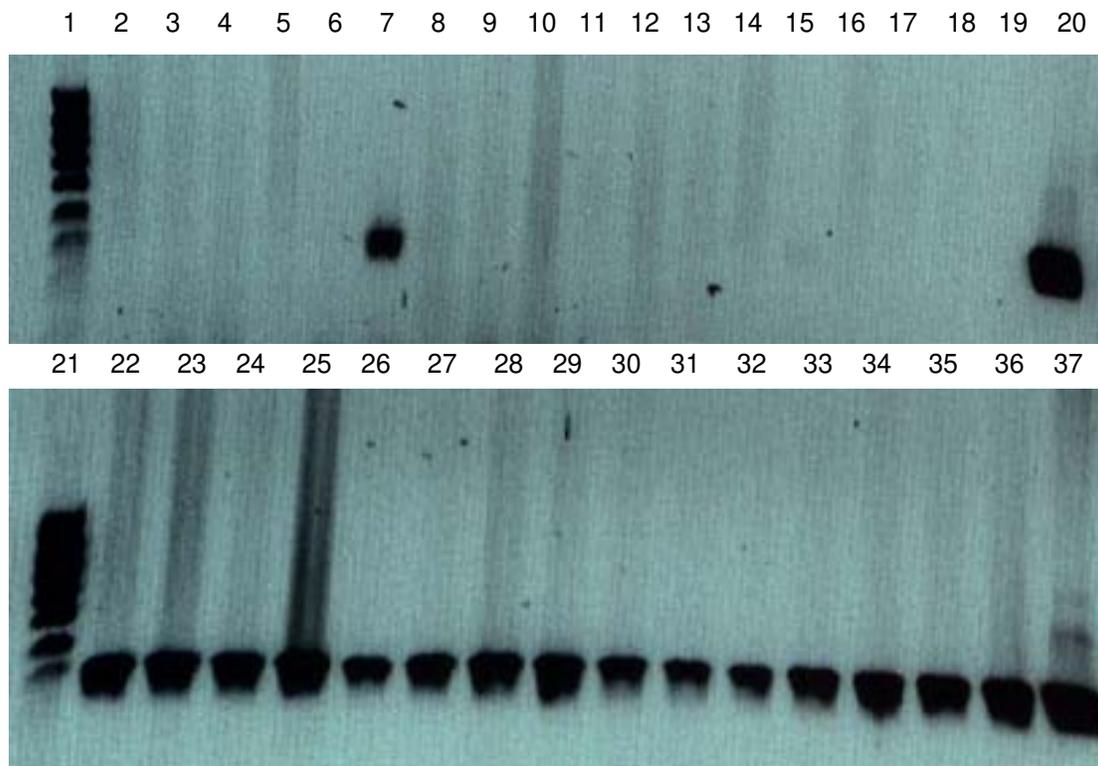


Figura 7- PCR diagnóstico para presença das construções em célula MEL após transfecção estável. 1 e 21- Marcador de peso molecular Hyperladder IV (Bioline); 20 e 37- Controle positivo; 2-19 PCR de DNA genômico extraído de células MEL transfectadas com a construção controle μ LCR- β^m - β ; 22-36- PCR de DNA genômico extraído de células MEL transfectadas com a construção μ LCR- β^m -R/PYR- β .

3.2.3.3- RT-PCR e digestão: O RNA total foi extraído de 5×10^7 células utilizando kit para extração de RNA (*RNAgents Total RNA Isolation System* - Promega) e quantificado utilizando espectrofotômetro MBA (*Molecular Biology Assistant* - Perkin Elmer). Em seguida, 2 μg de cada amostra foram submetidos a transcrição reversa, para síntese de cDNA, utilizando *SuperScript II* (Invitrogen) nas condições sugeridas pelo fabricante. A qualidade do cDNA foi avaliada por amplificação de α -globina murina Sense: 5'-GATTCTGACAGACTCAGGAAGAAAC-3'; Antisense: 3'-CCTTCCAGGGCTTCAGCTCCATAT-3'). As reações de RT-PCR para β globina foram realizadas com iniciadores específicos para RNA mensageiro, cobrindo na porção 5' a região não traduzida (5'-ACATTTGCTTCTGACACAACACTG-3'), e na porção 3' a junção entre exon III-exonII (5'-AGGAGCCTGAAGTTCTCAG-3'). As reações consistiram em 2mM de MgCl_2 , tampão de PCR 1X, 0.2mM de cada dNTP, 10 pM de cada iniciador, 1 U da enzima (Biolase), 2 μL do cDNA, para um volume final de 25 μL . Os parâmetros de amplificação foram desnaturação inicial a 95°C por 5 min, seguida por 40 ciclos de desnaturação a 95°C por 30 seg, anelamento a 68°C por 30 seg e extensão a 72°C por 30 seg, com uma fase final de extensão a 72°C por 10 min. Os produtos de PCR foram digeridos com as enzimas de restrição NcoI ou ClaI, e avaliados em gel de agarose.

3.2.4- Identificação de genes diferencialmente expressos em PHHF-2

3.2.4.1- Biblioteca de cDNA Subtrativa Supressiva (*Subtractive Suppression Hybridization-SSH*): A partir de RNA extraído de reticulócitos de sangue periférico de um sujeito portador da PHHF-2 e um indivíduo controle, utilizamos o método de SSH (*Subtractive Suppression Hybridization-SSH*) (Diatchenko *et al*, 1996) para obtenção de perfis transcricionais que pudessem identificar genes diferencialmente expressos naquela condição, e que pudessem ser correlacionados com a expressão aumentada de γ -globina em PHHF. O método consiste na construção de bibliotecas de cDNA que contenham apenas aqueles genes com expressão diferente entre as duas condições estudadas. Baseia-se nos seguintes princípios: 1- Hibridização subtrativa- as duas

populações de cDNA estudadas (controle e paciente, por exemplo) são misturadas, após digestão com uma enzima de restrição, sob condições que permitam a hibridização de moléculas complementares, correspondentes a transcritos do mesmo gene. Uma destas populações (denominada *Tester*) é ligada a adaptadores contendo seqüências para iniciadores de PCR, e misturada com excesso da outra população (*Driver*), sem adaptadores. Assim, após a hibridização (geralmente duas hibridizações consecutivas), apenas aqueles transcritos que não sofreram hibridização (com adaptadores) são amplificados por PCR, enquanto os transcritos comuns (sem adaptadores) não sofrem amplificação.

2- PCR supressivo- antes da primeira hibridização, a população *Tester* é separada em duas alíquotas e ligadas a adaptadores diferentes. Após a primeira e segunda hibridizações, as moléculas de cDNA dupla fita, contendo os mesmos adaptadores nas duas extremidades sofrem um efeito denominado PCR supressivo na reação (Diatchenko *et al*, 1996). Estas moléculas correspondem aos transcritos mais abundantes, já que estes tem maior probabilidade de hibridização dentro da mesma população de cDNA. Como adaptadores idênticos contém seqüências complementares invertidas, e a ligação intramolecular ocorre mais facilmente que entre moléculas (neste caso entre os iniciadores e o molde ou adaptadores), a reação de amplificação não ocorre a partir destes cDNAs. Apenas aquelas moléculas de cDNA dupla fita, com adaptadores diferentes nas extremidades sofrem amplificação exponencial e podem ser detectadas. Esta estratégia permite a normalização da biblioteca, evitando que os transcritos de alta expressão tenham amplificação preferencial, e prejudiquem a detecção dos transcritos menos expressos ou raros.

3- Clonagem e sequenciamento- após amplificação, os transcritos normalizados e diferencialmente expressos podem ser clonados e submetidos a sequenciamento para identificação e análise dos genes correspondentes. Esses procedimentos devem ser realizados para cada um dos perfis desejados. Assim, obtivemos duas bibliotecas supressivas subtrativas, representando os perfis transcripcionais (genes mais expressos) em indivíduos com PHHF-2 e sujeitos controles normais.

3.2.4.2- Extração de RNA de reticulócitos

Sangue periférico foi colhido com aprovação do Comitê de Ética local e após consentimento livre e esclarecido dos sujeitos da pesquisa. Os dados hematológicos do indivíduo portador da PHHF-2 estão descritos na Tabela 1, bem como os dados de outros indivíduos cujo RNA foi utilizado para análise de expressão dos genes encontrados nas bibliotecas. Cada 5mL de sangue colhido com anticoagulante EDTA foi transferido para tubo Falcon de 15mL. Após centrifugação em 3.000 rpm (centrífuga Sorvall, rotor SLA600TC), 10 minutos a 4°C, o plasma foi descartado e foi adicionado 5x o volume de solução de lise de hemácias (0.144M NH₄Cl; 0.01M NH₄HCO₃). As amostras foram incubadas 30 minutos em gelo, com agitação vigorosa a cada 5 minutos. Após centrifugação 3.000 rpm 15 minutos a 4°C, o sobrenadante contendo hemácias lisadas e reticulócitos foi transferido para novo tubo, homogeneizado com 1/10 de volume de uma solução de Sucrose/KCl (1.5M C₁₂H₂₂O₁₁; 0.15M KCl, filtrada ou preparada com H₂O tratada com DEPC) e centrifugado novamente 5.000 rpm por 20 minutos a 4°C. O sobrenadante contendo agora apenas reticulócitos foi transferido para novo tubo, tratado com Ácido acético 10% e centrifugado 5.000 rpm 20 minutos a 4°C. O precipitado (*pellet*) foi então ressuspendido e homogeneizado em 1mL de Trizol[®] (Invitrogen). O protocolo de extração de RNA foi seguido segundo recomendações do fabricante, sendo o precipitado final (*pellet*) ressuspendido em 5-10µL de H₂O livre de RNase. As amostras foram quantificadas em espectrofotômetro e avaliadas quanto à integridade em gel desnaturante de agarose.

3.2.4.3- Síntese e amplificação de cDNA para construção das bibliotecas

O cDNA utilizado para construções das bibliotecas subtrativas foi obtido por amplificação de moléculas de cDNA de tamanho completo. O procedimento se inicia com a síntese da primeira fita de cDNA utilizando um *oligo(dT)* modificado (3'*BD SMART CDS IIA*), complementar à cauda poli-A do RNAm. Quando a transcriptase reversa alcança a extremidade 5' da molécula, a atividade transferase terminal da enzima adiciona preferencialmente deoxicitidinas na porção 3' do cDNA. Um outro oligonucleotídeo

adicionado à reação (*SMART IIA*), que possui um *oligo(G)*, se hibridiza à seqüência de citidinas, criando um molde estendido. A transcriptase reversa muda então de molde e continua replicando até o final do oligo (Zhu *et al*, 2001). O cDNA fita simples resultante contém a seqüência completa 5' do RNAm, bem como seqüências complementares aos oligonucleotídios *BD SMART*, que podem ser utilizados com pontos de ancoramento para iniciadores em reações de RT-PCR, amplificando desta maneira as seqüências completas de cDNA dos transcritos da amostra em estudo. O protocolo que se segue é similar ao descrito no protocolo do Kit *BD SMART PCR cDNA Synhtesis Kit* (BD Biosciences), com adapatações. 1-2µg de RNA de reticulócitos (8µL) foi tratado primeiramente com DNase (Invitrogen) 15 minutos a temperatura ambiente, numa reação final de 10µL. Síntese da primeira fita: 2µL (20µM) de cada iniciador (3'BD SMART CDS IIA: 5'-AAGCAGTGGTATCAACGCAGAGTACT₍₃₀₎VN-3', onde V=A, G ou C E N=qualquer uma das bases; SMART IIA: 5'-AAGCAGTGGTATCAACGCAGAGTACGCGGG-3') foi adicionado ao RNA tratado e as amostras foram incubadas 70°C por 2 minutos. Foi adicionado então um mix com 2.5µL de tampão RT 10X, 1.5µL MgCl₂ 25mM, 2.5µL dNTPs 10mM, 0.5µL DTT 100µM, 2.0µL de "RNase OUT" e 2.0µL de *Superscript III* (Invitrogen). As amostras foram incubadas 42°C por 1h, foi adicionado 20µL de H₂O e incubadas 7 minutos a 72°C para inativação da enzima. Síntese da segunda fita/ amplificação do cDNA: as reações de amplificação foram padronizadas para reações de PCR utilizando enzima de alta fidelidade (*Platinum High Fidelity Polymerase* - Invitrogen), nas seguintes condições: tampão de PCR HiFi 1X, 2mM de MgSO₄, 10µM de cada dNTP, 20 µM de iniciador (5' PCR Primer IIA: 5'-AAGCAGTGGTATCAACGCAGAGT-3'), 0.5µL da enzima, 2µL do cDNA fita simples, para um volume final de 50µL. Os parâmetros de amplificação foram desnaturação inicial a 95°C por 1 min, seguida por 33 ciclos de desnaturação a 95°C por 10 seg, anelamento a 60°C por 15 seg e extensão a 68°C por 7 min. O número de ciclos ideal foi determinado com a realização de testes de amplificação das amostras em diferentes ciclos (Figuras 8 e 9). O ciclo escolhido deve estar antes do ponto de saturação da reação. Para nossos experimentos selecionamos 33 ciclos. As bandas de maior intensidade nos géis representam aqueles transcritos mais freqüentes.

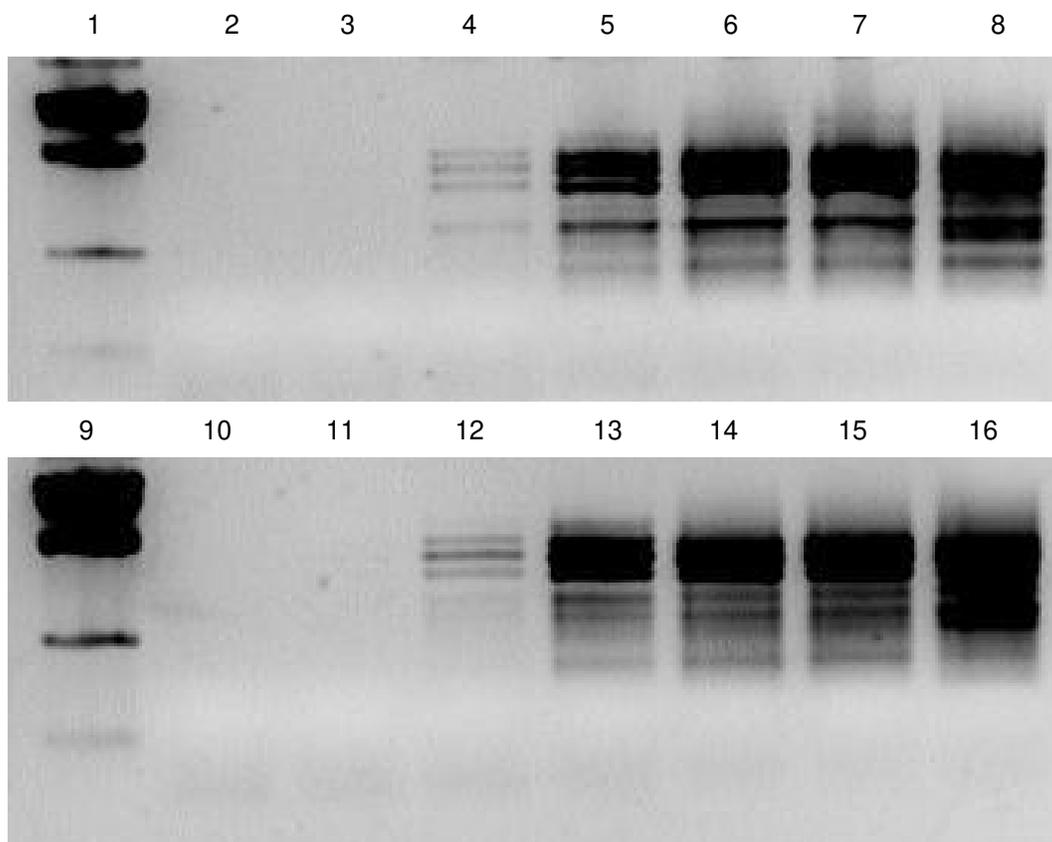


Figura 8- Padronização do número de ciclos de PCR para amplificação do cDNA a ser utilizado na construção da biblioteca subtrativa. 1 e 9- Marcador de peso molecular λ hindIII.; 2- cDNA Reticulócito Normal (RN) 20 ciclos; 3- RN 25 ciclos; 4- RN 30 ciclos; 5- RN 35 ciclos; 6- RN 40 ciclos; 7- RN 45 ciclos; 8- RN 50 ciclos; 10- cDNA Reticulócito PHHF-2 (RP) 20 ciclos; 11 RP 25 ciclos; 12- RP 30 ciclos; 13- RP 35 ciclos; 14- RP 40 ciclos; 15- RP 45 ciclos; 16- RP 50 ciclos.

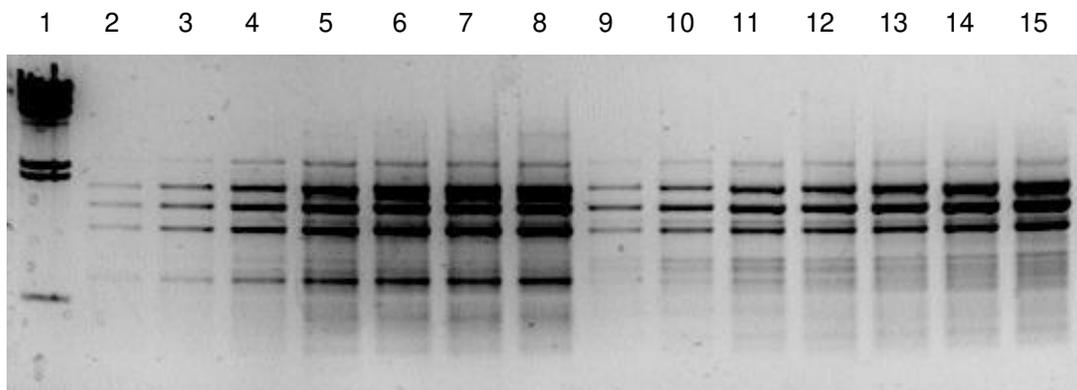


Figura 9- Padronização do número de ciclos de PCR para amplificação do cDNA a ser utilizado na construção da biblioteca subtrativa. 1- Marcador de peso molecular λ hindIII.; 2- cDNA Reticulócito Normal (RN) 30 ciclos; 3- RN 32 ciclos; 4- RN 34 ciclos; 5- RN 36 ciclos; 6- RN 38 ciclos; 7- RN 40 ciclos; 8- RN 42 ciclos; 9- cDNA Reticulócito PHHF-2 (RP) 30 ciclos; 10- RP 32 ciclos; 11- RP 34 ciclos; 12- RP 36 ciclos; 13- RP 38 ciclos; 14- RP 40 ciclos; 15- RP 42 ciclos.

3.2.4.4- Purificação do cDNA e Digestão com RsaI

As reações de PCR contendo o cDNA dupla fita das amostras foram purificadas e concentradas com colunas GFX (Amershan). Foram purificadas duas reações de 50 μ L por coluna, em um total de 3 colunas por amostra, recuperando um volume final de 120 μ L (Figura 10). Destes, 115 μ L foram digeridos com 13 μ L de enzima RsaI (5U/ μ L) e tampão adequado 1X, num volume final de 130 μ L, a 37°C 16-20hs. Uma alíquota de 5 μ L foi avaliada em gel de agarose para checar a eficiência da digestão (Figura 11). As amostras de digestão foram homogeneizadas com igual volume de uma mistura fenol:clorofórmio:álcool isoamílico e centrifugadas a velocidade máxima por 10 minutos. O sobrenadante foi transferido para novo tubo e o DNA precipitado com 2.0 M de acetato de sódio e 2 volumes de etanol absoluto. Centrifugado a velocidade máxima por 20 minutos e lavado com etanol 75%. Centrifugado novamente em velocidade máxima 5 minutos, deixado secar e ressuscitado em 7 μ L de H₂O. Foi feita checagem com 1 μ L em gel de agarose (Figura 12).

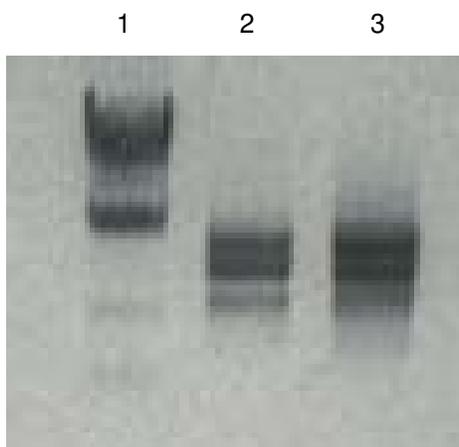


Figura 10- cDNA purificado em colunas GFX. 1- Marcador de peso molecular λ hindIII.;
2- RN; 3- RP.

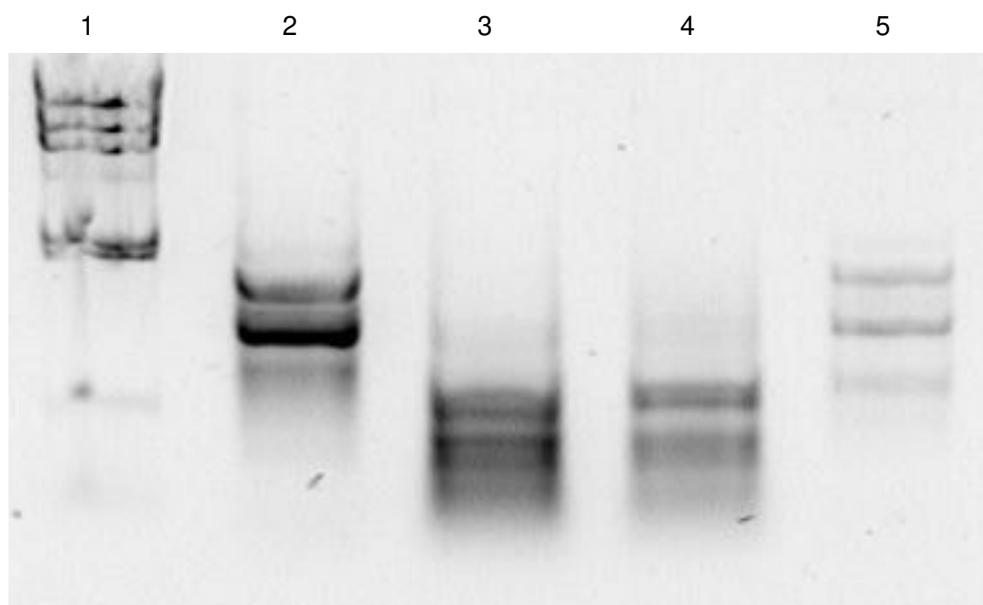


Figura 11- Avaliação da digestão do cDNA com RsaI. 1- Marcador de peso molecular
 λ hindIII.; 2- RP não digerido; 3- RP digerido; 4- RN digerido;
5- RN não digerido.

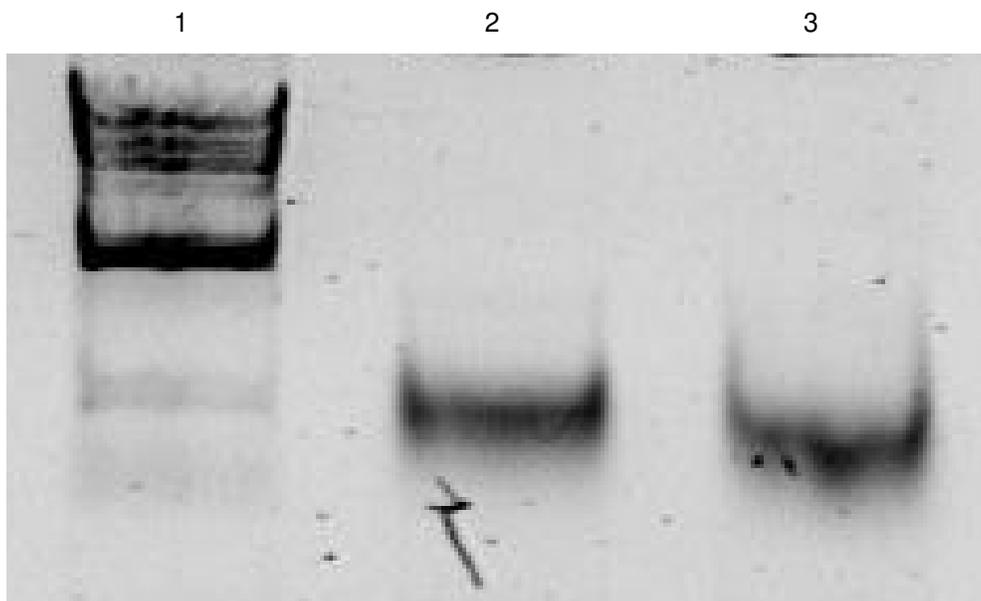


Figura 12- cDNA digerido com RsaI após purificação. 1- Marcador de peso molecular λ hindIII.; 2- RN purificado; 3- RP purificado.

3.2.4.5- Ligação dos adaptadores

Os reagentes e procedimentos que se seguem foram feitos segundo *Clontech PCR-Select cDNA Subtraction Kit* (Clontech). 1 μ L do cDNA *Tester* digerido com RsaI foi diluído em 5 μ L de H₂O. Desta diluição, 2 μ L foram misturados com 2 μ L de Adaptador 1 ou 2R (10 μ M), 2 μ L de tampão de ligação 5X, 1 μ L de T4 Ligase e 3 μ L de H₂O. 2 μ L da ligação com Adaptador 1 e 2 μ L da ligação com Adaptador 2R foram misturados e reservados como teste não subtraído para análise posterior (Figuras 13 e 14). As ligações foram incubadas 16-20hs 16°C, posteriormente a 72°C para inativar a ligase e estocadas a -20°C.

3.2.4.6- Hibridizações

Na 1ª hibridização, 1.5µL de cada ligação de cDNA *Tester* aos adaptadores foi misturado em separado com 1.5µL de cDNA *Driver* digerido com *RsaI*, mais 1µL de tampão de hibridização 4X. A mistura foi incubada 98°C por 1.5 minuto e posteriormente 68°C por 8 horas. Na 2ª hibridização, 1µL da mistura de 2µL de H₂O, 1µL de tampão de hibridização 4X e 1µL de cDNA *Driver* digerido com *RsaI* foi pré-aquecida a 98°C por 1.5 minuto e misturada com as amostras de cDNA *Tester* da 1ª hibridização, como segue: com a pipeta marcando 15µL e tubos da 1ª hibridização a 68°C, foi pipetado todo o volume do tubo contendo o adaptador 2R, puxado um pouco de ar na ponteira e pipetado o novo *Driver* pré-aquecido a 98°C. Estas amostras foram então transferidas para o tubo da 1ª hibridização contendo o adaptador 1 (68°C) e misturado. Este procedimento permite que o conteúdo dos 3 tubos (cDNA *Tester* - adaptador 1 após 1ª hibridização; cDNA *Tester* - adaptador 2R após 1ª hibridização e novo *Driver*) sejam misturados ao mesmo tempo. As amostras foram incubadas 68°C 16-20hs. 200µL do tampão de diluição foi adicionado às amostras, que foram incubadas novamente a 68°C por 7 minutos e estocadas a -20°C.

3.2.4.7- Amplificações por PCR

Nesta etapa, os cDNAs diferencialmente expressos são amplificados seletivamente, com iniciadores com seqüência homóloga aos adaptadores ligados às moléculas. Na primeira reação de PCR, somente cDNA de dupla fita com diferentes adaptadores ligados às extremidades são amplificados exponencialmente. Na segunda amplificação, iniciadores mais internos à seqüência do molde (*Nested*) são utilizados para reduzir ruído de fundo (*background*) e enriquecer aquelas seqüências diferencialmente expressas. Primeiramente, as amostras são incubadas a 75°C por 5', para preencher as extremidades dos adaptadores na molécula dupla fita, seguindo imediatamente ao programa de PCR que se deu nas seguintes condições: desnaturação a 94°C por 30 segundos, seguida de 28 ciclos de desnaturação a 94°C 30 segundos, anelamento a 65°C por 30 segundos e

extensão a 72°C por 1.5 minuto. 2µL do cDNA resultante da 2ª hibridização foi adicionado a uma reação contendo tampão de PCR 1X, 2mM de MgCl₂, 10mM de cada dNTP, 10µM de iniciador *PCR Primer 1* (5'-CTAATACGACTCACTATAGGGC-3'), 1U da enzima Taq polimerase (Invitrogen), para um volume final de 25µL. 8µL da reação foram verificados em gel de agarose 2% (Figura 13). Na segunda amplificação, 1µL do cDNA gerado na primeira reação de PCR foi adicionado a uma reação contendo tampão de PCR 1X, 2mM de MgCl₂, 10mM de cada dNTP, 10µM de iniciador *Nested PCR Primer 1* (5'-TCGAGCGGCCCGCCCGGGCAGGT-3'), 10µM de iniciador *Nested PCR Primer 2R* (5'-AGCGTGGTCGCGGCCGAGGT-3'), 1U da enzima Taq polimerase (Invitrogen), para um volume final de 25µL. A amplificação ocorreu com desnaturação inicial a 94°C por 30 segundos, seguida de 13 ciclos de desnaturação a 94°C 30 segundos, anelamento a 68°C por 30 segundos e extensão a 72°C por 1.5 minuto. 8µL da reação foram verificados em gel de agarose 2% (Figura 14). Para estas duas séries de amplificações também foram avaliados o cDNA não subtraído (item 3.2.4.5) em diluição de 1µL para 1mL de H₂O.

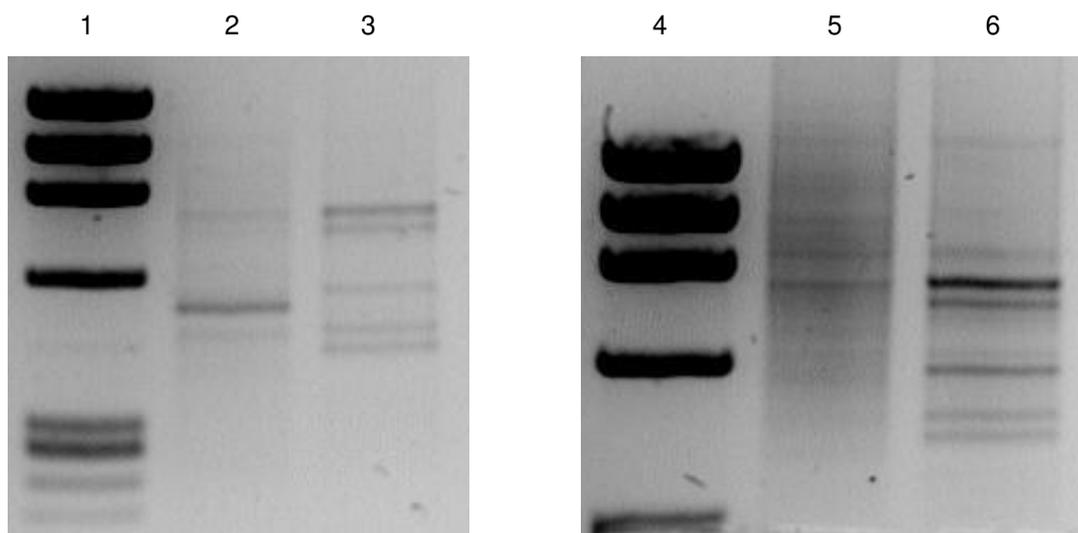


Figura 13- 1ª amplificação do cDNA subtraído. 1 e 4- Marcador de peso molecular ϕ X 174 haeIII; 2- cDNA RN após subtração; 3- cDNA RN não subtraído; 5- cDNA RP após subtração; 6- cDNA RP não subtraído.

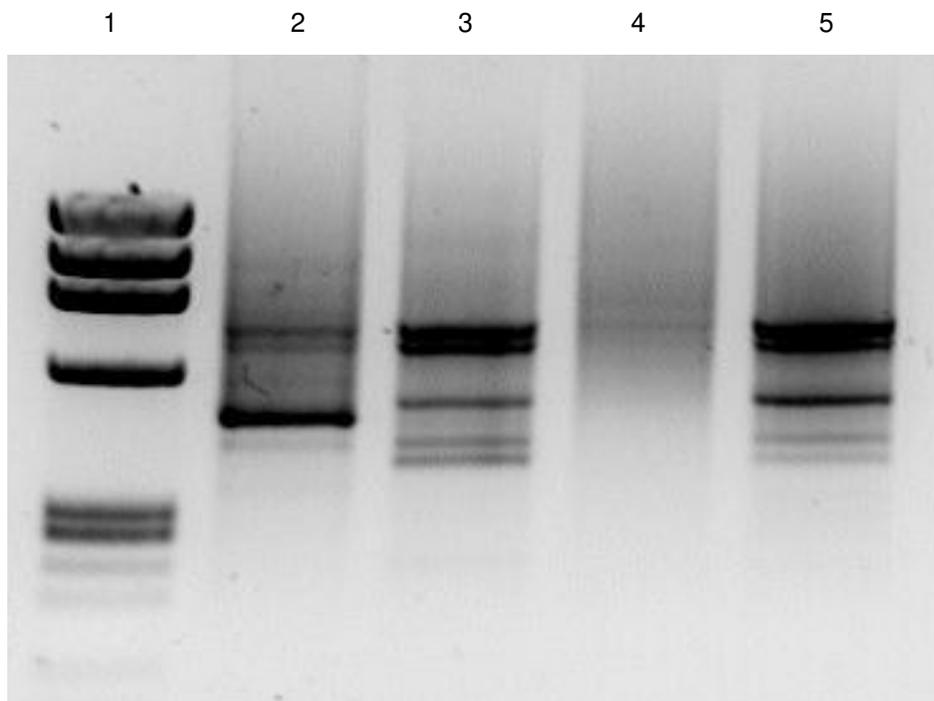


Figura 14- 2ª amplificação do cDNA subtraído (PCR *Nested*). 1- Marcador de peso molecular ϕ X 174 haeIII; 2- cDNA RN após subtração; 3- cDNA RN não subtraído; 4- cDNA RP após subtração; 5- cDNA RP não subtraído.

3.2.4.8- Clonagem e sequenciamento

As amostras de cDNA resultantes da segunda amplificação foram clonadas em vetor para clonagem de PCR pGEMT (Promega). 3 μ L do cDNA foi adicionado a 5 μ L de tampão ligase 2X, 1 μ L de vetor pGEMT e 1 μ L de T4 ligase e a reação incubada a 4°C 16-20hs. 3-5 μ L da ligação foi transformado em 100 μ L de células DH5 α competentes por choque térmico (20 minutos em gelo, 1.5 minutos a 42°C e 2 minutos em gelo). As células foram transferidas para 1mL de meio LB ou SOC sem ampicilina, incubadas sob agitação a 37°C, e plaqueadas em meio LB sólido com marcador ampicilina e sistema de galactosidade para seleção de colônias com inserto. As colônias positivas para inserto (colônias de cor branca) foram inoculadas em placas de 96 poços contendo 120 μ L de meio LB com ampicilina por poço. Estas placas foram incubadas sob agitação a 120rpm a

37°C 12-16 hs. 2µL do inóculo foi submetido a reação de PCR contendo iniciadores padrões M13 para amplificação dos insertos (Figura 15). 1µL do PCR foi seqüenciado com 5pmol de iniciador sense do PCR (M13 sense), utilizando pré-mix (Amersham - Kit para MegaBace), em reações de 14µL por poço. As condições foram desnaturação 94°C por 3 minutos, seguidos de 35 ciclos de desnaturação 94°C 20 segundos, 58°C 15 segundos e 60°C 1 minuto. O seqüenciamento foi purificado como segue: 2µL de acetato de amônio 7.5M e 50µL de etanol absoluto foi adicionado a cada poço e homogeneizado por *vortex*. As placas foram incubadas 20 minutos em gelo, em ambiente escuro e centrifugadas 45 minutos 4.000rpm 4°C. O etanol foi retirado por inversão das placas. 100µL de etano 70% foi adicionado a cada poço, as placas centrifugadas 4.000 rpm 4°C por 15 minutos, o etanol 70% removido por inversão e depois por *spin* (200 rpm por 5-7 segundos) com as placas invertidas e secadas a 65°C por aproximadamente 10 minutos. O DNA da reação de seqüenciamento foi ressuspendido por *vortex* em 10µL de tampão de aplicação em *MegaBace* (Amersham) e estocadas a 4°C até aplicação no seqüenciador automático *MegaBace*. Foram processadas 10 placas de 96 poços para cada biblioteca.

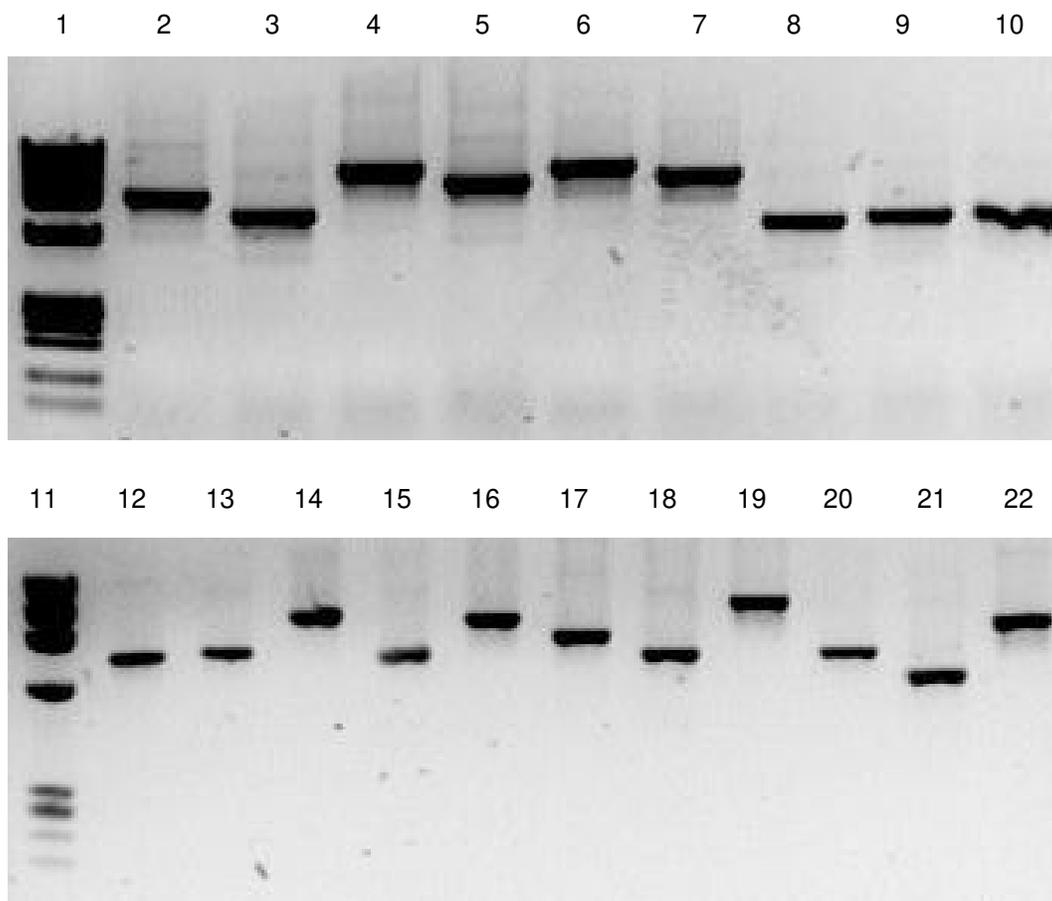


Figura 15- Amplificação dos insertos das bibliotecas subtrativas com iniciadores M13. 1 e 11- Marcador de peso molecular ϕ X 174 haeIII. 2-10- Clones da biblioteca subtrativa RN; 12-22- Clones da biblioteca subtrativa RP.

3.2.4.9- Análise das seqüências geradas

As seqüências geradas pelo *MegaBace* foram submetidas a *Blast* (Basic Local Alignment Search Tool) para seqüências de nucleotídios (*Blastn*), contra o banco de dados disponível em <http://www.ncbi.nlm.nih.gov>. A categorização funcional dos genes foi baseada no *Gene Ontology* e realizada com o auxílio do programa DAVID disponível em <http://apps1.niaid.nih.gov/David/>.

3.2.4.10- Diferenciação de progenitores eritróides humanos em cultura líquida de duas fases

Foi utilizado *buffy-coat* (células mononucleares) de doadores saudáveis. Estas células foram isoladas por centrifugação em gradiente Ficoll-Hypaque (Sigma), onde a camada intermediária foi coletada em tubo estéril e lavada com meio de cultura *Iscove's Modified Dulbecco's Medium* (IMDM, Gibco) por 3 vezes, através de centrifugações sucessivas. O *pellet* de células foi ressuscitado no mesmo meio de cultura (IMDM) e a quantidade e viabilidade das células verificada com azul de trypan em *Câmara de Neubauer*. As células foram cultivadas em uma densidade de 10^6 cels/ ml em meio IMDM, suplementado com soro bovino fetal inativado e estéril (Gibco), meio condicionado coletado da linhagem de células de carcinoma de bexiga 5637 e ciclosporina A (Sandoz). As culturas foram incubadas a 37°C numa atmosfera úmida com 5% de CO₂ por 5 a 7 dias (Fase I). Após esse período, as células não aderentes foram coletadas e lavadas com meio IMDM através de centrifugação e então quantificadas por exclusão de células mortas coradas com azul de trypan em *Câmara de Neubauer*. As células foram novamente colocadas em cultura, em placas de 6 poços numa densidade ótima de $0,5 \times 10^6$ cels/ ml com meio IMDM fresco, suplementado com soro bovino fetal (Gibco), soro albumina bovina deionizada (BSA-Sigma), 2-mercaptoetanol (Sigma), Glutamina (Sigma), dexametazona, transferrina saturada com ferro (Sigma), eritropoietina (EPO) recombinante (Cilag), *Stem Cell Factor* (SCF - Sigma), fungizona (Gibco), estreptomicina (Gibco) e gentamicina (Gibco). As culturas foram incubadas a 37°C em atmosfera úmida contendo 5% de CO₂ e 5% O₂. A morfologia foi avaliada através do preparo do *citospin* (preparo de lâmina por centrifugação contendo uma amostra das células da cultura) corado com *Leishman*. As células foram coletadas para extração de RNA no término da fase I (0 h / sem EPO) e em 11 dias da fase II (216 hs após EPO). A diferenciação de células com produção de Hb F foi avaliada por citometria de fluxo com anticorpos contra HbF.

3.2.4.11- PCR quantitativo em tempo Real (qRT-PCR)

A expressão de alguns genes encontrados nas bibliotecas foi analisada por PCR em tempo real (*Real-Time PCR*). Estes genes foram selecionados com base em seus aspectos funcionais e sua possível correlação com os fenótipos estudados. Os iniciadores foram desenhados com o uso do software *Primer Express* (Applied Biosystems) e estão dispostos na Tabela 2. Foram feitos testes para padronização da concentração e eficiência de cada conjunto de iniciadores. A eficiência de amplificação é obtida da fórmula $10^{(-1/slope)}$, onde *slope* significa o coeficiente de inclinação da curva (Meijerink *et al.*, 2001; Pfaffl, 2001) A detecção de amplificação em tempo real foi realizada no equipamento *ABI 5700 Sequence Detector System*® (Applied Biosystems) utilizando o reagente *SYBRGreen PCR Master Mix*® (Applied Biosystems), que além de conter todos os reagentes necessários para a realização de uma reação de PCR (*dNTPs*, *MgCl₂*, *Tampão*, *Taq Ampli-Gold*) contém também o corante *SYBRGreen*, componente intercalante de DNA dupla-fita necessário para a detecção da reação ciclo a ciclo pelo equipamento acima citado.

Todas as amostras foram amplificadas sempre em duplicata em placas de 96 poços (Sorenson, BioScience Inc) com tampas plásticas que permitem a passagem de luz. As reações foram realizadas em um volume total de 25 µl contendo 10 ng de cDNA dos pacientes, obtido a partir da transcrição reversa com *oligo(dT)* e *Superscript III* (Invitrogen), segundo recomendações do fabricante, 12.5 µl de *SYBRGreen Master Mix PCR* (Applied Biosystems), e os iniciadores nas concentrações pré-estabelecidas. Os dados são coletados na forma de emissão de fluorescência durante os ciclos da reação e analisados em forma de curva. A determinação dos valores para quantificação da expressão é feita com a determinação do *threshold*, que representa o momento em que o sistema detecta luz de maneira “inequívoca” nas amostras, diferenciando o sinal real do ruído de fundo. O ponto de encontro entre esta linha e a curva de amplificação é denominado *threshold cycle*, ou *Ct*. Quanto maior a expressão do gene, mais precocemente ele será amplificado e, conseqüentemente, menor será o *Ct*. Os dados de *Ct* de cada gene foram exportados para uma planilha de Excell (Microsoft Corporation), para análise. O método de quantificação da expressão gênica utilizado é denominado “análise relativa dos dados”, onde a expressão dos genes escolhidos é quantificada em relação a uma situação

considerada como referência, que constitui a amostra calibradora. Todos os dados obtidos foram normalizados a partir da expressão de 2 genes constitutivos utilizados como controles endógenos (GAPDH e ACTB). A análise relativa dos dados foi obtida segundo Vandesompele e cols. (Vandesompele *et al.* 2002), da seguinte maneira: cada média de Ct é subtraída do Ct com menor média entre as amostras analisadas para cada gene, controles (GAPDH e ACTB) ou estudado, obtendo-se assim o ΔCt ($\Delta Ct = Ct \text{ mínimo} - Ct \text{ amostra}$). Em seguida calculamos Q ($2^{\Delta Ct}$), baseado na eficiência de amplificação de cada conjunto de iniciadores, onde 2 corresponde a uma eficiência de 100 %. Encontramos então o fator de normalização (FN) representando a média geométrica entre os valores de Q obtidos para os genes em cada amostra. Os valores de expressão gênica normalizados são então calculados a partir da divisão entre os valores de Q para o gene de estudo e o fator de normalização (Q gene estudo/ FN). Os dados finais são apresentados como a razão entre estes valores normalizados de cada amostra (RP1, RP2, RDB ou 216 hs) sobre os valores da amostra calibradora (RN ou 0 h). As amostras controles (calibradoras) terão então sempre valor relativo de 1 (RN/RN por exemplo) e as amostras em estudo como valores de diferença de expressão (*fold change*) relativa.

Para certificação de que a fluorescência emitida não estava sendo gerada por ampliações inespecíficas, foi feita a análise da curva de dissociação, baseada na *Tm* (*Melting Temperature*) específica de cada produto, que é determinada pela sua composição de nucleotídeos. A cada variação da temperatura, que acontece entre 60°C à 95°C logo após o término do PCR, são coletados dados de fluorescência. A curva de dissociação ocorre quando em determinada temperatura ocorre desnaturação do produto de PCR e conseqüente queda de fluorescência gerada pelo *SYBRGreen*, já que este se intercala somente fita dupla de DNA.

Tabela 2- Iniciadores utilizados para análise de expressão de alguns genes encontrados nas Biblioteca subtrativas.

Genes (Sigla)	Iniciadores
ACTB (F)	5'- AGGCCAACCGCGAGAAG - 3'
ACTB (R)	5'- ACAGCCTGGATAGCAACGTACA - 3'
GAPDH (F)	5'- GCACCGTCAAGGCTGAGAAC - 3'
GAPDH (R)	5'- CCACTTGATTTTGGAGGGATCT - 3'
HBB (F)	5'- AGGCACCGAGCACTTTCTTG - 3'
HBB (R)	5'- ATCTGTCCACTCCTGATCCAGTT - 3'
HBG (F)	5'- CATGGCAAGAAGGTGCTGACT - 3'
HBG (R)	5'- GCAAAGGTGCCCTTGAGATC - 3'
HBA1 (F)	5'- TGGTCCCCACAGACTCAGAGA - 3'
HBA1 (R)	5'- CGGCCTTGACGTTGGTCTT - 3'
ARID1B (F)	5'- TCCAAAGATATCGTTACTCCTGAGG - 3'
ARID1B (R)	5'- TTAATAGTGTCCAAAGCCCACGTAC - 3'
NAP1L4 (F)	5'- GGAGGGAGAAGACGAGGATGA - 3'
NAP1L4 (R)	5'- GCTGGCTGGGTTCTTCTAGA - 3'
SH3BGRL2 (F)	5'- GCGACCGATACTGTGGAGATTAT - 3'
SH3BGRL2 (R)	5'- CATCTTCCACTCTTCTTCTAAGGTTC - 3'
TSPYL1 (F)	5'- AAAGGTGCCCTAAGCAGGAGA - 3'
TSPYL1 (R)	5'- GCTCTTAACGACAGCACAGTCTTC - 3'
ZHX2 (F)	5'- GAACTGGCTGAATCAGACTCCG - 3'
ZHX2 (R)	5'- CACAGCAGTTCTAACAGACTTCCCT - 3'
CD47 (F)	5'- GAAAGCTGTAGAGGAACCCCTTAA - 3'
CD47 (R)	5'- ATCGGAGTCCATCACTTCACTTC - 3'

4- RESULTADOS

4.1- Avaliação Funcional da região R/PYR

As duas cópias do gene β na construção $\mu\text{LCR-}\beta^m\text{-R/PYR-}\beta$ podem ser distinguidas por padrões diferentes de restrição, já que em β^m , o sítio de Nco I, presente na cópia normal, foi substituído, por inserção de um fragmento de 6 pb, pelo sítio da enzima Cla I. A Figura 16 mostra os padrões de digestão obtidos da amostra $\mu\text{LCR-}\beta^m\text{-R/PYR-}\beta$ e do controle $\mu\text{LCR-}\beta^m\text{-}\beta$ nos estudos com transfecção transiente. Após as digestões, o controle (linhas 4 e 6) apresenta as duas bandas correspondentes às duas cópias do gene (β^m e β), sendo que β^m , por estar mais próximo ao μLCR apresentou intensidade maior no gel, como esperado. A amostra $\mu\text{LCR-}\beta^m\text{-R/PYR-}\beta$ (linhas 5 e 7) seguiu um padrão de expressão parecido, porém não foi possível identificar a banda correspondente à segunda cópia (β) no gel de agarose. As Figuras 17 e 18 mostram os padrões de digestão obtidos das amostras contendo $\mu\text{LCR-}\beta^m\text{-R/PYR-}\beta$ e do controle $\mu\text{LCR-}\beta^m\text{-}\beta$, nas células transfectadas de maneira estável. Após as digestões, o controle apresenta o padrão esperado de expressão preferencial do gene de β “marcado” (β^m), localizado mais próximo a região de controle LCR. Neste caso, somente ClaI produz um fragmento de restrição, que aparece com maior intensidade no gel. (Figura 18). Entretanto, as amostras contendo a construção $\mu\text{LCR-}\beta^m\text{-R/PYR-}\beta$ apresentaram variabilidade na expressão das duas cópias de β -globina, com expressão preferencial de β normal, β “marcado” ou de ambos (Figuras 17 e 18). Estes dados diferem dos dados encontrados para transfecção transiente (linhas 8 e 9 das Figuras 17 e 18), podendo resultar da conformação linear apresentada pelos fragmentos na transfecção estável, e podem indicar interferência da região R/PYR na expressão dos genes de β -globina.

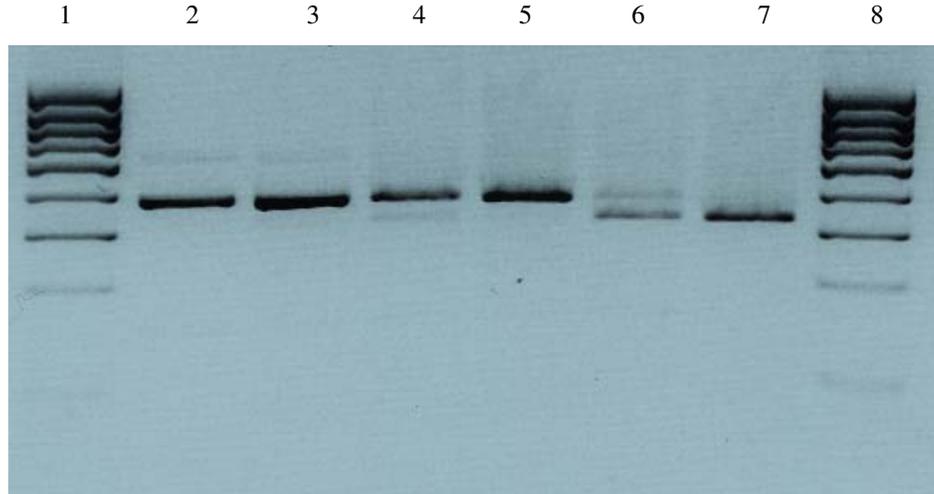


Figura 16- Avaliação do padrão de expressão das duas cópias do gene β por RT-PCR e digestão, após transfecção transiente. 1 e 8- Marcador de peso molecular Hyperladder IV (Bioline). 2 – RT-PCR do controle ($\mu\text{LCR-}\beta^m\text{-}\beta$) não digerido; 3- RT-PCR da amostra ($\mu\text{LCR-}\beta^m\text{-R/PYR-}\beta$) não digerido; 4- $\mu\text{LCR-}\beta^m\text{-}\beta$ digerido com Nco I; 5- $\mu\text{LCR-}\beta^m\text{-R/PYR-}\beta$ digerido com Nco I; 6- $\mu\text{LCR-}\beta^m\text{-}\beta$ digerido com Cla I; $\mu\text{LCR-}\beta^m\text{-R/PYR-}\beta$ digerido com Cla I.

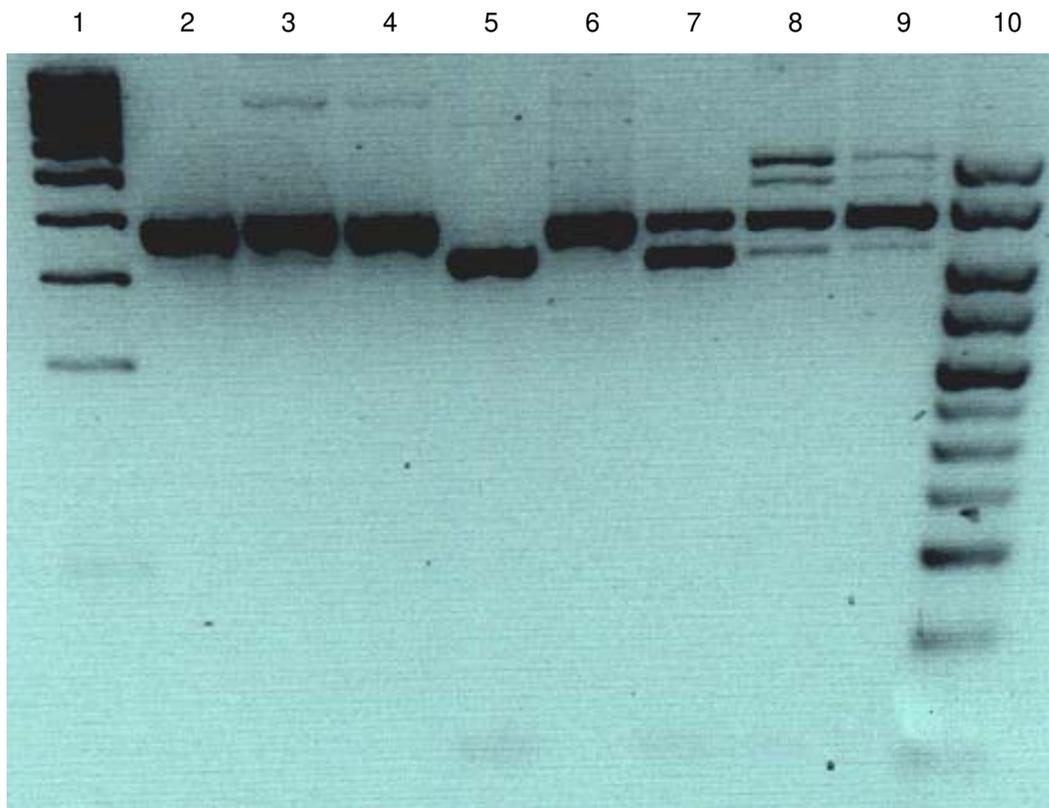


Figura 17- Avaliação do padrão de expressão das duas cópias do gene β por RT-PCR e digestão (NcoI) após transfecção estável em células MEL. 1- Marcador de peso molecular Hyperladder IV (Bioline). 2- RT-PCR do controle (μ LCR- β^m - β) não digerido; 3- RT-PCR do controle (μ LCR- β^m - β) digerido com NcoI; 4-7- RT-PCR de amostras (μ LCR- β^m -R/PYR- β) digeridas com NcoI; 8- RT-PCR do controle (μ LCR- β^m - β) em transfecção transiente digerido com NcoI I; 9- RT-PCR da amostra (μ LCR- β^m -R/PYR- β) em transfecção transiente digerido com NcoI I; 10- Marcador de peso molecular Hyperladder V (Bioline).

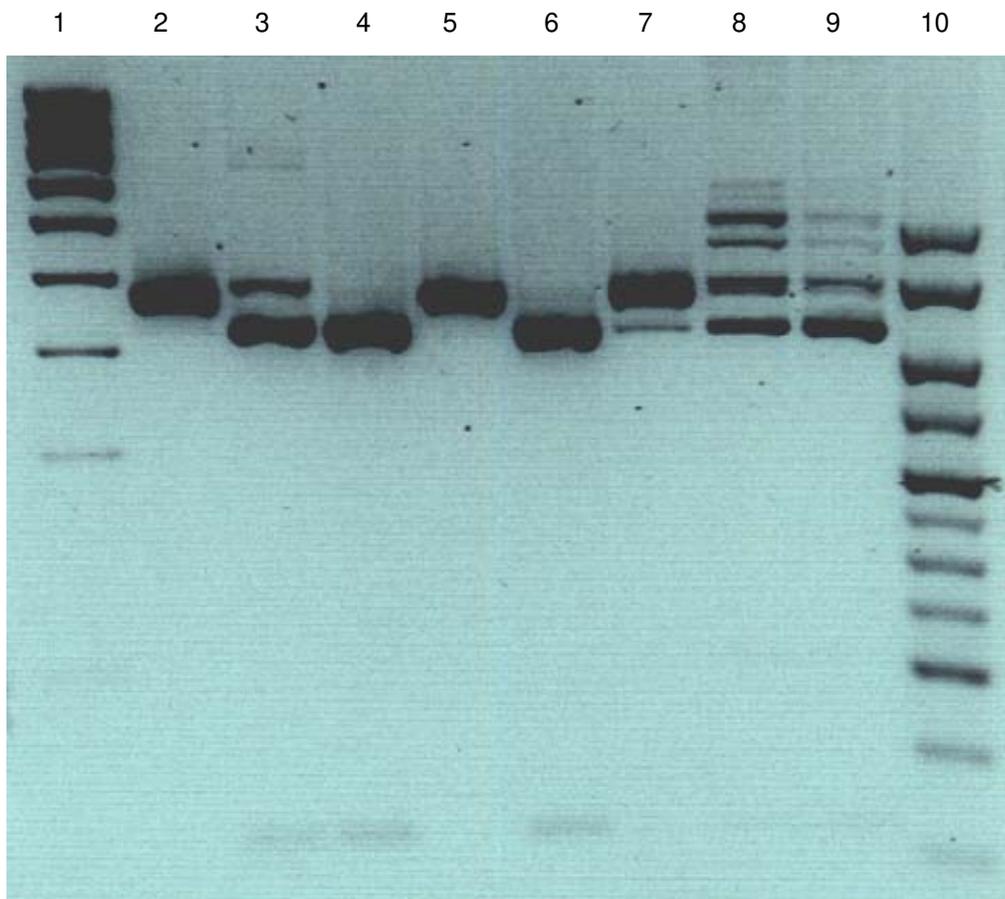


Figura 18- Avaliação do padrão de expressão das duas cópias do gene β por RT-PCR e digestão (ClaI) após transfeção estável em células MEL. 1- Marcador de peso molecular Hyperladder IV (Bioline). 2- RT-PCR do controle ($\mu\text{LCR-}\beta^{\text{m}}\text{-}\beta$) não digerido; 3- RT-PCR do controle ($\mu\text{LCR-}\beta^{\text{m}}\text{-}\beta$) digerido com ClaI; 4-7- RT-PCR de amostras ($\mu\text{LCR-}\beta^{\text{m}}\text{-R/PYR-}\beta$) digeridas com ClaI; 8- RT-PCR do controle ($\mu\text{LCR-}\beta^{\text{m}}\text{-}\beta$) em transfeção transiente digerido com ClaI; 9- RT-PCR da amostra ($\mu\text{LCR-}\beta^{\text{m}}\text{-R/PYR-}\beta$) em transfeção transiente digerido com ClaI; 10- Marcador de peso molecular Hyperladder V (Bioline).

4.2- Biblioteca de cDNA Subtrativa Supressiva

Foram geradas 890 seqüências com similaridade significativa (valor $p > 10^{-5}$), sendo 541 seqüências para reticulócitos normais (RN) e 349 para reticulócitos de PHHF-2 (RP). Destas, 78 representam seqüências únicas em RN e 159 em RP. Foram encontrados 46 genes, 1 ORF (*Open Reading Frame*), 8 proteínas hipotéticas e 7 ESTs (*Expressed Sequences Tags*) diferencialmente expressos em RN, podendo representar genes com expressão diminuída em PHHF-2. A biblioteca RP apresentou 89 genes anotados, 4 ORFs, 13 proteínas hipotéticas e 12 ESTs, indicando genes com expressão aumentada em PHHF-2. Foram ainda identificados clones similares à seqüências genômicas depositadas que não representavam genes categorizados (16 para Normal e 42 para PHHF). A menor redundância na biblioteca RP pode indicar que há um maior número de genes que são ativados ou tem expressão aumentada nesta condição. A lista completa dos transcritos encontrados está disposta nas Tabelas 3 e 4 com o número de ocorrências absoluta e normalizada (clones por 1000) na biblioteca correspondente. Alguns transcritos foram identificados nas duas bibliotecas (Tabela 5) e podem representar ruído de fundo (*background*), ou seja, transcritos não subtraídos completamente. Os genes descritos correspondem a diversas funções celulares e em destaque estão aqueles potencialmente envolvidos com a regulação gênica das globinas. O transcrito com maior representação foi SLC25A37, um membro da família de carreadores de soluto na membrana de mitocôndrias, aparecendo 356 e 78 vezes em RN e RP, respectivamente. Gama globina foi detectada na biblioteca RP e β globina na biblioteca RN, em baixos níveis, o que indica uma subtração e normalização adequada das bibliotecas.

Tabela 3- Transcritos encontrados na biblioteca subtrativa normal (RN), correspondendo a genes com expressão diminuída em PHHF-2.

Gene	Símbolo	N (n/1000)	Localização
Genes Eritróide específicos			
Membrane protein, palmitoylated 1, 55kda	MPP1	4(7.39)	Xq28
Hemoglobin, beta	HBB	2(3.69)	11p15.5
Erythrocyte membrane protein band 4.1	EPB41	1(1.84)	1p33-p32
Glycophorin C	GYPC	1(1.84)	2q14-q21
Fatores de Transcrição			
DNA directed RNA polymerase II polypeptide j-related gene	POLR2J2	3(5.54)	7q11.22
Activating signal cointegrator 1 complex subunit 2	ASCC2	2(3.69)	22q12.1
Zinc fingers and homeoboxes 2	ZHX2	2(3.69)	8q24.13 20q11.21-
Zinc finger protein 335	ZNF335	2(3.69)	q13.12
Nuclear factor I/A	NFIA	1(1.84)	1p31.3-p31.2
Remodelamento de Comatina			
Nucleosome assembly protein 1-like 4	NAP1L4	2(3.69)	11p15.5
Replicação/reparo/recombinação de DNA			
Bloom syndrome	BLM	12(22.18)	15q26.1
PAP associated domain containing 1	PAPD1	1(1.84)	10p11.23
Outras funções			
Sh3 domain binding glutamic acid-rich protein like 2	SH3BGRL2	19(35.12)	6q13-16
Selenium binding protein 1	SELENBP1	12(22.18)	1q21-q22
Choline phosphotransferase 1	CHPT1	8(14.78)	12q
Calcium binding protein p22	CHP	4(7.39)	15q13.3
Development and differentiation enhancing factor 1	DDEF1	3(5.54)	8q24.1-q24.2
ARP3 actin-related protein 3 homolog (yeast)	ACTR3	2(3.69)	2q14.1
Coagulation factor v (proaccelerin, labile factor)	F5	2(3.69)	1q23
Membrane-associated ring finger (c3hc4) 6	MARCH6	2(3.69)	5p15.2
Phosphatidylinositol-4-phosphate 5-kinase, type II, alpha	PIP5K2A	2(3.69)	10p12.32
Ribose 5-phosphate isomerase a	RPIA	2(3.69)	2p11.2
Ubiquitin-conjugating enzyme e2b (rad6 homolog)	UBE2B	2(3.69)	5q23-q31
ATP-binding cassette, sub-family f (gcn20), member 1	ABCF1	1(1.84)	6p21.33
Arachidonate 15-lipoxygenase	ALOX15	1(1.84)	17p13.3
Antizyme inhibitor 1	AZIN1	1(1.84)	8q22.3

Deoxycytidine kinase	DCK	1(1.84)	4q13.3-q21.1
Epithelial stromal interaction 1 (breast)	EPSTI1	1(1.84)	13q13.3
Lymphocyte adaptor protein	LNK	1(1.84)	12q24
5'-nucleotidase, cytosolic III	NT5C3	1(1.84)	7p14.3
Phosphoribosyl aminoimidazole carboxylase	PAICS	1(1.84)	4pter-q21
Ribosomal protein s11	RPS11	1(1.84)	19q13.3
Sec14-like 3 (<i>S. cerevisiae</i>)	SEC14L3	1(1.84)	22q12.2
Sumo1/sentrin specific protease 7	SEN7	1(1.84)	3q12
Soc-2 suppressor of clear homolog (<i>C. elegans</i>)	SHOC2	1(1.84)	10q25
Transportin 2 (importin 3, karyopherin beta 2b)	TNPO2	1(1.84)	19p13.13
Tropomyosin 1 (alpha)	TPM1	1(1.84)	15q22.1
Tripartite motif-containing 56	TRIM56	1(1.84)	7q22.1
Tetratricopeptide repeat domain 11	TTC11	1(1.84)	7q22.1
Open Reading Frames			
Chromosome 6 open reading frame 48	C6orf48	1(1.84)	6p21.3
Proteínas hipotéticas			
Hypothetical protein DKFZP761D221	DKFZp761D221	10(18.48)	1p31.2
Hypothetical protein KIAA1434	KIAA1434	2(3.69)	20p12.3
Hypothetical protein LOC80006	FLJ13611	2(3.69)	5q12.3
Hypothetical protein LOC124801	FLJ30656	2(3.69)	17q21.31
Hypothetical protein FLJ32009	FLJ32009	1(1.84)	11q12.2
Hypothetical gene supported by AL833137	LOC285872	1(1.84)	7q36.3
ESTs (Expressed Sequence Tags)			
Homo sapiens HCM5409 gene, virtual transcript		2(3.69)	
Homo sapiens HCM1154 gene, virtual transcript		2(3.69)	
Homo sapiens cDNA FLJ32900 fis, clone TESTI2005492		1(1.84)	
Homo sapiens cDNA FLJ35212 fis, clone PROST1000136		1(1.84)	
Homo sapiens clone 23700 mRNA sequence		1(1.84)	
Homo sapiens cDNA FLJ46806 fis, clone TRACH3034414		1(1.84)	
Homo sapiens mRNA; cDNA DKFZP566N0924		1(1.84)	
Seqüências genômicas			
Chromosome 11q, clone RP11-259D17		5(9.24)	
Chromosome 3, clone RP11-655N11		4(7.39)	
Chromosome 1, clone RP11-520H14		2(3.69)	
Chromosome 22, clone RP4-539M6		2(3.69)	
Chromosome 9, clone RP11-115J7		1(1.84)	
Chromosome 20, clone RP11-465L10		1(1.84)	

Chromosome 12 BAC RP11-161H23	1(1.84)
Chromosome 12 BAC RP11-181C3	1(1.84)
Chromosome 18, clone RP11-1015N5	1(1.84)
Chromosome 2 BAC clone RP11-273D4	1(1.84)
Chromosome 4 BAC clone RP11-1197E19	1(1.84)
Chromosome 10 clone RP11-390B4	1(1.84)
Chromosome 16 clone RP11-167P11	1(1.84)
Chromosome X clone RP3-473B4	1(1.84)

N- ocorrência absoluta de cada transcrito; n/1000- ocorrência normalizada por 1000.

Tabela 4- Transcritos encontrados na biblioteca subtrativa para PHHF-2 (RP), correspondendo a genes com expressão aumentada em PHHF-2.

Nome	Símbolo	N(n/1000)	Localização
Genes Eritróide específicos			
Ferrochelataze (protoporphyria)	FECH	2(5.69)	18q21.3
Hemoglobin, gamma G/ A	HBG	1(2.84)	11p15.5
Rhesus blood group-associated glycoprotein	RHAG	1(2.84)	6p21.1-p11
Fatores de transcrição			
Zinc finger protein 286	ZNF286	3(8.54)	17p11.2
Homeo box B13	HOXB13	2(5.69)	17q21.2
Transducin (beta)-like 1X-linked receptor 1	TBL1XR1	2(5.69)	3q26.32
Transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)	TLE3	2(5.69)	15q22
Zinc finger protein 135 (clone phz-17)	ZNF135	2(5.69)	19q13.4
Breast cancer 1, early onset	BRCA1	1(2.84)	17q21
			20q11.22-
Itchy homolog E3 ubiquitin protein ligase (mouse)	ITCH	1(2.84)	q11.23
Max interactor 1	MXI1	1(2.84)	10q24-q25
Transcription factor 7 (T-cell specific, HMG-box)	TCF7	1(2.84)	5q31.1
Zinc finger protein 165	ZNF165	1(2.84)	6p21.3
Remodelamento de cromatina			
AT rich interactive domain 1B (SWI1-like)	ARID1B	4(11.39)	6q25.1
TSPY-like 1	TSPYL1	2(5.69)	6q22-q23
SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4	1(2.84)	19p13.2
Replicação/reparo/recombinação de DNA			
Fanconi anemia, complementation group c	FANCC	2(5.69)	9q22.3
Chromosome condensation protein g	HCAP-G	1(2.84)	4p15.33
Topoisomerase (DNA) III beta	TOP3B	1(2.84)	22q11.22

Outras funções			
Spermidine/spermine n1-acetyltransferase	SAT	9(25.64)	Xp22.1
B-cell cll/lymphoma 10	BCL10	5(14.24)	1p22
Caspase 4, apoptosis-related cysteine protease	CASP4	5(14.24)	11q22.2-q22.3
Zinc finger, DHHC-type containing 2	ZDHHC2	5(14.24)	8p21.3-p22
Protein phosphatase 1, catalytic subunit, beta isoform	PPP1CB	4(11.39)	2p23
Phosphorylase, glycogen; liver (hers disease, glycogen storage disease type vi)	PYGL	4(11.39)	14q21-q22
Androgen-induced proliferation inhibitor	APRIN	3(8.54)	13q12.3
Potassium inwardly-rectifying channel, subfamily j, member 2	KCNJ2	3(8.54)	17q23.1-q24.2
Mitogen-activated protein kinase kinase kinase 1	MAP3K1	3(8.54)	5q11.2
Mesoderm induction early response 1	MI-ER1	3(8.54)	1p31.2
Programmed cell death 4 (neoplastic transformation inhibitor)	PDCD4	3(8.54)	10q24
Rap2 interacting protein 8	RPIP8	3(8.54)	17q21.31
Transducer of erbb2, 1	TOB1	3(8.54)	17q21
Amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 2	ALS2CR2	2(5.69)	2q33-q34
Adaptor-related protein complex 2, alpha 1 subunit	AP2A1	2(5.69)	19q13.33
CD47 antigen (rh-related antigen, integrin-associated signal transducer)	CD47	2(5.69)	3q13.1-q13.2
C-type lectin domain family 7, member a	CLEC7A	2(5.69)	12p13.2-p12.3
Deltex 3-like (Drosophila)	DTX3L	2(5.69)	3q21.1
Gli pathogenesis-related 1 (glioma)	GLIPR1	2(5.69)	12q21.1
Major histocompatibility complex, class II, DR alpha	HLA-DRA	2(5.69)	6p21.3
Inhibitor of kappa light polypeptide gene enhancer in b-cells, kinase beta	IKBKB	2(5.69)	8p11.2
Mortality factor 4 like pseudogene 1	MORF4LP1	2(5.69)	1q41-q42
Platelet/endothelial cell adhesion molecule (CD31 antigen)	PECAM1	2(5.69)	17q23
Phosphoribosyl pyrophosphate synthetase 2	PRPS2	2(5.69)	Xp22.3-p22.2
Rap2b, member of ras oncogene family	RAP2B	2(5.69)	3q25.2
Rp42 homolog	RP42	2(5.69)	3q26.3
Spastin	SPAST	2(5.69)	2p24-p21
Tousled-like kinase 1	TLK1	2(5.69)	2q31.1
T cell receptor alpha locus	TRA@	2(5.69)	14q11.2
WD repeat domain 43	WDR43	2(5.69)	2p23.2
Arrestin domain containing 3	ARRDC3	1(2.84)	5q14.3
Copine VIII	CPNE8	1(2.84)	12q12
Dead (asp-glu-ala-asp) box polypeptide 3, x-linked	DDX3X	1(2.84)	Xp11.3-p11.23
Eukaryotic translation initiation factor 2, subunit 1 alpha, 35kda	EIF2S1	1(2.84)	14q23.3
Epoxide hydrolase 1, microsomal (xenobiotic)	EPHX1	1(2.84)	1q42.1
Family with sequence similarity 33, member a	FAM33A	1(2.84)	17q23.2
Homo sapiens major histocompatibility complex, class J, GHKAJ (HLA-GHKAJ) on chromosome 6	HLA-GHKAJ	1(2.84)	6p21.3
Hydroxysteroid (17-beta) dehydrogenase 3	HSD17B3	1(2.84)	9q22
IBR domain containing 2	IBRDC2	1(2.84)	6p22.3
Insulin-like growth factor 1 receptor	IGF1R	1(2.84)	15q26.3
Immunoglobulin heavy variable (ii)-33-1	IGHV11-33-1	1(2.84)	14q32.33
Integrin, alpha x (antigen cd11c (p150), alpha polypeptide)	ITGAX	1(2.84)	16p11.2
Janus kinase 3 (a protein tyrosine kinase, leukocyte)	JAK3	1(2.84)	19p13.1
Lipoprotein lipase	LPL	1(2.84)	8p22
Homo sapiens gene for transmembrane mucin MUC20S	MUC20S	1(2.84)	3q

Neurexin 3	NRXN3	1(2.84)	14q31
Nudix (nucleoside diphosphate linked moiety x)-type motif 4 pseudogene 1	NUDT4P1	1(2.84)	1q21.1
Oxoglutarate (alpha-ketoglutarate) dehydrogenase (lipoamide)	OGDH	1(2.84)	7p14-p13
Platelet-activating factor acetylhydrolase, isoform ib, beta subunit 30kda	PAFAH1B2	1(2.84)	11q23
Peptidylprolyl isomerase d (cyclophilin d)	PPID	1(2.84)	4q31.3
Protein phosphatase 1a (formerly 2c), magnesium-dependent, alpha isoform	PPM1A	1(2.84)	14q23.1
Protein phosphatase 1, regulatory (inhibitor) subunit 14c	PPP1R14C	1(2.84)	6q24.3-q25.3
RAB2B, member ras oncogene family	RAB2B	1(2.84)	14q11.2
Retinoblastoma-associated protein 140	RAP140	1(2.84)	3p14.3
RNA binding motif protein 8a	RBM8A	1(2.84)	1q12
Ribosomal protein l23	RPL23	1(2.84)	17q
Sh3 domain binding glutamic acid-rich protein like	SH3BGRL	1(2.84)	Xq13.3
Solute carrier family 35, member b1	SLC35B1	1(2.84)	17q21.33
Signal transducing adaptor molecule (SH3 domain and itam motif) 2	STAM2	1(2.84)	2q23.3
Staufen, RNA binding protein (Drosophila)	STAU	1(2.84)	20q13.1
THO complex 4	THOC4	1(2.84)	17q25.3
UTP14, u3 small nucleolar ribonucleoprotein, homolog a (yeast)	UTP14A	1(2.84)	Xq25
Vomer nasal 1 receptor 7 pseudogene	VN1R7P	1(2.84)	21p11.2
Exportin 7	XPO7	1(2.84)	8p21
Open Reading Frames			
Chromosome 16 open reading frame 54	C16orf54	7(19.94)	16p11.2
Chromosome 6 open reading frame 62	C6orf62	3(8.54)	6p22.2
Chromosome 16 open reading frame 35 ; -14 gene	C16orf35	1(2.84)	16p13.3
Chromosome 22 open reading frame 13	C22orf13	1(2.84)	22q11.2
Proteínas hipotéticas			
Hypothetical protein FLJ13910	FLJ13910	2(5.69)	2p11.2
Hypothetical protein FLJ10808	FLJ10808	1(2.84)	4q13.2
Hypothetical protein FLJ11155	FLJ11155	1(2.84)	4q32.1
Hypothetical protein FLJ23323	FLJ23323	1(2.84)	1p36.31
KIAA0247	KIAA0247	1(2.84)	14q24.1
Hypothetical protein LOC222070	LOC222070	1(2.84)	7p13
LOC441048	LOC441048	1(2.84)	4q31.3
Hypothetical gene supported by AB002449	LOC441104	1(2.84)	5q22.3
Hypothetical protein MGC34648	MGC34648	1(2.84)	1p35.3
FLJ25476 protein	FLJ25476	1(2.84)	1p35.1
DKFZP564J0863 protein	DKFZP564J0863	1(2.84)	11q13.1
ESTs (Expressed Sequence Tags)			
Homo sapiens cDNA FLJ22642 fis, clone HSI06970		6(17.09)	
Homo sapiens cDNA clone IMAGE:6195280		1(2.84)	
Homo sapiens cDNA FLJ11570 fis, clone HEMBA1003309		1(2.84)	
Homo sapiens cDNA FLJ14162 fis, clone NT2RM4002504		1(2.84)	
Homo sapiens cDNA FLJ33203 fis, clone ADRGL2007080		1(2.84)	
Homo sapiens cDNA FLJ37855 fis, clone BRSSN2014636		1(2.84)	
Homo sapiens cDNA FLJ43934 fis, clone TESTI4013735		1(2.84)	
Homo sapiens cDNA: FLJ21333 fis, clone COL02535		1(2.84)	
Homo sapiens cDNA: FLJ21967 fis, clone HEP05652		1(2.84)	
Homo sapiens mRNA; cDNA DKFZp686K15150		1(2.84)	

Homo sapiens, clone IMAGE:4798349, mRNA	1(2.84)
Seqüências genômicas	
Chromosome 4, clone RP11-429O22	10(28.49)
Chromosome 11q, clone RP11-215H18	6(17.09)
Chromosome 20q, clone RP5-914P20	3(8.54)
Chromosome 16 clone RP11-529K1	2(5.69)
Chromosome 19, CIT-HSP-444n24	2(5.69)
Chromosome 11q, clone RP11-691L4	1(2.84)
Chromosome 2, BAC 854M03 of library RPCI-11	1(2.84)
Chromosome 12, BAC RP11-282O18	1(2.84)
Chromosome 12, BAC RP11-611O2	1(2.84)
Chromosome 12, BAC RP11-629N8	1(2.84)
Chromosome 3, BAC RP11-669C7	1(2.84)
Chromosome 2, clone RP11-347P5	1(2.84)
Chromosome 2, clone RP11-417F21	1(2.84)
Chromosome 1, clone RP11-438H8	1(2.84)
Chromosome 11q13, clone 18h3	1(2.84)
Chromosome 15, clone RP11-394G8	1(2.84)
Chromosome 16, clone CTC-479C5	1(2.84)
Chromosome 16 clone RP11-167P11	1(2.84)
Chromosome 16 clone RP11-322D14	1(2.84)
Chromosome 17, clone RP11-1070B7	1(2.84)
Chromosome 18, clone RP11-1015N5	1(2.84)
Chromosome 18, clone RP11-618K16	1(2.84)
Chromosome 4 clone RP11-603B8	1(2.84)
Chromosome 5 clone RP11-229C3	1(2.84)
Chromosome 19-cosmid f24109	1(2.84)
Chromosome 18 clone RP13-270P17	1(2.84)
Chromosome 7 clone RP4-668E10	1(2.84)
Chromosome 17, clone RP11-143K11	1(2.84)
Chromosome 14 BAC R-247L20	1(2.84)
Chromosome 6, clone DADB-129D20	1(2.84)
Chromosome 10, clone RP11-175O19	1(2.84)
Chromosome 10, clone RP11-192N10	1(2.84)
Chromosome 10, clone RP11-192N10	1(2.84)
Chromosome 9, clone RP11-409K20	1(2.84)
Chromosome 9, clone RP11-443P11	1(2.84)
Chromosome 1p, clone RP1-224A6	1(2.84)
Chromosome 6q, clone RP1-273N12	1(2.84)
Chromosome 20, clone RP3-447F3	1(2.84)
Chromosome 6q, clone RP3-468K18	1(2.84)
Chromosome 6p, clone RP3-514K20	1(2.84)
Chromosome 20, clone RP5-1167H4	1(2.84)

N- ocorrência absoluta de cada transcrito; n/1000- ocorrência normalizada por 1000.

Tabela 5- Transcritos encontrados nas duas bibliotecas.

Gene	Símbolo	RN	RP
Inhibitor of growth family, member 3	ING3	1(1.84)	1(2.84)
Pelota homolog (Drosophila)	PELO	1(1.84)	1(2.84)
WW domain containing oxidoreductase	WVOX	1(1.84)	1(2.84)
30 kDa protein	LOC55831	2(3.69)	1(2.84)
Protein phosphatase 1, catalytic subunit, beta isoform	PPP1CB	2(3.69)	4(11.39)
Proteasome (prosome, macropain) subunit, beta type, 4	PSMB4	3(5.54)	1(2.84)
Fragile histidine triad gene	FHIT	4(7.39)	4(11.39)
Hypothetical protein FLJ20489	FLJ20489	13(24.02)	9(25.64)
Solute carrier family 25, member 37	SLC25A37	356(658)	78(222)

N- ocorrência absoluta de cada transcrito; n/1000- ocorrência normalizada por 1000.

Uma análise da distribuição dos genes quanto à localização genômica (Figura 19) mostra que o cromossomo 1, juntamente com o cromossomo 17 possuem a maior percentagem de genes encontrados, sendo que o maior número de genes agrupados no cromossomo 17 (10 genes) estão distribuídos em 4 Mb no braço longo (17q). Estes dados não se referem a níveis de expressão de genes individuais, mas ao número de genes diferencialmente expressos encontrados. Alguns genes estão presentes apenas na biblioteca normal ou em PHHF-2, o que significa genes com expressão diminuída ou aumentada na PHHF-2, respectivamente (Tabela 6).

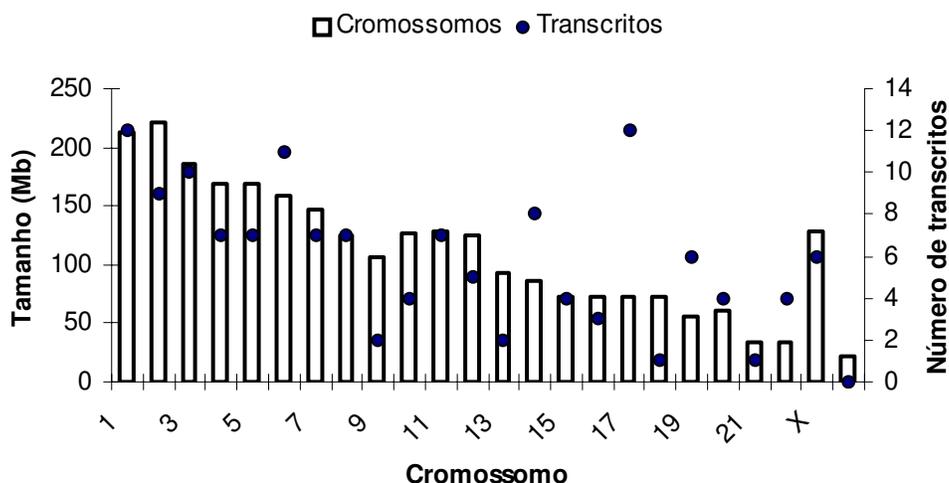


Figura 19- Transcritos diferencialmente expressos distribuídos por localização genômica.

Tabela 6- Distribuição dos transcritos encontrados nos cromossomos 11, 16 e 17 nas bibliotecas de reticulócito Normal (RN) e de PHHF-2 (RP).

Transcrito	Localização cromossômica	N Ocorrências na Biblioteca RN	N Ocorrências na Biblioteca RP
Cromossomo 11			
HBG	11p15.5	-	1
HBB	11p15.5	2	-
NAP1L4	11p15.5	2	-
FLJ32009	11q12.2	1	-
DKFZP564J0863	11q13.1	-	1
CASP4	11q22.2-q22.3	-	5
PAFAH1B2	11q23	-	1
Cromossomo 16			
C16orf54	16p11.2	-	7
ITGAX	16p11.2	-	1
C16orf35	16p13.3	-	1
Cromossomo 17			
ZNF286	17p11.2	-	3
ALOX15	17p13.3	1	-
RPL23	17p21	-	1
BRCA1	17p21	-	1
TOB1	17p21	-	3
HOXB13	17p21.2	-	2
RPIP8	17p21.31	-	3
FLJ30656	17p21.31	2	-
SLC35B1	17p21.33	-	1
PECAM1	17p23	-	2
KCNJ2	17q23.1-q24.2	-	3
FAM33A	17p23.2	-	1
THOC4	17p25.3	-	1

A expressão de alguns genes foi avaliada por PCR em tempo Real, sendo os dados apresentados nas Figuras 20 e 21 como a diferença de expressão relativa à amostra calibradora (reticulócito normal). Como esperado, γ globina apresentou altos níveis de expressão nos reticulócitos dos indivíduos portadores de PHHF-2 e de $\delta\beta$ -talassemia estudados, enquanto a diferença de expressão de β globina foi aproximadamente duas vezes menor que a referência, o que reflete a perda de um dos alelos pela deleção (Figura 20). O nível de gama globina na amostra da paciente portadora de $\delta\beta$ -talassemia Siciliana, embora alto (200-300 vezes maior que o controle), foi de duas a quatro vezes menor que aqueles encontrados em reticulócitos de PHHF-2. Surpreendentemente, α globina também apresentou níveis de RNAm similares aos de β globina. Os genes ARID1B e TSPYL1, associados a remodelamento de cromatina apresentaram um padrão de expressão muito similar nas amostras de PHHF-2, confirmando o aumento de expressão destes genes nesta condição (Figura 21). Entretanto, em $\delta\beta$ -talassemia, a expressão foi de aproximadamente 0.11 (ARID1B) e 0.15 (TSPYL1) vezes quando comparada com reticulócitos normais. SH3BGRL2 e ZHX2 também tiveram seus padrões de expressão confirmados, com diminuição da expressão em PHHF-2, estando também diminuídos na talassemia estudada (Figura 21). O aumento de expressão de CD47 em PHHF-2 foi confirmado nas duas amostras e este gene ainda apresentou um aumento de expressão acentuado em talassemia (Figura 21). Não foi possível em nossos experimentos, entretanto, corroborar a diminuição da expressão de NAP1L4. Os dados para α globina e ZHX2 foram confirmados com dois ensaios independentes e são apresentados como a média destes experimentos. A expressão dos genes estudados foi normalizada tendo como referência dois genes constitutivos (GAPDH e ACTB).

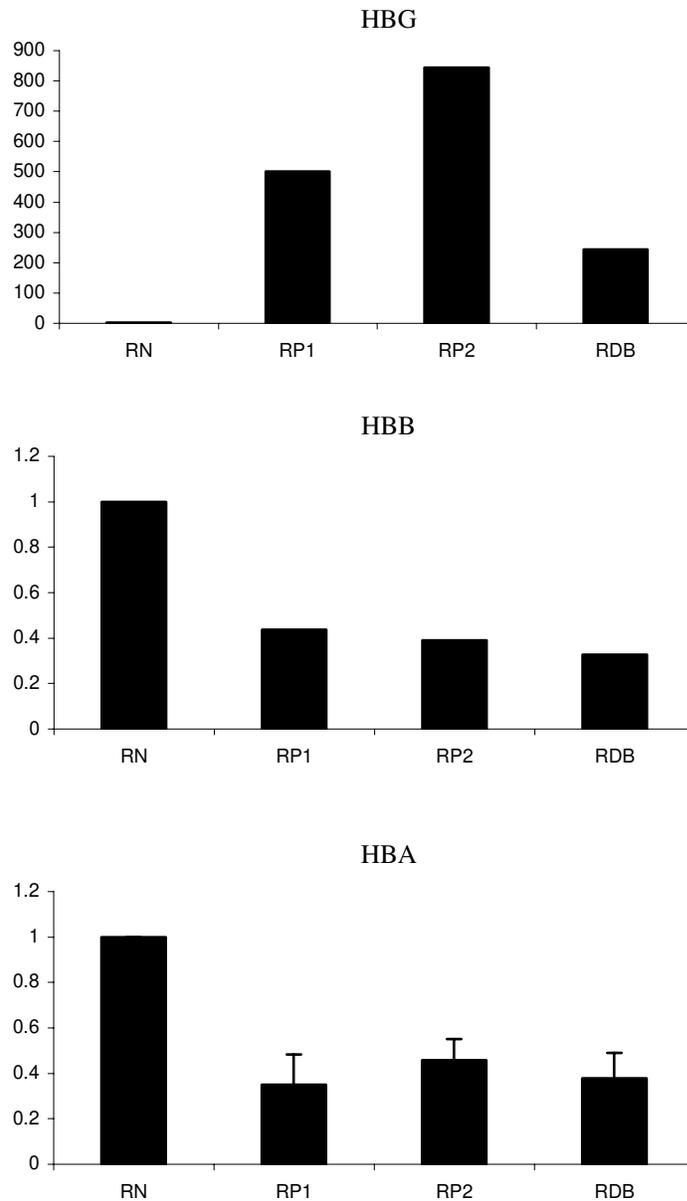


Figura 20- Padrão de expressão gênica avaliado por PCR em tempo real dos genes das globinas alfa (HBA), beta (HBB) e gama (HBG). As amostras tiveram expressão normalizada tendo como referência a média geométrica entre dois genes constitutivos (ACTB e GAPDH). Os valores indicam expressão relativa à amostra calibradora (RN). Os dados para HBA são apresentados como a média de 2 experimentos independentes. RN- Reticulócito Normal; RP-Reticulócito PHHF-2 de 2 indivíduos diferentes. RDB- Reticulócitos de $\delta\beta$ -talassemia Siciliana.

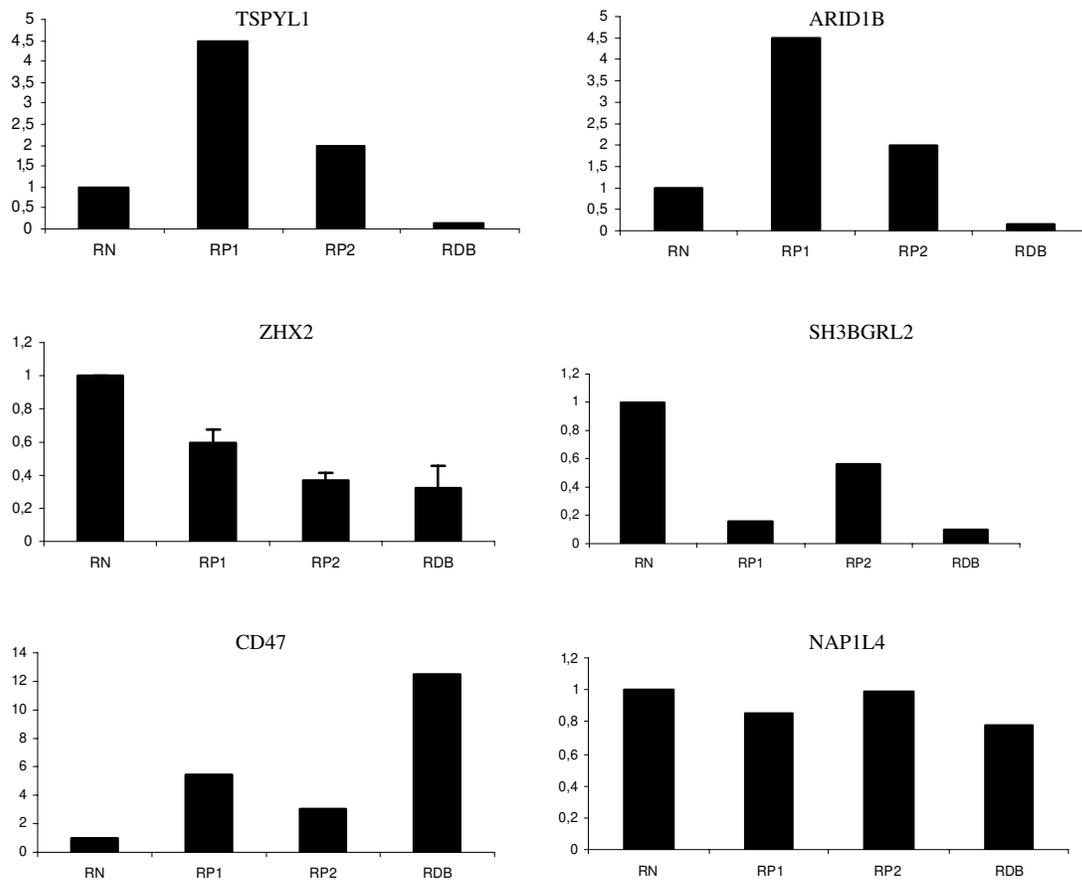


Figura 21- Padrão de expressão gênica avaliado por PCR em tempo real dos genes TSPYL1, ARID1B, ZHX2, SH3BGRL2, NAP1L4 e CD47. As amostras tiveram expressão normalizada tendo como referência a média geométrica entre dois genes constitutivos (ACTB e GAPDH). Os valores indicam expressão relativa à amostra calibradora (RN). Os dados para ZHX2 são apresentados como a média de 2 experimentos independentes. RN-Reticulócito Normal; RP-Reticulócito PHHF-2 de 2 indivíduos diferentes. RDB-Reticulócitos de $\delta\beta$ -talassemia Siciliana.

Investigamos a expressão dos genes ZHX2, SH3BGRL2, TSPYL1 e ARID1B, além, das globinas β e γ em um modelo *in vitro* de diferenciação eritróide (Fibach *et al.* 1989; Pope *et al.*, 2000). Neste sistema, células mononucleares de sangue periférico são cultivadas em meio contendo Eritropoietina (EPO), permitindo a maturação dos progenitores eritróides. Para análise, extraímos RNA de células de indivíduos normais coletadas antes da adição de EPO (0 h) e após 11 dias de diferenciação (216 hs). Nesta fase, é possível encontrar células eritróides em diferentes estágios de maturação, como mostra a Figura 22A. A análise por citometria de fluxo contra Hb F, mostra um aumento de células produzindo Hb F, resultantes do processo de diferenciação (Figura 22B e C).

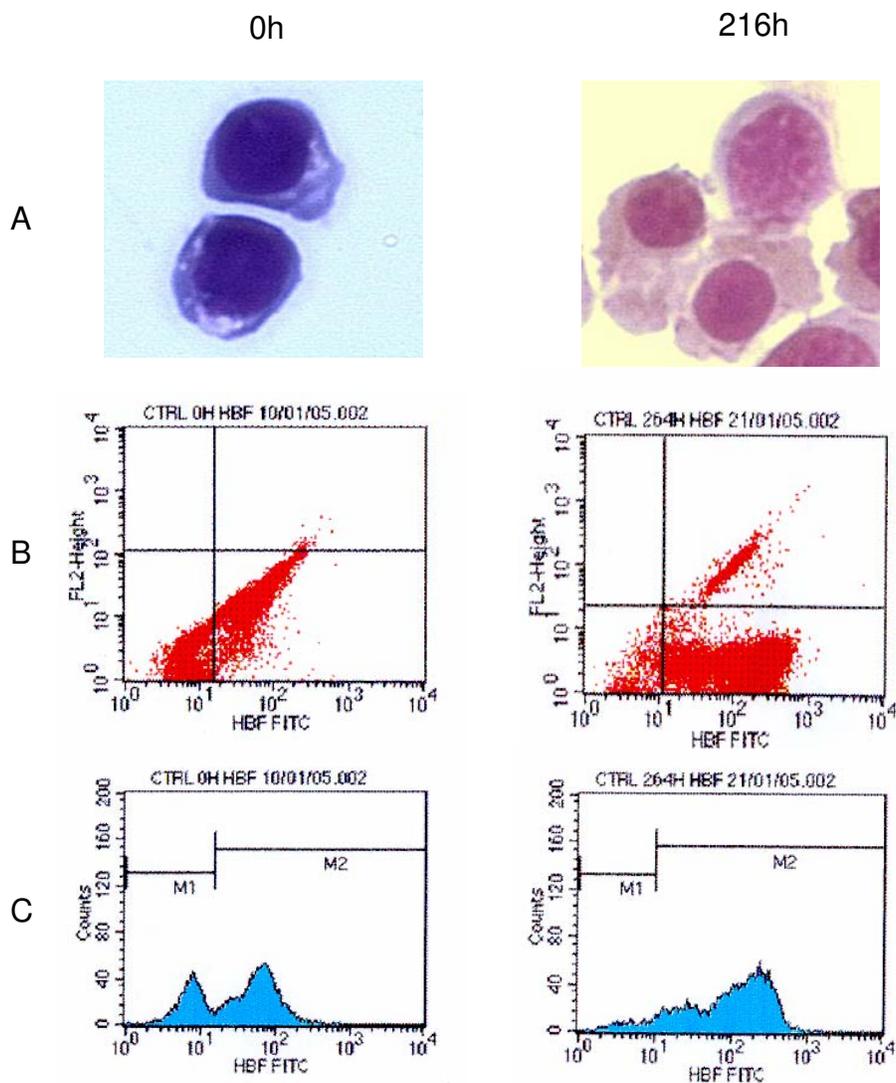


Figura 22- *Citospin* (A) e citometria de fluxo com anticorpo para HbF (B e C) das amostras colhidas de cultura primária de sangue periférico de um indivíduo controle. B- Porcentagem de células no quadrante inferior direito (HbF): 65.89% (0h), 83.88% (216hs). C- Porcentagem de células em M2: 66.23 % com média de fluorescência 69.74 (0h); 94.33 % com fluorescência 159.11.

A análise de expressão gênica na diferenciação *in vitro* (Figura 23) mostra, como esperado, aumento da síntese das globinas. Novamente, ARID1B e TSPYL1 apresentaram padrões similares de expressão, com um incremento de aproximadamente 5 vezes após 216 hs de diferenciação. SH3BGRL2 aumentou a expressão em torno de 27 vezes, enquanto ZHX2 teve expressão diminuída.

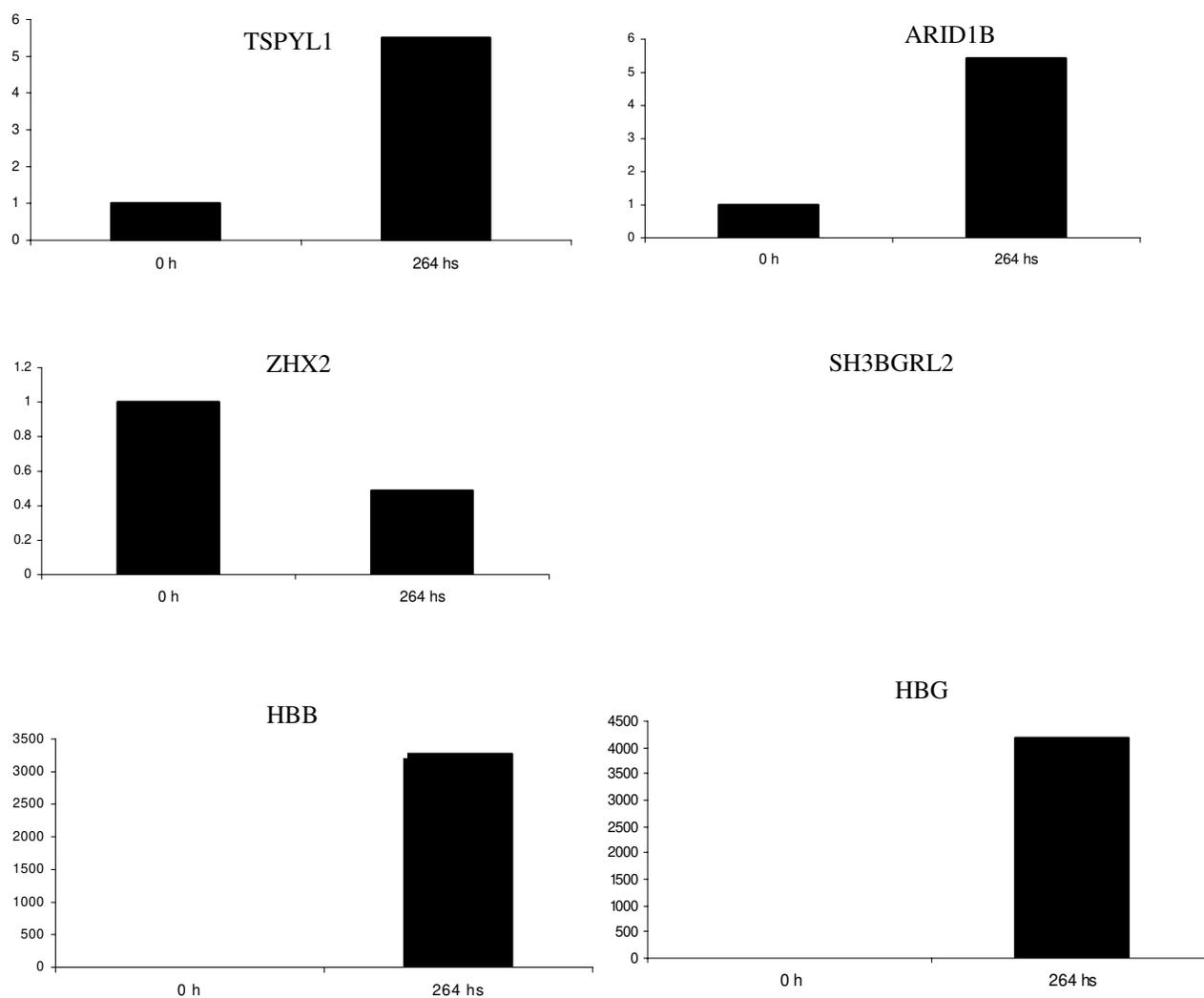


Figura 23- Padrão de expressão gênica avaliado por PCR em tempo real dos genes TSPYL1, ARID1B, ZHX2, SH3BGRL2 e das globinas beta (HBB) e gama (HBG) em cultura primária de II fases, coletadas em 0 h, sem EPO e após 11 dias (216 hs) da adição de EPO. As amostras tiveram expressão normalizada tendo como referência a média geométrica entre dois genes constitutivos (ACTB e GAPDH).

5- DISCUSSÃO

5.1- Avaliação de seqüências regulatórias *in cis*

Alguns estudos demonstram que uma região (fragmento R) potencialmente silenciadora, que distingue a ($\delta\beta$)^o talassemia de origem Africana da PHHF-5 na porção 5', e localizada aproximadamente 3Kb à montante de δ , quando colocada à jusante de um gene repórter que contem promotores dos genes fetal e adulto, exerce um efeito negativo na expressão gênica (Vitale *et al*, 1994; Vitale *et al*, 2001). Este dado reforça a hipótese de que silenciadores são removidos da região. Além disto, o fragmento R possui algumas características interessantes como a presença de duas repetições *Alu* e sítios de ligação para Sp1 e GATA-1, e possivelmente GATA-1 interage de maneira estágio específica com os promotores das globinas, em relação a região (Vitale *et al*, 2001).

Entretanto, estudos em animais transgênicos envolvendo todo o cluster, demonstram que este elemento, quando ausente, não afeta a expressão do gene γ (não participa no silenciamento de γ), mas estimula a expressão de β (Calzolari *et al*, 1999). Este estímulo poderia se dar através da ligação na região de complexos protéicos que interagem com o LCR, tornando a região importante para o recrutamento da maquinaria de transcrição durante o estágio adulto.

Estes resultados levam a uma aparente contradição, onde, em um primeiro momento, tem-se um efeito inibitório, quando a seqüência é colocada à jusante de um gene repórter, e em um outro, propõe-se que sirva como ativador de transcrição. Contudo, outros estudos, com uma região adjacente ao elemento R, localizada aproximadamente 1Kb à montante de δ , demonstram que há interação com um complexo PYR, que aparece durante o estágio adulto, responsável pela modificação da estrutura cromatínica (O'Neill *et al*, 1991). Este complexo possui, como subunidades identificadas, Ikaros, uma proteína *zinc-finger*, específica de linhagens hematopoiéticas, e que funciona como subunidade de ligação ao DNA, subunidades do complexo SWI/SNF, associadas com a ativação da transcrição e MI-2, uma subunidade do complexo NuRD (*Nucleosome remodeling deacetylase*), associadas com a repressão da transcrição gênica (O'Neill *et al*, 1999; O'Neill *et al*, 2000). Estes dados demonstram a participação, em um mesmo complexo protéico, de fatores associados tanto na ativação quanto na

repressão da transcrição gênica. Além disso, a deleção desta região em animais transgênicos (O'Neill *et al*, 1999), e a avaliação de animais nulos (*Knockout*) para o fator Ikaros (Lopez *et al*, 2002), que são importantes para a formação do complexo PYR, apresentam atraso na troca (*switching*) dos genes fetais para os adultos, mostrando que este complexo é relevante para a regulação normal dos genes das globinas durante o desenvolvimento. Um outro trabalho com camundongos transgênicos, onde os autores removem a região intergênica entre γ e δ globina em um Cromossomo Artificial de Levedura (YAC- *Yeast Artificial Chromosome*), não apresenta alterações na expressão dos genes de globina (Gaensler *et al*. 2003). Entretanto, este trabalho carece de uma análise estrutural correta dos transgenes nas linhagens obtidas, o que dificulta a interpretação dos dados.

É possível também que haja interação entre as duas regiões analisadas, mediante associação entre diversos fatores protéicos, como PYR e GATA-1, dentre outros, o que permite um padrão complexo de regulação da transcrição gênica, por reestruturação da cromatina (Vitale *et al*, 2001). Assim, estas duas regiões adjacentes podem atuar em conjunto na regulação do *switching*, fazendo com que seja necessária a presença de toda a região a 5' de δ para uma regulação normal. Como a seqüência linear em que estão dispostos os genes ao longo do cluster da β -globina parece ser importante na regulação total de sua expressão (Peterson *et al*, 1999), é natural cogitar que a posição destas regiões regulatórias em relação aos genes (promotores) também seja relevante. Desta forma, uma região poderia, teoricamente, desempenhar duas funções, mesmo que antagônicas, a depender de sua orientação em relação aos promotores dos genes. É interessante destacar que a região que compreende o fragmento R e o sítio de PYR situa-se justamente em uma posição entre os genes fetais (γ G e γ A) e os adultos (δ e β). Neste caso, esta região poderia participar na inativação de genes dispostos à montante, enquanto aqueles genes à sua jusante poderiam ser acentuados.

Para investigar esta hipótese, ou o efeito da disposição destas seqüências na regulação da expressão dos genes das globinas, realizamos ensaios de expressão gênica *in vitro*, em linhagem celular eritróide, utilizando uma seqüência que engloba o fragmento R e o sítio de ligação a PYR (R/PYR), disposta entre duas cópias do gene da β globina.

Além disto, diferentemente do trabalho que demonstrou efeito silenciador, a construção possui um “micro” LCR (seqüência contendo as regiões centrais dos quatro sítios hipersensíveis à DNase do LCR), importante para a regulação do cluster, e que pode interagir com a região estudada. Estudos demonstram que a cópia do gene β mais próxima do LCR é favorecida na competição, tendo um nível maior de expressão (Peterson & Stamatoyannopoulos, 1993). Neste trabalho, avaliamos se a segunda cópia é favorecida na competição pelo LCR, na presença da região R/PYR e do complexo PYR, ou se há um efeito inibitório na expressão da primeira cópia. Utilizamos como controle uma construção sem o fragmento (Peterson & Stamatoyannopoulos, 1993).

Os resultados obtidos com as construções transfectadas em células MEL podem indicar interferência da região R/PYR na expressão dos genes de β -globina. As duas cópias do gene β na construção μ LCR- β^m -R/PYR- β podem ser distinguidas por padrões diferentes de restrição, já que em β^m , o sítio de Nco I, presente na cópia normal, foi substituído, por inserção de um fragmento de 6 pb, pelo sítio da enzima Cla I. Nos estudos com células transfectadas de maneira estável, o controle apresenta o padrão esperado de expressão preferencial do gene de β “marcado” (β^m), localizado mais próximo a região de controle LCR. Entretanto, as amostras contendo a construção μ LCR- β^m -R/PYR- β apresentaram variabilidade na expressão das duas cópias de β -globina, com expressão preferencial de β normal, β “marcado” ou de ambos. Estes dados diferem dos dados encontrados para transfecção transiente, podendo resultar da conformação linear apresentada pelos fragmentos na transfecção estável. Uma possível explicação para a variabilidade seria um padrão estocástico de interação de complexos protéicos responsáveis por modificações de cromatina e transcrição gênica, na região R/PYR, em cada “clone” de células MEL. Neste aspecto, estes dados podem fornecer indícios de que a região em estudo interage com fatores de transcrição importantes na regulação das globinas. No entanto, não podemos descartar a possibilidade de esta variabilidade ser explicada por artefatos inerentes ao modelo de estudo utilizado. Alguns trabalhos recentes têm demonstrado que a expressão de genes sob controle de seqüências regulatórias em células de linhagem eritróide tem comportamento diferente daquele encontrado em modelos *in vivo* (Li *et al.*, 2005). Outros estudos com estas células transfectadas, como interferência para subunidades do

complexo PYR, ou ainda, estudos *in vivo*, com animais transgênicos contendo a construção elaborada neste trabalho poderiam melhor elucidar a participação desta seqüência intergênica no controle da transcrição dos genes de globina.

5.2- Identificação de fatores de transcrição potencialmente envolvidos com a regulação gênica das globinas

Também foi objeto deste estudo a identificação de genes diferencialmente expressos em células eritróides de sujeitos portadores da PHHF-2, que pudessem ser correlacionados com a regulação da transcrição das globinas e o aumento de hemoglobina fetal nesta condição. Utilizamos o método SSH (*Subtractive Supression Hybridization*) para construir dois perfis de expressão gênica, representados por duas bibliotecas de cDNA (ESTs), para genes expressos preferencialmente em reticulócitos normais (RN) e aqueles com maior freqüência em reticulócitos de PHHF-2 (RP). Este método permite a subtração, ou seja, exclusão dos genes comuns ou com expressão similar, e normalização dos transcritos, possibilitando que os genes menos expressos possam ser detectados. Estas características foram úteis na análise da expressão gênica em reticulócitos, já que a maioria dos transcritos neste tipo celular corresponde às globinas.

Os genes correspondentes aos transcritos encontrados nos dois perfis estão distribuídos no genoma em todos os cromossomos, com exceção do cromossomo Y. Os cromossomos 1 e 17 possuem a maior quantidade de genes. Entretanto, o cromossomo 17 possui 72.8 Mb, aproximadamente três vezes menor que o cromossomo 1, o segundo maior cromossomo humano. Estes dados são condizentes com o fato do cromossomo 17, juntamente com os cromossomos 19 e 22, possuir uma das maiores densidades gênicas entre os cromossomos humanos (Venter *et al*, 2001).

Parte dos transcritos também estão descritos no *Hembase* (<http://hembase.niddk.nih.gov>), repositório de dados de ESTs desenvolvido a partir de uma biblioteca subtrativa de células eritróides em diferenciação *in vitro* (Gubin *et.al*, 1999). Os genes detectados estão associados a diversas funções celulares.

Genes eritróide específicos como globinas, Glicoforina C, banda 4.1 e Ferroquelatase foram encontrados nas bibliotecas. De maneira condizente com a perda de um alelo de β globina e o quadro de persistência de hemoglobina fetal, um clone de globina γ foi detectado na biblioteca de PHHF-2, enquanto 2 seqüências correspondentes à globina β foram encontradas apenas na biblioteca normal. A baixa ocorrência das globinas indica que houve uma normalização adequada das bibliotecas. Alfa globina não foi detectada em nenhum dos perfis.

5.2.1- Globinas

A análise de expressão gênica por PCR quantitativo em tempo real (qRT-PCR) das globinas confirmou a diminuição de β globina em aproximadamente metade da expressão relativa à referência normal, e altos níveis de expressão de γ globina em reticulócitos de PHHF-2, sendo a diferença de expressão (*fold change*) de até aproximadamente 800 vezes maior que o controle normal. O nível de γ globina em uma amostra de $\delta\beta$ -talassemia Siciliana estudada também se apresentou elevado, embora em menor proporção (200-300 vezes) quando comparado aos valores de PHHF-2. Este fato pode ser decorrente da distribuição heterogênea de células F nesta condição e não propriamente da expressão deste gene na célula. Surpreendentemente, α globina apresentou diminuição dos níveis de RNAm similar à globina β em todas as amostras analisadas. Este dado foi confirmado com dois experimentos independentes, sendo cada ensaio realizado com as amostras em duplicata, e os níveis de expressão normalizados com dois genes constitutivos, como controles endógenos da reação. Esta diminuição de α globina pode ser decorrente de processos de regulação comuns dos níveis de globina do tipo α e β . Como avaliamos a expressão gênica em reticulócitos maduros e os iniciadores utilizados foram desenhados em junções de *exons*, as moléculas de RNA estudadas correspondem a RNAm processados e não transcritos primários. Assim, é possível que os dados de expressão por qRT-PCR não reflitam alterações na transcrição gênica de α globina, mas resulte de mecanismos de controle pós-transcricionais dos níveis de RNAm correspondentes aos dois genes. Estes mecanismos pós-transcricionais tem sido relatados

na literatura envolvendo os transcritos de globina (Russel *et al*, 1997), incluindo um trabalho recente envolvendo a expressão de γ e β globina em células contendo a deleção responsável pela $\delta\beta$ -talassemia Corfu (Chakalova *et al* 2005). Na deleção Corfu, o gene da β globina não é removido e alterações na expressão de α globina não foram detectadas. Assim, a diminuição dos níveis de RNAm de α globina encontrados na PHHF-2 e na $\delta\beta$ -talassemia Siciliana pode ser dependente da remoção do fragmento de DNA correspondente ao gene da β globina. Além disto, estes dados podem ajudar a compreender os níveis normais entre globinas α e não- α encontrados em PHHF.

Alternativamente, a própria transcrição do gene α poderia estar alterada nestas células, como resultado de perturbações nas interações entre os cromossomos, provocadas pela remoção de regiões regulatórias importantes. Estudos recentes têm demonstrado que além do fato da localização dos cromossomos no núcleo obedecer a padrões não randômicos, alças de cromatina contendo genes ativos podem se estender além destes territórios definidos e possivelmente interagir com outras regiões do genoma, inclusive outros cromossomos (Spilianakis *et al* 2005). Em trabalho publicado na *Nature Genetics* (Osborne *et al* 2004), os autores demonstram que genes transcripcionalmente ativos colocalizam-se em sub-regiões do núcleo (*transcription factories*), numa frequência maior do que a seria esperada por acaso (alguns genes chegam a ter uma taxa de colocalização de 60%). Este trabalho foi realizado em células eritróides de camundongo e demonstra que mesmo genes localizados em cromossomos diferentes podem estar situados nestas sub-regiões de atividade transcripcional. Este é o caso, por exemplo dos genes que codificam a α globina e β globina, localizados nos cromossomos 11 e 7 murino, respectivamente.

No braço curto do cromossomo 16 humano, onde se localiza o *cluster* de genes da α globina, foram encontrados 3 transcritos expressos na biblioteca de PHHF-2, ou genes supostamente com expressão aumentada nestas células, sendo que C16orf35, uma ORF (*Open Reading Frame*) de função desconhecida, também denominada -14 (ou *conserved gene telomeric to alpha globin cluster*), possui no *intron* 5 o elemento HS40, responsável pela regulação dos genes de α globina (Vyas *et al*, 1995). Caso a expressão desta ORF esteja realmente alterada, podemos ter indícios de como estas

subregiões cromossômicas são capazes de interação na regulação da transcrição das globinas. Estes genes originaram-se por duplicação e devem manter uma regulação fina da transcrição e tradução das proteínas, que seja capaz de manter os níveis adequados das cadeias α e β para a composição correta da hemoglobina.

5.2.2- Fatores de transcrição e remodeladores de cromatina

O objetivo principal da utilização de SSH foi a identificação de genes candidatos a reguladores da síntese de globina, que estivessem diferencialmente expressos em PHHF-2. De fato, encontramos uma série de fatores de transcrição e remodeladores de cromatina diretamente relacionados com o processo de transcrição gênica e que podem, potencialmente, estar envolvidos com a regulação das globinas. Foram encontrados *zinc fingers* da família *Kruppel* e fatores com domínio *homeobox*, relacionados em muitos casos a transcrição dos genes de globina, dentre outros.

Selecionamos alguns genes com descrições na literatura que os correlacionam potencialmente com a transcrição dos genes de globina, para análise por qRT-PCR. ARID1B (também conhecido como BAF250b), por exemplo, é uma proteína do tipo SWI, que auxilia no processo de remodelamento de cromatina, e é uma das subunidades que compõe o complexo PYR, cuja região de associação no *cluster* da β globina foi estudada neste trabalho, em células eritróides (O'Neill *et al.*, 1999; Nie *et al.* 2000; Wang *et al.*, 2004). Os ensaios de expressão gênica mostram um aumento de ARID1B em dois sujeitos com PHHF-2, embora em níveis diferentes (aproximadamente 5 e 2 vezes), confirmando os dados encontrados nas bibliotecas subtrativas. TSPYL1 (Vogel *et al.* 1998), da família de proteínas do tipo NAP (*Nucleosome Assembly Protein*) também está associado a remodelamento de cromatina e localiza-se em uma região cromossômica (6q22-q23) freqüentemente associada com PHHF não ligada ao cromossomo 11, apesar de adjacente a delimitação definida por marcadores polimórficos em trabalhos recentes (Close *et al.* 2004). Curiosamente, a expressão deste gene nas amostras estudadas seguiu um padrão muito similar ao de ARID1B, com a mesma diferença de expressão. Este padrão se manteve nos estudos de diferenciação eritróide com cultura primária de um indivíduo

controle, o que sugere um mecanismo comum de regulação destes remodeladores de cromatina na célula, seja pela localização próxima no genoma (cromossomo 6q) ou por sua similaridade funcional. O aumento da expressão destes genes poderia auxiliar na manutenção de um estado de cromatina aberto na região constituída pelos genes de γ globina, permitindo que estes genes mantenham sua expressão elevada nas células adultas. A expressão de ARID1B e TSPYL1, entretanto, parece estar diminuída ou inalterada em $\delta\beta$ -talassemia. Isto pode resultar dos fragmentos que são removidos em PHHF-2 e mantidos na $\delta\beta$ -talassemia Siciliana, como R/PYR.

Um outro gene que codifica uma proteína do tipo NAP (NAP1L4) é encontrado na mesma região cromossômica (11p15.5) do *cluster* da β globina, a uma distância aproximada de 2Mb. Este gene parece ser responsável pela transferência de grupos de histona para os cromossomos (Rodriguez *et al*, 1997), e uma perturbação na expressão deste gene poderia alterar complexos de cromatina, com conseqüente alteração na expressão dos genes localizados nestas regiões (Ohkuni *et al*, 2003). Esta hipótese poderia explicar a alteração na expressão de genes em células de PHHF-2. No entanto, nossos ensaios de avaliação da expressão gênica por qRT-PCR não foram capazes de corroborar a diminuição deste gene em PHHF-2.

ZHX2, que possui um domínio *zinc finger* e um *homeobox*, tem descrição recente, foi categorizado como um repressor de transcrição (Kawata *et al*. 2003) e está localizado em uma região genômica (8q24.13) onde fora reconhecido um QTL (*Quantitative Trait Loci*) associado a PHHF heterocelular, não ligada ao cromossomo 11 (Garner *et al*. 2002; Garner *et al*. 2004). Além disto, uma alteração de *zhx2*, seu homólogo em camundongo, que normalmente tem aumento de expressão após o nascimento, é responsável pelo quadro de Persistência Hereditária de Alfa-fetoproteína e H19 (Perincheri *et al.*, 2005), proteínas expressas em fígado fetal, interessantemente, sítio da eritropoese neste estágio de desenvolvimento. Dois ensaios independentes de qRT-PCR confirmaram a diminuição de ZHX2 em PHHF-2 e $\delta\beta$ -talassemia Siciliana. Além disto, a expressão de ZHX2 está inversamente correlacionada com os níveis de γ globina nas duas amostras de PHHF-2 estudadas (Figura 24).

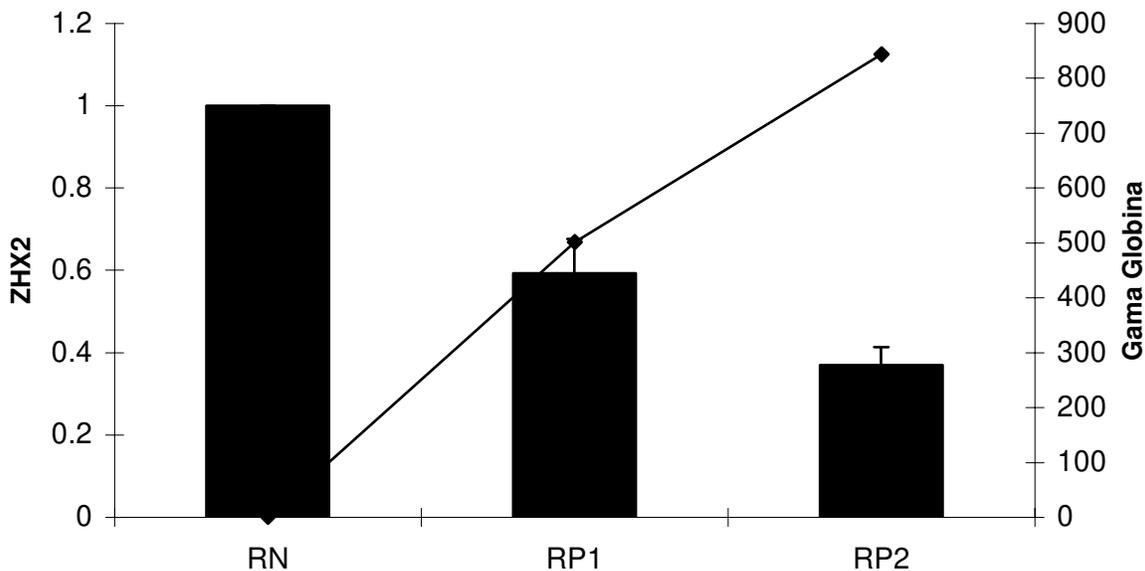


Figura 24- Correlação inversa entre o padrão de expressão de ZHX2 (barras) e gama globina (linha) nas amostras de PHHF estudadas.

Não existem descrições na literatura da participação direta destes fatores de transcrição e remodeladores de cromatina na regulação da síntese de globinas, sendo descritos pela primeira vez como genes candidatos neste trabalho. Além disto, os genes estudados por qRT-PCR, além de outros descritos na biblioteca subtrativa, poderiam atuar em sinergismo para a produção do fenótipo de Persistência de Hb F. ZHX2, por exemplo, poderia em parte ser responsável pelo silenciamento de γ globina no estágio adulto e sua supressão neste período poderia reativar a expressão de γ . Entretanto, para que esta reativação seja consistentemente mantida, talvez seja necessária a presença de fatores que produzam estado de cromatina aberto nestas regiões, como ARID1B e TSPYL1. Isto poderia explicar os fenótipos heterocelular e pancelular encontrados em $\delta\beta$ -talassemia e PHHF deletional, respectivamente. Caso estudos funcionais confirmem a participação de ZHX2 na repressão, e de ARID1B e TSPYL1 no aumento de γ globina em PHHF, neste último caso através de modificações na cromatina, estes genes poderiam vir a ser considerados alvos potenciais para estratégias de reativação de γ globina nas hemoglobinopatias onde o aumento de Hb F tem efeito benéfico.

Outros fatores de transcrição detectados neste estudo incluem membros da família *zinc finger* do tipo *Krüppel* (ZNF135, ZNF165 e ZNF286) e HOXB13, todos encontrados no perfil RP. Outros membros da família *Krüppel*, como KLF1 (EKLF) e KLF11 (FKLF), estão envolvidos na regulação gênica das globinas (Perkins *et al*, 1996; Suske *et al*, 2005). HOXB13 codifica um fator de transcrição que pertence à família *homeobox*, essencial para o desenvolvimento embrionário dos vertebrados. Este gene foi implicado no desenvolvimento da derme fetal e regeneração cutânea (Stelnicki *et al*, 1998). Os autores demonstraram que há uma quantidade substancialmente maior de HOXB13 em fibroblastos fetais cultivados quando comparados com fibroblastos adultos cultivados, o que pode implicar que sua função é mais importante neste estágio de desenvolvimento. Não há relatos da expressão destes genes em tecidos hematopoéticos.

Também detectamos, no perfil RP, SMARCA4 (BGR1), um membro da família de proteínas SWI/SNF, que foi descrito como uma subunidade do complexo PYR (O'Neill *et al*, 2000). Notavelmente, também identificamos genes envolvidos com reparo e recombinação de DNA potencialmente associados a SMARCA4 em complexos protéicos. Estes incluem FANCC, cuja deficiência causa anemia Fanconi, uma desordem genética que predispõe os pacientes a deficiências hematopoéticas, feteitos ao nascimento e câncer (Chen *et al*, 1996; Otsuki *et al*, 2001); BLM, relacionado à família de helicases RecQ, sendo o determinante da Síndrome *Bloom*, que é caracterizada por instabilidade genômica acentuada (Meetei *et al*, 2003); e TOP3B, uma DNA topoisomerase que controle e altera estados topológicos do DNA durante a transcrição e possivelmente interage com helicases da família RecQ (Wu *et al*, 2002). Seria interessante investigar se estes genes possuem expressão alterada como consequência da remoção do sítio de ligação de PYR no *cluster* da β globina e/ou se há uma relação entre esta expressão alterada e mecanismos de reparo de DNA nestas células.

5.2.3- Genes associados a outras funções

No cromossomo 11, sítio da deleção estudada, foram encontrados duas proteínas hipotéticas, um gene associado a metabolismo de lipídio, e CASP4, envolvido com processo de apoptose. CASP4 foi encontrado na biblioteca de PHHF-2,

juntamente com outros genes que também podem estar associados à morte celular, como, por exemplo, BCL-10. Além destes, foi detectado apenas na biblioteca RN o gene NT5C3, cuja carência está associada à Deficiência de Pirimidina 5' Nucleotidase, uma das anemias hemolíticas mais frequentes causadas por defeitos enzimáticos nas células vermelhas (Marinaki *et al.* 2001).

Não podemos afirmar se as células eritróides dos sujeitos portadores do tipo deletional estudado estão submetidas a algum tipo de estresse que possa induzir a um quadro de apoptose ou complicações que levem à anemia hemolítica, apesar da perda de um fragmento extenso de DNA. Além disto, os indivíduos com PHHF não possuem complicações clínicas, e seu quadro hematológico é similar ao normal. Entretanto, dentro do conjunto de genes encontrados nos perfis subtrativos, estão presentes genes que hipoteticamente poderiam atuar com efeito compensatório contra um possível quadro de morte celular ou anemia. CD47, por exemplo, encontrado na biblioteca de PHHF, tem sido correlacionado com uma maior sobrevivência de células eritróides na corrente sanguínea (Oldenborg *et al.*, 2000). CD47 é uma proteína de membrana e se liga a receptores SIRP α nos macrófagos encontrados no baço. Esta ligação ativa vias celulares nos macrófagos que inibem a fagocitose. O qRT-PCR confirmou o aumento de expressão deste gene nas amostras de PHHF-2 e ainda em níveis mais elevados em $\delta\beta$ -talassemia Siciliana (Figura 21). Possivelmente o estudo da expressão deste gene em diferentes formas de talassemias poderia elucidar se CD47 possui uma função importante na sobrevivência de células vermelhas no sangue, com conseqüências clínicas, e estabelecer sua possível utilidade como marcador de gravidade para estas desordens.

Ainda, alguns estudos com o gene ALS2CR2 mostram possível efeito protetor contra apoptose (Sanna *et al.*, 2002). FANCC, cuja deficiência causa a anemia Fanconi, retarda o processo de apoptose e está relacionado a reparo de DNA (Chen *et al.*, 1996; Donahue *et al.* 2002). Este gene está possivelmente aumentado em PHHF-2. Estudos posteriores com estes genes em pacientes portadores de talassemias, como as $\delta\beta$ -talassemias causadas por deleções, poderiam esclarecer se há algum papel diferencial e importante destas vias na patologia e clínica destas anemias.

Podemos também destacar a presença dos genes SH3BGRL2 (Mazzocco *et al.* 2002) e SELENBP1 (Chang *et al.* 1997) que aparecem com alta frequência na biblioteca RN e que estariam assim deficientes em PHHF-2. Embora suas funções não sejam totalmente conhecidas, eles podem ter atividade de redoxina, o que poderia indicar uma alteração da homeostase de oxigênio em PHHF, seja pela alta concentração de HbF nestas células ou ainda em resposta a estresse oxidativo. O qRT-PCR confirmou a diminuição da expressão de SH3BGRL2 em reticulócitos de PHHF-2 e $\delta\beta$ -talassemia Siciliana, e apontou um aumento de expressão deste gene na diferenciação eritróide (cultura primária), o que pode indicar uma participação importante deste gene nas células vermelhas maduras. Esta é, ao nosso conhecimento, a primeira descrição destes genes em células eritróides, e que podem ser importantes na fisiologia normal dos eritrócitos contra os efeitos nocivos de O₂.

5.3- Considerações finais

Embora os modelos experimentais com animais transgênicos envolvendo seqüências regulatórias *in cis*, muitas vezes incluindo o *locus* completo da globina do tipo β em Cromossomos Artificiais de Levedura (YACs), terem sido de fundamental importância na elucidação dos mecanismos de regulação gênica das globinas e outros genes, eles não reproduzem o contexto genômico completo no qual os genes implicados evoluíram. Os dados disponíveis com o camundongo transgênico contendo a PHHF-2 em um YAC (Xiang *et al.*, 2005) claramente demonstram que a simples remoção da região que compreende a deleção no contexto do *locus* completo e a justaposição do acentuador 3' não são suficientes para reproduzir o fenótipo de PHHF. A presença de fatores de transcrição e remodeladores de cromatina, com expressão alterada em PHHF, pode indicar que a participação destes fatores *trans* nos mecanismos envolvidos na elevação de globina gama nestas células, seja mais importante do que previamente suposto, e tenha papel decisivo no fenótipo conseqüente. Obviamente, estes fatores devem interagir com regiões específicas do *cluster*, e possivelmente, em condições normais, com aquelas que são removidas pelas deleções. Ainda, a remoção destas regiões deve interferir na expressão de outros genes e

possivelmente são a causa destas alterações. Assim, um modelo unificado para explicar os níveis aumentados de γ globina em PHHF, além de $\delta\beta$ -talassemias, deve agregar tanto o papel de elementos regulatórios, como os promotores, silenciadores removidos pela deleção, acentuadores justapostos e LCR, com fatores de transcrição e remodeladores de cromatina que sofrem alterações de expressão nas células contendo estas deficiências gênicas (Figura 25). Estes mecanismos também poderiam explicar os dados contraditórios encontrados nos estudos envolvendo a região R/PYR em animais transgênicos.

Embora parte da expressão gênica alterada descrita nas bibliotecas subtrativas poder ser resultado indireto do aumento da síntese de Hb F nas células estudadas, alterações genéticas similares às aqui apresentadas podem implicar perturbações em interações cromossômicas e/ou nos níveis de RNAs regulatórios não codificantes, que podem levar a modificações na expressão gênica em larga escala. A hipótese aqui levantada sugere a possibilidade de deleções como as descritas neste trabalho removerem regiões importantes para o recrutamento e manutenção da maquinaria transcricional em regiões específicas do núcleo, e interferirem, desta maneira, na expressão de genes localizados na vizinhança (*cis*) ou em outros cromossomos (*trans*). Uma região candidata corresponde ao elemento R/PYR estudado neste trabalho em ensaios de transfecção em células eritróides. Ainda, a diminuição de transcritos *intergênicos* ou *intrônicos*, que possuam função regulatória certamente poderia ser um mecanismo explicativo para a perturbação nos níveis de RNAm de genes localizados em outras regiões do genoma. Transcritos intergênicos já foram descritos para o *cluster* da β -globina, porém seu possível papel funcional ainda não é conhecido (Plant *et al*, 2001). Um estudo mais aprofundado sobre a interferência de alterações cromossômicas na expressão global e interações transcricionais entre os genes encontrados poderia fornecer pistas sobre a organização dos cromossomos e dos genes ativos no núcleo celular.

Além disto, possivelmente outros genes são sujeitos a estas perturbações genômicas e não foram detectados em nenhum dos perfis, seja pela quantidade de clones seqüenciados ou pela subtração eficiente de transcritos com menor diferença de expressão. Ao nosso conhecimento, este é a primeira descrição de alterações de expressão gênica em larga escala em células contendo alterações no DNA como as deleções apresentadas. Finalmente, a dissecação dos dados de expressão diferencial apresentados neste trabalho

pode ajudar a elucidar vias importantes na diferenciação eritróide, na manutenção das células vermelhas no sangue e ainda na melhor compreensão dos mecanismos genéticos envolvidos com as patologias das células vermelhas como anemia falciforme e talassemias.

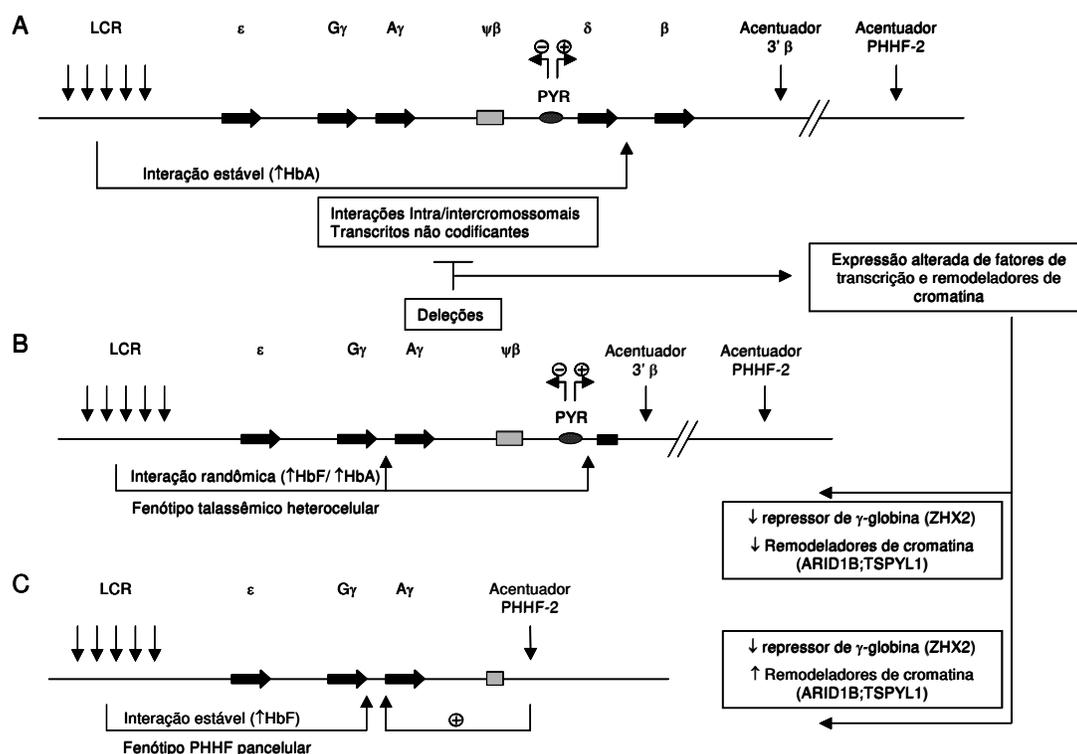


Figura 25- Modelo para explicar o efeito de diferentes formas de deleções dentro do *cluster* da β -globina. Sob condições normais (A), o LCR interage com os genes δ e β durante o estágio adulto, possivelmente envolvendo o complexo PYR. Deleções podem interferir na expressão de fatores *trans* através da interferência em interações inter-cromossômicas ou diminuição da expressão de RNAs regulatórios não codificantes. Estas alterações *cis* e *trans* podem modificar a maneira em que LCR, promotores, silenciadores, acentuadores e fatores de transcrição interagem para a regulação da transcrição. Em adição, diferentes fatores de transcrição e/ou remodeladores de cromatina podem interagir sinergisticamente com as regiões remanescentes para produzir um padrão randômico ($\delta\beta$ -talassemia) (B) ou estável (PHHF) (C) de expressão dos genes de γ globina nas células.

6- CONCLUSÕES

Os estudos funcionais com a região R/PYR podem indicar interferência deste elemento na regulação dos genes de globina. Possivelmente os complexos protéicos que interagem com esta região e com outras seqüências de relevância na regulação da transcrição, como promotores e LCR, criam condições para que γ globina seja silenciada e β globina passe a se expressar no estágio adulto. Além disto, a construção elaborada neste trabalho pode vir a ser utilizada em outros ensaios, como a obtenção de linhagens transgênicas de camundongos, que podem fornecer uma avaliação mais precisa a respeito da função deste elemento na transcrição gênica;

A análise dos perfis diferenciais de expressão gênica em reticulócitos de PHHF-2 e normais mostra alterações em larga escala que podem estar associadas com o fenótipo de PHHF. Como hipóteses, estas modificações podem ser decorrentes de perturbações nas interações cromossômicas e/ou nos níveis de RNAs regulatórios não codificantes;

A observação, por PCR em tempo real, de que os níveis de RNAm de α globina estão diminuídos em PHHF-2 e $\delta\beta$ -talassemia Siciliana sugere um mecanismo de regulação pós-transcricional compensatório para as cadeias de globina ou de regulação gênica por interação física direta entre os cromossomos 11 e 16;

A presença de fatores de transcrição e remodeladores de cromatina, como ARID1B, TSPYL1 e ZHX2, dentre outros, com expressão alterada indica fortemente que outros mecanismos genéticos estão envolvidos com o aumento da síntese de γ globina no estágio adulto, além dos modelos previamente descritos na literatura. Além disto, são candidatos potenciais a reguladores de globina que ainda não haviam sido descritos, e podem ajudar a compreender as diferenças fenotípicas encontradas entre $\delta\beta$ -talassemia e PHHF deletional;

Identificamos ainda outros genes com expressão diferencial, além de seqüências sem similaridade a genes categorizados, como ORFs, ESTs e seqüências genômicas, que podem ter papel importante na fisiologia normal das células vermelhas ou ainda nas hemoglobinopatias hereditárias;

Ao nosso conhecimento, esta é a primeira descrição demonstrando as implicações na expressão gênica em larga escala em células contendo alterações no DNA como as deleções apresentadas. O aprofundamento dos dados aqui apresentados pode resultar em mudanças de paradigmas não só na explicação causal das hemoglobinopatias estudadas, bem como no entendimento da disposição e organização dos genes transcricionalmente ativos no genoma, e de que maneira são regulados.

7- REFERÊNCIAS BIBLIOGRÁFICAS

ALTER, B. P.; GOLL, S. C.; FRENOV, G. D.; GRAVELY, E.; HUISMAN, T. H. J. Globin chain electrophoresis: a new approach to the determination of the G γ / A γ ratio of globin synthesis. **Br. J. Haematology** **44**: 525-534, 1980.

ANAGNOU, N. P.; PAPAYANNOPOULOU, T.; STAMATOYANNOPOULOS, G.; NINHEIUS, A. W. Structurally diverse molecular deletions in the β -globin gene cluster exhibit an identical phenotype in interaction with the β^S -gene. **Blood** **65**: 1245-1251, 1985.

ANAGNOU, N. P.; PEREZ-STABLE, C.; GELINAS, R.; CONSTANTINI, F.; LIAPAKI, K.; CONSTANTOPOULOU, M.; KOSTEAS, T.; MOSCHONAS, N. K. E STAMATOYANNOPOULOS, G. Sequences located 3' to the breakpoint of the hereditary persistence of fetal hemoglobin-3 deletion exhibit enhancer activity and can modify the developmental expression of the human fetal A γ -globin gene in transgenic mice. **J. Biol. Chem.**, **270** (17): 10256-10263. 1995.

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**.**362 (6422)**: 722-8. 1993.

ARCASOY, M. O, ROMANA, M; FABRY, M.E; SKARPIDI, E; NAGEL, R.L; FORGET, B.G. High levels of human γ -globin gene expression in adult mice carrying a transgene of deletion-type hereditary persistence of fetal hemoglobin. **Mol. Cell Biol.**, **17**: 2076-2089. 1997.

ASANO H, LI XS, STAMATOYANNOPOULOS G. FKLf-2: a novel Kruppel-like transcriptional factor that activates globin and other erythroid lineage genes. **Blood**. **95 (11)**: 3578-84. 2000.

BANK, A. Regulation of human fetal hemoglobin: new players, new complexities, **Blood** 2005.

BERG PE, WILLIAMS DM, QIAN RL, COHEN RB, CAO SX, MITTELMAN M, SCHECHTER AN. A common protein binds to two silencers 5' to the human beta-globin gene. **Nucleic Acids Res.** **17 (21)**: 8833-52. 1989.

BETKE, K.; MARTI, H. R.; SCHLICHT, I. Estimation of small percentage of fetal haemoglobin. **Nature** **184**: 1877-1878, 1959.

BODINE DM, LEY TJ. An enhancer element lies 3' to the human A gamma globin gene. **EMBO J.** (10): 2997-3004. 1987.

BOLLEKENS, J. A. & FORGET, B. G. $\delta\beta$ -thalassemia and Hereditary Persistence of Fetal hemoglobin. **Hematology/ Oncology Clinics of North America** **3** (5):399-422. 1991.

BRADLEY, T. B. & RANNEY, H. M.: Acquired disorders of hemoglobin. In: Brown, E. B.: progress in Hematology, vol. VIII, Grune and Stratton, New York, S. Francisco, London. Pp 77-98, 1973.

BUNN, H. & FORGET, B. G. Hemoglobin: molecular, genetic and clinical aspects. Philadelphia, W. B. Saunders Company, 1ª edição, 1986.

BUNN, H. F. Reversing ontogeny. **New England Journal of Medicine**, **328**: 129-131. 1993.

BUNN, H. F. Pathogenesis and treatment of sickle cell disease. **New England Journal of Medicine**, **377**: 762-769. 1997

BULGER. M. & GROUDINE, M. Looping versus linking: toward a model for globin gene activation. **Genes & Development**. **13**:2465-2477. 1999

CALZOLARI, R.; MCMORROW, T.; YANNOUSOS, N.; LANGEVELD, A.; GROSVELD, F. Deletion of a region that is a candidate for the difference between the deletion forms of hereditary persistence of fetal hemoglobin and $\delta\beta$ -thalassemia affects β - but not γ -globin gene expression. **EMBO J.**, **18** (4):949-958. 1999

CAMASCHELLA, C.; SERRA, A.; GOTTARDI, E.; ALFARANO, A.; REVELLO, D.; MAZZA, U.; SAGLIO, G. A new hereditary persistence of fetal hemoglobin has the breakpoint within the 3' β -globin gene enhancer. **Blood**, **74** (4):1000-1005.1990

CHAKALOVA L.; OSBORNE C.S; DAY, Y-F.; GOYENECHEA B.; METAXOTOU-MAVROMATI, A.; KATTAMIS, A.; KATTAMIS, C.; FRASER P The Corfu deltabeta thalassemia deletion disrupts gamma-globin gene silencing and reveals post-transcriptional regulation of HbF expression. **Blood**. **105**:2154-60. 2005.

CHANG, P. W. G.; TSUI, S. K. W.; LIEW, C.; LEE, C.; WAYE, M. M. Y.; FUNG, K. Isolation, characterization, and chromosomal mapping of a novel cDNA clone encoding human selenium binding protein. **J. Cell. Biochem**. **64**: 217-224, 1997.

CHEN, M.; TOMKINS, D. J.; AUERBACH, W.; MCKERLIE, C.; YOUSSEOUFIAN, H.; LIU, L.; GAN, O.; CARREAU, M.; AUERBACH, A.; GROVES, T.; GUIDOS, C. J.; FREEDMAN, M. H.; CROSS, J.; PERCY, D. H.; DICK, J. E.; JOYNER, A. L.; BUCHWALD, M. Inactivation of Fac in mice produces inducible chromosomal instability and reduced fertility reminiscent of Fanconi anaemia. **Nature Genet**. **12**: 448-451, 1996

CLOSE, J.; GAME, L.; CLARK, B.; BERGOUNIOUX, J. GEROVASSILI, A.; THEIN, S. L. Genome annotation of a 1.5 Mb region of human chromosome 6q23 encompassing a quantitative trait locus for fetal hemoglobin expression in adults. **BMC Genomics**. **5**:33. 2004.

COLLINS, F. S.; COLE, J. L.; LOCKWOOD, W. K.; IANNUZZI, M. C. The deletion in both common types of hereditary persistence of fetal hemoglobin is approximately 105 kilobases. **Blood**, **70** (6):1797-1803. 1987

CRAIG, J. E.; BARNETSON, R. A.; PRIOR, J.; RAVEN, J. L.; THEIN, S. L. Rapid detection of deletions causing $\delta\beta$ Thalassemia and Hereditary Persistence of Fetal Hemoglobin by enzymatic amplification. **Blood**. **86** (6):1673-1682. 1994

CRAIG, J. E., EFREMOV, GD, FISHER, C, THEIN, SL. Macedonian (delta beta) zero thalassemia has the same molecular basis as Turkish inversion-deletion (delta beta) zero thalassemia. **Blood**, **85**(4):1146-7. 1995

DACIE, J. V. & LEWIS, S. M. Practical haematology. 6th ed.. Churchill Livingstone, Edinburgh, London, Mulbourne, New York, 1984.

De ANDRADE, T.G.; FATTORI, A.; SAAD, S.T.O; SONATTI, M F; COSTA, F.F. Molecular identification of Sicilian (deltabeta) degrees-thalassemia associated with beta-thalassemia and hemoglobin S in Brazil. **Braz J Med Biol Res.** **35**:873-6. 2002.

De ANDRADE, T.G.; SAAD, S.T.O; SONATTI, M F; COSTA, F.F. Simple fluorescent PCR method for detection of large deletions in the beta-globin gene cluster. **Am J. Hematol.** **72**:225-7. 2003.

DEL RIO, T; URBÁN, Z; CSISZÁR, K.; BOYD, CD.: A gene dosage PCR method for detection of elastin gene deletions in patients with Williams syndrome. **Clin Genet.,** **54**: 129-135, 1998

DIATCHENKO L, LAU YF, CAMPBELL AP, CHENCHIK A, MOQADAM F, HUANG B, LUKYANOV S, LUKYANOV K, GURSKAYA N, SVERDLOV ED, SIEBERT PD. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. **Proc Natl Acad Sci U S A,** **93**:6025-30, 1996

DONAHUE, S. L.; CAMPBELL, C. A DNA double strand break repair defect in Fanconi anemia fibroblasts. **J. Biol. Chem.** **277**: 46243-46247, 2002.

DONZE D, TOWNES TM, BIEKER JJ. Role of erythroid Kruppel-like factor in human gamma- to beta-globin gene switching. **J Biol Chem.** **270** (4): 1955-9. 1995.

EFREMOV GD, NICOLOV N, DUNA H, SCHOEDER WA, MILLER A; HUISMAN THJ. $\delta\beta$ -Thalassemia in two Yugoslavian families: **Scand. J. haematol.** **14**, 226, 1975.

ELDER, J. T.; FORRESTER, W. C. ; THOMPSON, C.. Translocation of an erythroid-specific hypersensitive site in deletion type hereditary persistence of fetal hemoglobin. **Mol. Cell Biol.** **10**: 1382-1389. 1990

ERLICH, H. A. PCR Technology: principles and applications for DNA amplification. Stockton press, New York, 1989.

FEINGOLD, E. A. & FORGET, B. G. The breakpoint of a large deletion causing hereditary persistence of fetal hemoglobin occurs within an erythroid DNA domain remote from the β -globin gene cluster. **Blood,** **74** (6):2178-2186. 1989

FEINGOLD, E. A.; PENNY, L. A.; NIENHUIS, A. W.; FORGET, B. G. An olfactory receptor gene is located in the extended human β -globin gene cluster and is expressed in erythroid cells. **Genomics**, **61**:15-23. 1999

FIBACH, E.; MANOR, D.; OPPENHEIM, A. RACHMILEWITZ, E. A. Proliferation and maturation of human erythroid progenitors in liquid culture. **Blood**. **73**:100-3. 1989.

FORGET, B. G. Molecular basis of hereditary persistence of fetal hemoglobin. **Ann N Y Acad. Sci.** **850**:38-44. 1998

FORGET, B. G. Developmental control of human globin gene expression. **Prog. Clin. Biol. Res.**, 352:313-322. 1990.

FORGET, B. G. Molecular genetics of the human globin genes. In: *Disorders of hemoglobin*. Steinberg, MH; Forget, BG; Higgs, D.; Nagel, RL. (Eds). Cambridge University Press. Cambridge, UK, 2001

FORSBERG EC, DOWNS KM, BRESNICK EH. Direct interaction of NF-E2 with hypersensitive site 2 of the beta-globin locus control region in living cells. **Blood (1)**: 334-9. 2000.

GAENSLER, K. M.; ZHANG, Z.; LIN, C.; YANG, S.; HARDT, K.; FLEBBE-REHWALDT, L. Sequences in the (A)gamma-delta intergenic region are not required for stage-specific regulation of the human beta-globin gene locus. **Proc Natl Acad Sci USA**. **100**:3374-3379. 2003.

GARNER, C.; TATU, T.; BEST, S.; CREARY, L; THEIN, S. L. Evidence of genetic interaction between the beta-globin complex and chromosome 8q in the expression of fetal hemoglobin. **Am J Hum Genet**. **70**:793-9. 2002.

GARNER, C.; SILVER, N.; BEST, S.; MENZEL, S.; MARTIN, C. SPECTOR, T. D.; THEIN, S. L. Quantitative trait locus on chromosome 8q influences the switch from fetal to adult hemoglobin. **Blood**. **104**:2184-6. 2004.

GILMAN, JG, BRINSON, EG, MISHIMA, M. The 32.6 kb Indian delta beta-thalassaemia deletion ends in a 3.4 kb L1 element downstream of the beta-globin gene. **Br J Haematol.** **82(2):**417-21. 1992

GONÇALVES, M. S. Caracterização molecular da persistência hereditária da hemoglobina fetal e da $\delta\beta$ -talassemia em uma população brasileira. Tese-Doutorado-Universidade Estadual de Campinas. Campinas, S.P.,150p.,1993.

GONÇALVES, M. S, FAHEL, S.; FIGUEIREDO, M. S.; KIMURA, E. J.; NECHTMAN, F.; STOMING, T. A; ARRUDA, V. R.; SAAD S. T. O.; COSTA, F. F. Molecular identification of hereditary persistence of fetal hemoglobin type 2 (HPFH type 2) in patients from Brazil. **Ann. Hematol.** **70:** 159-161, 1995.

GONG QH, MCDOWELL JC, DEAN A. Essential role of NF-E2 in remodeling of chromatin structure and transcriptional activation of the epsilon-globin gene in vivo by 5' hypersensitive site 2 of the beta-globin locus control region. **Mol Cell Biol.** **(11):**6055-64. 1996.

GRIBNOU, J; DIDERICH, K; PRUZINA, S.; CALZOLARI, R.; FRASER, P. Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. **Mol Cell**, **5 (2):** 377-86. 2000.

GROSVELD F, VAN ASSENDELFT GB, GREAVES DR, KOLLIAS G. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. **Cell.** **51(6):** 975-85. 1987

GUBIN, A.; NJOROGE, J. M.; BOUFFARD, G. G.; MILLER, J. L. Gene expression in proliferating human erythroid cells. **Genomics.** **59:**168-77. 1999.

HARJU S, MCQUEEN KJ, PETERSON KR. Chromatin structure and control of beta-like globin gene switching. **Exp Biol Med.** **(9):** 683-700. 2002.

HENTHORN, P. S.; MAGER, D. L.; HUISMAN, T. H. J. A gene deletion ending within a complex array of repeated sequences 3' to the human β -globin gene cluster. **Proc. Natl Acad Sci USA**, **83:** 5194-5198. 1986

HENTHORN, P.; SMITHIES, O.; MAGER, D. L. Molecular analysis of deletions in the human β -globin gene cluster: deletion junctions and locations of breakpoints. **Genomics**. **6**:226-237. 1990

HUISMAN, T.; SCHROEDER, W.; EFREMOV, G.; DUMA, H.; MLADENOVSKI, B.; HYMAN, C.; RACHMILEWITZ, E.; VOLVER, N.; MILLER, H.; BRODIE, A.; SHELTON, J. R.; SHELTON, J. B.; APPEL, G. The present status of the heterogeneity of fetal hemoglobin in β -thalassemia: an attempt to unify some observations in thalassemia and related conditions. **Ann. N Y Acad Sci**, **232**:107-124. 1974

JANE, S. M. Understanding fetal globin gene expression: a step towards effective Hb F reactivation in haemoglobinopathies. **Bri. J. Haematology**. **102**: 415-422. 1998.

JIMENEZ, G.; GRIFFITHS, SP.; FORD, AM.; GREAVES, MF.; ENVER, T. Activation of the beta-globin locus control region precedes commitment to the erythroid lineage. **Proc. Natl. Acad. Sci U S A**. **89**(22): 10618-22. 1992

KATSANTONI, E.Z. LANGEVELD, A.; WAI, AW.; DRABEK, D.; GROSVELD, F.; ANAGNOU, NP.; STROUBOULIS, J. Persistent gamma-globin expression in adult transgenic mice is mediated by HPFH-2, HPFH-3, and HPFH-6 breakpoint sequences, **Blood**. **102**: 3412-9. 2003.

KAWATA, H.; YAMADA, K.; SHOU, Z.; MIZUTANI, T.; YAZAWA, T.; YOSHINO, M.; SEKIGUCHI, T.; KAJITANI, T.; MIYAMOTO, K. Zinc-fingers and homeoboxes (ZHX) 2, a novel member of the ZHX family, functions as a transcriptional repressor. **Biochem J**. **373**:747-57. 2003.

KELLUM, R.; SCHEDL, P.A. Position-effect assay for boundaries of higher order chromosomal domains. **Cell**. **64** (5): 941-50. 1991

KOSTEAS, T.; PALENA, A.; ANAGNOU, N. P. Molecular cloning of the breakpoints of the hereditary persistence of fetal hemoglobin type-6 (HPFH-6) deletion and sequence analysis of the novel juxtaposed region from the 3' end of the β -globin gene cluster. **Hum Genet.**, **100**:441-445. 1997

KUTLAR, A.; GARDINE, M. B.; HEADLEAU, M. G.; REESE, A. L.; CLERK, M. P.; NAGLE, S.; SUKUMARAN, P. K.; HUISMAN, T. H. J. Heterogeneity in the molecular basis of three types of HPFH and the relative synthesis of G γ and A γ types of γ chain. **Biochem. Genet.**, **22**: 21-35. 1984

LEE CH, MURPHY MR, LEE JS, CHUNG JH. Targeting a SWI/SNF-related chromatin remodeling complex to the beta-globin promoter in erythroid cells. **Proc Natl Acad Sci U S A.** **96 (22)**:12311-5. 1999.

LEE JS, LEE CH, CHUNG JH. The beta-globin promoter is important for recruitment of erythroid Kruppel-like factor to the locus control region in erythroid cells. **Proc Natl Acad Sci U S A.** **96 (18)**:10051-5. 1999.

LI J, NOGUCHI CT, MILLER W, HARDISON R, SCHECHTER AN. Multiple regulatory elements in the 5'-flanking sequence of the human epsilon-globin gene. **J Biol Chem.** **273(17)**:10202-9. 1998.

LI, J, PETERSOM, KR, FANG, X, STAMATOYANNOPOULOS G. Locus control regions. **Blood**, **100(9)**:3077-86. 2002

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**, **105**:3346-52, 2005.

LOPEZ, R. A.; SCHOETZ, S.; DEANGELIS, K.; O'NEILL. D.; BANK, A. Multiple hematopoietic defects and delayed globin switching in Ikaros null mice. **Proc. Natl. Acad. Sci. USA**, **99**:602-607. 2002.

MANIATTIS, T.; FRITSCH, E. F.; SAMBRROK, J. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory-2^a ed. 1989.

MARINAKI, A. M.; ESCUREDO, E.; DULEY, J. A.; SIMMONDS, H. A.; AMICI, A.; NAPONELLI, V.; MAGNI, G.; SEIP, M.; BEN-BASSAT, I.; HARLEY, E. H.; THEIN, S. L.; REES, D. C. Genetic basis of hemolytic anemia caused by pyrimidine 5-prime-nucleotidase deficiency. **Blood** **97**:3327-3332. 2001.

MAZZOCCO, M.; MAFFEI, M.; EGEO, A.; VERGANO, A.; ARRIGO, P.; Di LISI, R.; GHIOTTO, F.; SCARTEZZINI, P. The identification of a novel human homologue of the SH3 binding glutamic acid-rich (SH3BGR) gene establishes a new family of highly conserved small proteins related to Thioredoxin Superfamily. **Gene**. **291**:233-9. 2002.

McNAUGHTON, J. C. ; COCKBURN, D. J.; HUGHES, G.; JONES, W. A; LAING, N. G.; RAY, P. N.; STOCKWELL, P. A; PETERSEN, G. B. Is gene deletion in eukaryotes sequence-dependent? A study of nine deletion junctions and nineteen other deletion breakpoints in intron 7 of the human dystrophin gene. **Gene**, **222**:41-51, 1998.

McPHERSON, M. J; HAMES, B. D. ; TAYLOR, G. R. PCR 2: a practical approach. *The practical approach series*. IRL Press, New York. 1995.

MEETEI AR, SECHI S, WALLISCH M, YANG D, YOUNG MK, JOENJE H, HOATLIN ME, WANG W. A multiprotein nuclear complex connects Fanconi anemia and Bloom syndrome, **Molec. Cell. Biol.** **23**:3417-3426. 2003.

MEIJERINK, J.; MANDIGERS, C.; VAN DE LOCHT, L.; TONNISSEN, E.; GOODSID, F.; RAEMAEEKERS, J. A novel method to compensate for different amplification efficiencies between patient DNA samples in quantitative real-time PCR. **J Mol Diagn**, **3**:55-61, 2001.

MERIKA M, ORKIN SH. DNA-binding specificity of GATA family transcription factors. **Mol Cell Biol.** **13 (7)**:3999-4010. 1993.

MILLER, IJ; BIEKER, JJ. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. **Mol Cell Biol.** **13 (5)**:2776-86. 1993.

MULLIS, K. B. & FALCONA, F. Specific synthesis of DNA in vitro via polymerase-catalysed reaction. **Methods Enzymol.**, **155**:335-345. 1987

MORGAN, NV; TIPPING, AJ; MATHEW, CG: High frequency of large intragenic deletions in the Falconi Anemia Group A Gene. **Am. J. Hum. Genet.**, **65**:1330-1341, 1999.

NIE, Z.; XUE, Y.; YANG, D.; ZHOU, S.; DERRO, B. J.; ARCHER, T. K.; WANG, W. A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex. **Mol Cell Biol.** **20**:8879-88. 2000.

OHKUNI, K.; SHIRAHIGE, K.; KIKUCHI, A. Genome-wide expression analysis of NAP1 in *Saccharomyces cerevisiae*. **Biochem Biophys Res Commun.** **306**:5-9. 2003.

OLDENBORG, PA; ZHELEZNYAK, A; FANG, YF; LAGENAUR, CF; GRESHAM HD, LINDBERG FP. Role of CD47 as a marker of self on red blood cells. **Science**, **288**:2051-4, 2000.

O'NEILL, D.; BORNSCHLEGEL, K.; FLAMM, M. & CASTLE, M. A DNA- binding factor in adult hematopoietic cells interacts with a pyrimidine-rich domain upstream from the human δ -globin gene. **Proc. Natl. Acad. Sci. USA**, **88**: 8953-8957. 1991.

O'NEILL, D.; YANG, J.; ERDJUMENT-BROMAGE, H.; BORNSCHLEGEL, K.; TEMPST, P. & BANK, A. Tissue-specific and developmental stage-specific DNA binding by a mammalian SWI/SNF complex associated with human fetal-to-adult globin gene switching. **Proc. Natl. Acad. Sci. USA**, **96**:349-354 .1999.

O'NEILL, D. W.; SHOETZ, S.S.; LOPEZ,R.A.; CASTLE, M.;RABINOWITZ, L.; SHOR, E.; KRAWCHUK, D.; GOLL, M.G.; RENZ, M.; SEELIG, H.; HAN, S.; SEONG, R.H.; PARK, S. D.; AGALIOT, T.; MUNSHI, N.; THANOS, D.; ERDJUMENT, H.; TEMPST, P.& BANK, A. Na Ikaros-containing chromatin-remodeling complex in adult-type erythroid cells. **Mol. Cell. Biol.**, **20** (20):7572-7582. 2000.

ÖNER, R, ÖNER, C, ERDEM, G, BALKAN, H, OZDAG, H., ERKAN, M, GUMRUK, F, GURGEY, A, ALTAY, C. A novel (delta beta)(0)-thalassemia due to a approximately 30-kb deletion observed in a Turkish family. **Acta Haematol.****96(4)**:232-6. 1996

OSBORNE CS, CHAKALOVA L, BROWN KE, CARTER D, HORTON A, DEBRAND E, GOYENECHEA B, MITCHELL JA, LOPES S, REIK W, FRASER P. Active genes dynamically colocalize to shared sites of ongoing transcription. **Nat Genet**, **36**:1065-71, 2004.

OTSUKI, T; FURUKAWA, Y; IKEDA, K; ENDO, H; YAMASHITA, T; SHINOHARA, A; IWAMATSU, A; OZAWA, K; LIU, JM. Fanconi anemia protein, FANCA, associates with BRG1, a component of the human SWI/SNF complex, **Hum. Molec. Genet.** **10**:2651-2660. 2001.

PALENA, A, BLAU, A, STAMMATOYANOPOULOS, G, ANAGNOU, NP. Eastern European (delta beta) zero-thalassemia: molecular characterization of a novel 9.1-kb deletion resulting in high levels of fetal hemoglobin in the adult. **Blood.**, **83(12)**:3738-45. 1994.

PATRINOS, GP, DE KROM, M, DE BOER, E, LANGEVELD, A, IMAM, AM, STROUBOULIS, J, DE LAAT, W, GROSVELD, FG. Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. **Genes Dev.**,**18(12)**:1495-509. 2004

PERINCHERI, S.; DINGLE, R. W. C.; PETERSON, M. L.; SPEAR, B. T. Hereditary persistence of alpha-fetoprotein and H19 expression in liver of BALB/cJ mice is due to a retrovirus insertion in the Zhx2 gene. **Proc Natl Acad Sci U S A.** 102:396-401. 2005.

PERKINS, A.C.; GAENSLER, K.M.; ORKINS, S.H. Silencing of human fetal globin expression is impaired in the absence of the adult beta-globin gene activator protein EKLF, **Proc Natl Acad Sci U S A.** **93**:12267-71. 1996

PETERSON KR, CLEGG CH, HUXLEY C, JOSEPHSON BM, HAUGEN HS, FURUKAWA T & STAMATOYANNOPOULOS G. Transgenic mice containing a 248-kb yeast artificial chromosome carrying the human β -globin locus display proper developmental control of human globin genes. **Proc. Natl. Acad. Sci. USA** **90**:7593-7597, 1993.

PETERSON KR & STAMATOYANNOPOULOS G. Role of gene order in developmental control of human gamma- and beta-globin gene expression. **Mol Cell Biol.** **13(8)**:4836-43, 1993.

PETERSON KR, LI QL, CLEGG CH, FURUKAWA T, NAVAS PA, NORTON EJ, KIMBROUGH TG & STAMATOYANNOPOULOS G. Use of yeast artificial chromosomes (YACs) in studies of mammalian development: production of β -globin locus YAC mice carrying human globin developmental mutants. **Proc. Natl. Acad. Sci. USA** **92**:5655-5659, 1995.

PETERSON, K. R.; NAVAS, P. A. & STAMATOYANNOPOULOS, G. Cis-control of globin gene switching resides in gene-specific sequences, as well as in gene order. **Blood**, **94**:614 a, 1999.

PFAFFL, M. W. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic Acids Res**, **29**:e45, 2001.

PLANT, K.E. ROUTLEDGE, S.J. PROUDFROOT, N.J. Intergenic transcription in the human beta-globin gene cluster, **Mol Cell Biol**. **21**, 6507-14. 2001

POPE, S. H.; FIBACH, E.; SUN, J.; CHIN, K. RODGERS, G. P. Two-phase liquid culture system models normal human adult erythropoiesis at the molecular level. **Eur J Haematol**. **64**:292-303. 2000.

RAICH, N.; PAPAYANNOPOULOU, T.; STAMATOYANNOPOULOS, G.; ENVER, T. Demonstration of a human epsilon-globin gene silencer with studies in transgenic mice. **Blood**. **79**(4):861-4. 1992.

RAICH N, CLEGG CH, GROFTI J, ROMEO PH, STAMATOYANNOPOULOS G. GATA1 and YY1 are developmental repressors of the human epsilon-globin gene. **EMBO J**. **14** (4):801-9. 1995.

RODRIGUEZ, P.; MUNROE, D.; PRAWITT, D.; CHU, L.L.; BRIC, E.; KIM, J.; REID, L. H.; DAVIES, C.; NAKAGAMA, H.; LOEBBERT, R.; WINTERPACHT, A; PETRUZZI, M-J; HIGGINS, M.; NOWAK, N.; EVANS, G.; SHOWS, T.; WEISSMAN, B. E.; ZABEL, B.; HOUSMAN; D. E.; PELLETIER, J. Functional characterization of human nucleosome assembly protein-2 (NAP1L4) suggests a role as a histone chaperone. **Genomics**. **44**:253-65. 1997

ROTH, D. B.; PORTER, T. N; WILSON, J. H. Mechanisms of non-homologous recombination in mammalian cells. **Mol. Cell Biol**, **5**:2599-2607. 1985

ROTH, D. B. & WILSON, J. H. Non-homologous recombination in mammalian cells: role of short sequence homologies in the joining reaction. **Mol. Cell Biol**, **6**:4295-4304. 1986.

RUSSEL, J. E.; MORALES, J.; LIEBHABER, S.A. The role of mRNA stability in the control of globin gene expression, **Prog. Nucleic Acids Res. Mol. Biol** **57**:249-287. 1997.

SAGLIO, G.; CAMASCHELLA, C.; SERRA, A.; BERTERO, T.; CAMBRIN, G. R.; GUERRASIO, A.; MAZZA, U.; IZZO, P.; TERRAGNI, F.; GIGLIONI, B.; COMI, P. E OTTOLENGHI, S. Italian type of deletional hereditary persistence of fetal hemoglobin. **Blood**, **68** (3):646-651. 1986.

SANNA, M. G.; Da SILVA CORREIA, J.; LUO, Y.; CHUANG, B.; PAULSON, L. M.; NGUYEN, B.; DEVERAUX, Q. L.; ULEVITCH, R. J. ILPIP, a novel anti-apoptotic protein that enhances XIAP-mediated activation of JNK1 and protection against apoptosis. **J Biol Chem**. **277**:30454-62. 2002.

SHAVIT JA, MOTOHASHI H, ONODERA K, AKASAKA J, YAMAMOTO M, ENGEL JD. Impaired megakaryopoiesis and behavioral defects in mafG-null mutant mice. **Genes Dev.** (14):2164-74. 1998.

SERJEANT, G. R. & SERJEANT, B. E. Sickle Cell Disease. 3 ° Ed. Oxford Medical Publications, New York, 2001.

SHIOKOWA S.; YAMADA H.; Taki hard Y. Molecular analysis of japanese $\delta\beta^0$ -thal.. **Blood**, **72**:1771-1775, 1988.

SPILIANAKIS, C. G; LALIOTI, M. D.; TOWN, T.; LEE, G. R.; FLAVELL, R. A. Interchromosomal associations between alternatively expressed loci. **Nature**. **435**:637-45. 2005.

STAMATOYANNOPOULOS G, JOSEPHSON B, ZHANG JW, LI Q. Developmental regulation of human gamma-globin genes in transgenic mice. **Mol Cell Biol**. **13** (12):7636-44. 1993.

STAMATOYANNOPOULOS, G. & NIENHUIS, A. W. Hemoglobin switching. *Molecular basis of blood diseases*. 2^a edição (ed. Stamatoyannopoulos *et al.*) pp. 107-156. Saunders, Philadelphia. 1994.

STAMATOYANNOPOULOS, G. MAJEAURS, P. W.; PERLMUTTER, R. M.; VARMUS, H. & NIENHUIS, A. W. Hemoglobin switching. *Molecular basis of blood diseases*. 3^a edição (ed. Stamatoyannopoulos *et al.*) pp. 107-156. Saunders, Philadelphia. 2001.

STELNICKI, E.J.; ARBEIT, J.; CASS, D.L.; SANER, C. HARRISON, M.; LARGMAN, C. Modulation of the human homeobox genes PRX-2 and HOXB13 in scarless fetal wounds, **J. Invest. Derm.** **111**:57-63. 1998.

SUSKE, G. BRUFORD, E. PHILIPSEN, S. Mammalian SP/KLF transcription factors: bring in the family, **Genomics**. **85**:551-6. 2005

TRENT R. J.; SUIRKLYS L.; JONES P. Thai $\delta\beta$ -thal and its interaction with γ -Thal. **Hemoglobin**, **12**:101-110, 1988.

TSANG AP, VISVADER JE, TURNER CA, FUJIWARA Y, YU C, WEISS MJ, CROSSLEY M, ORKIN SH. FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. **Cell**. **90 (1)**: 109-19. 1997.

TSANG AP, FUJIWARA Y, HOM DB, ORKIN SH. Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. **Genes Dev**. **12 (8)**:1176-88. 1998.

TUAN, D.; MURNAME, M J.; DE RIEL, J. K.; FORGET, B. G. Heterogeneity in the molecular basis of HPFH. **Nature**, **285**:335-337, 1980.

TUAN, D.; FEINGOLD, E.; NEWMAN, M.; WEISSMAN, S. M.; FORGET, B. G. Different 3' end points of deletions causing $\delta\beta$ -thalassemia and hereditary persistence of fetal hemoglobin: implications for the control of γ -globin gene expression in man. **Proc. Natl. Acad. Sci USA**, **83**:6937-6941. 1983

VANDESOMPELE, J.; De PRETER, K.; PATTYN, F.; POPPE, B.; ROY, N. V.; De PAEPE, A.; SPELEMAN, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **Genome Biol.** **3**:RESEARCH0034. 2002.

VENTER, J.C.; ADAMS, M. D.; MYERS, E. W.; LI, P. W.; MURAL, R. J.; SUTTON, G. G. ET AL. The sequence of the human genome. **Science**, **291**:1304-51, 2001.

VITALE, M.; MARZO, R.; CALZOLARI, R.; ACUTO, S.; O'NEILL, D.; BANK, A; MAGGIO, A .Evidence for a globin promoter-specific silencer element located upstream of the human δ -globin gene. **Biochemical and Biophysical Research Communications.** **204** (1): 413-418, 1994.

VITALE, M.; CALZOLARI, R; MARZO, R.; ACUTO, S. & MAGGIO, A. A region upstream of the human δ -globin gene shows a stage-specific interaction with globin promoters in erythroid cell lines. **Blood Cell. Mol. Dis.** **74** (5):874-881. 2001.

VOGEL, T.; DITTRICH, O.; MEHRAEIN, Y.; DECHEND, F.; SCHNIEDERS, F.; SCHMIDTKE, J. Murine and human TSPYL genes: novel members of the TSPY-SET-NAP1L1 family. *Cytogenet Cell Genet.* **81**:265-70. 1998.

VYAS, P.; VICKERS, M. A.; PICKETS, D. J.; HIGGS, D. R. Conservation of position and sequence of a novel, widely expressed gene containing the major human alpha-globin regulatory element. **Genomics.** **29**:679-89. 1995.

WALL L, DEBOER E, GROSVELD F. The human beta-globin gene 3' enhancer contains multiple binding sites for an erythroid-specific protein. **Genes Dev.** (**9**):1089-100. 1988.

WANG, X.; NAGL, N. G.; WILSKERD, D.; VAN SCOY, M.; PACCHIONE, S.; YACIUK, P.; DALLAS, P. B.; MORAN, E. Two related ARID family proteins are alternative subunits of human SWI/SNF complexes. **Biochem J.** **383**:319-25. 2004.

WEATHERAL, DJ & CLEGG, JB: *The Thalassemia Syndromes.* Oxford, Blackwell Science Ltd, 4th ed., 2001.

WIJGERD, M.; GROSVELD, F. E FRASER, P. Transcription complex stability and chromatin dynamics in vivo. **Nature**. **377**:209-213, 1995.

WINICHAGOON, P.; FUCHAROEN, S.; THONGLAIROAN, V.; WASI, P. Thai G γ (A $\gamma\delta\beta$) $^{\circ}$ thalassemia and its interaction with a single γ -globin gene on a chromosome carrying β° thalassemia. **Hemoglobin**, **14**:185-197. 1990.

WU, L HICKSON, I.D. The Bloom's syndrome helicase stimulates the activity of human topoisomerase III alpha. **Nucleic Acids Res**. **30**: 4823-9. 2002.

XIANG P, HAN H, BARKESS G, OLAVE I, FANG X, YIN W, STAMATOYANNOPOULOS G, LI Q. Juxtaposition of the HPFH2 enhancer is not sufficient to reactivate the gamma-globin gene in adult erythropoiesis, **Hum Mol Genet**. **14 (20)**:3047-56. 2005

YAU, S. C.; BOBROW, C; MATHEW, C, G.; ABBS, S. J.: Accurate Diagnosis of carriers of deletions and duplications in Duchenne/ Becker musculae dystrophy by fluorescent dosage analysis. **J. Med. Genet.**, **33**:550-558, 1996.

ZHANG J.W.; STAMATOYANNOPOULOS G.; ANAGNOU P. Laotian ($\delta\beta$) $^{\circ}$ -thal. Molecular characterization of anovel deletion associated with increased production of HbF. **Blood**, **72 (3)**:983-988, 1988.

ZHU YY, MACHLEDER EM, CHENCHIK A, LI R, SIEBERT PD. Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. **Biotechniques**, **30**:892-7, 2001.

ZUELKER WW, ROBINSON AR & BOOKE CR. Reciprocal relationship of hemoglobins A2 and F in beta chain thalasseмииs: a key to the genetic control of hemoglobin F. **Blood**. **17**, 393, 1961.

8- ANEXOS



Diagnóstico molecular por dosagem gênica fluorescente (PCR Multiplex Fluorescente)

Para identificarmos as deleções estudadas neste trabalho, desenvolvemos uma modificação da técnica descrita previamente para Distrofia Muscular Duchenne/Becker (Yau *et al*, 1996), baseada na possibilidade de se identificar diferenças quantitativas na amplificação de amostras por PCR fluorescente, em presença de deleções ou duplicações, onde o produto final de amplificação é diretamente proporcional à quantidade inicial de cópias da seqüência alvo. Na Tabela 7 estão listados os primers desenhados para o mapeamento de deleções.

Tabela 7- Primers marcados com 6-FAM (*) para rastreamento de deleções no cluster da β -globina.

Primers	Região	Seqüência (5' → 3')
BACTE *	β -Actina	ACT CTT CCA GCC TTC CTT C
BACTD	β -Actina	CAG GTC TTT GCG GAT GTC
EPSE *	Gene ϵ	AGA CAG AGA GGC AGC AGC AC
EPSD	Gene ϵ	CTT CCA CAT TCA TCT TGC TCC
BETAE*	Gene β	GTG GAT GAA GTT GGT GGT GAG
BETAD	Gene β	GCC CAG TTT CTA TTG GTC TCC
3'GAMAE *	3' Gene γ A	GCC TGT TTG TGT TTG TGG TG
3'GAMAD	3' Gene γ A	TAC TCC CAC TTG CAG AAC TCC
PSEUE *	$\psi\beta$	ACA AGA AGG CTG CTG CCA C
PSEUD	$\psi\beta$	CTG CCT TTC TGC ACC TTG TC
5'DELE *	5' Gene δ	CCC TCT TCC TTC CTT CCT TTC
5'DELD	5' Gene δ	AAT GGC AGA CAA ATA CAC ACC C
3' BETAE*	3Kb 3' BETA	TAA CGC AGC TCC TCA CCA G
3' BETAD	3Kb 3' BETA	AGC CTT CCT CTC TCA ACT CG
5' EPSE *	16kb5' Gene ϵ	TCC TGC TGG AAT CTC CTC AC
5' EPSD	16kb5' Gene ϵ	AGG CCA CTA ACG AAG AAG ATG
HPF6E*	3' PHHF-6	TCT TTC TCC TGA GGC GTC TG
HPF6D	3' PHHF-6	GGC TGA TCA TGG GCA CAT AG
HPF1E*	3' PHHF-1	AGA CTG AAA GGA AGA GAG GTG G
HPF1D	3' PHHF-1	CTT GGC TAC TCA TCC ATC ATT C

Determinação da fase exponencial da reação: Foi realizada uma escala de ciclos (15-30 ciclos) de amplificação, para determinação da fase exponencial da reação de PCR (Figura 26), que foi obtida pelos valores das áreas dos picos fluorescentes em cada ciclo. Desta forma, podemos determinar a quantidade de ciclos (23 ciclos) que são necessárias para identificarmos diferenças quantitativas nas amostras estudadas.

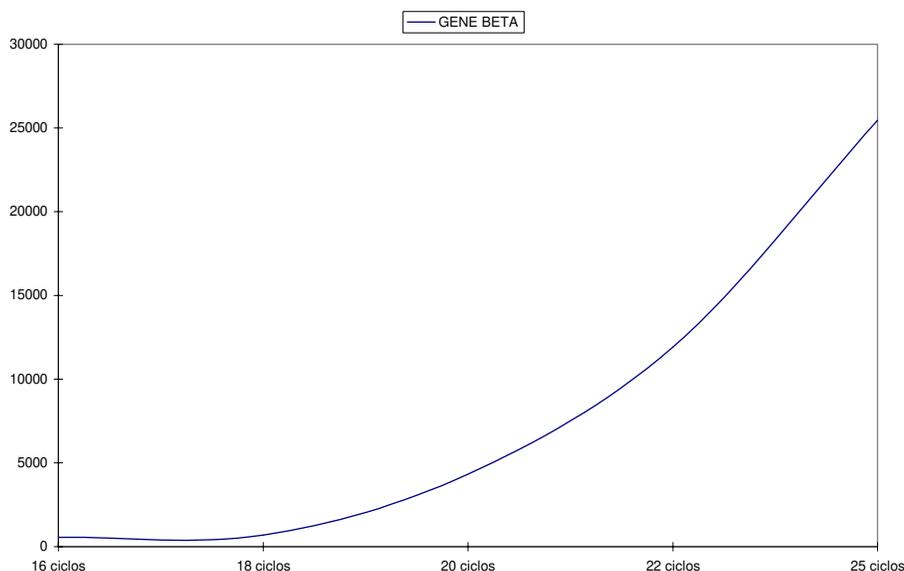


Figura 26- Determinação da fase exponencial da reação de PCR para o gene beta (De Andrade 2002).

PCR multiplex

Para a primeira série de iniciadores (multiplex I), foram utilizados BACTE e BACTD (como controle interno da β -Actina), BETAE e BETAD (para o gene β), EPSE e EPSD (para o gene ϵ), 5'DELE e 5'DELD (para região 5' do gene δ), PSEUE e PSEUD (para $\psi\beta$), 3'GAMAE e 3'GAMAD (região 3' do gene γ A). Uma segunda série de primers (multiplex II) possui BACTE e BACTD, 5'EPSE e 5'EPSD (para uma região 16 Kb 5' do gene ϵ), HPF6E e HPF6D (para a região 3' do ponto de quebra da PHHF-6) e

HPF1E e HPF1D (para a região 3' do ponto de quebra da PHHF-1). A terceira série de iniciadores (multiplex III) contem 3'BETAE e 3'BETAD (para uma região 3 KB 3' do gene β), e BACTE e BACTD (como controle interno da β -Actina). Os parâmetros para as reações foram: Desnaturação inicial a 94° C 5 min., seguidos de 23 ciclos de desnaturação a 94° C 1min; anelamento a 60° C por 30 segundos e extensão a 72° C por 15 segundos; com extensão final a 72° C por 30 min. Cada reação teve um volume final de 25 μ l, com 100ng de DNA molde; 0.2 mmol/L de cada dNTP; 0.5 U de Taq DNA polimerase e 5 pmoles de cada oligonucleotídeo, diluídos em tampão de PCR, realizadas no termociclador Perkin Elmer Cetus. As Tabelas 8, 9 e 10 mostram o arranjo dos multiplex, e a Figura 27, a posição aproximada das regiões estudadas.

Tabela 8- Iniciadores que formam Multiplex I.

Primers (Multiplex I)	Tamanho do Fragmento
BETAE e BETAD	83 pb
5' DELE e 5' DELD	93 pb
BACTE e BACTD	100 pb
PSEUE e PSEUD	110 pb
3' GAMAE e 3' GAMAD	120pb
EPSE e EPSD	140 pb

Tabela 9- Iniciadores que formam Multiplex II.

Primers (Multiplex II)	Tamanho do Fragmento
BACTE e BACTD	100 pb
HPF1E e HPF1D	199 pb
5' EPSE e 5' EPSD	278 pb
HPF6E e HPF6D	140 pb

Tabela 10- Iniciadores que formam Multiplex III.

Primers (Multiplex III)	Tamanho do Fragmento
3' BETAE e 3' BETAD	73 pb
BACTE e BACTD	100 pb

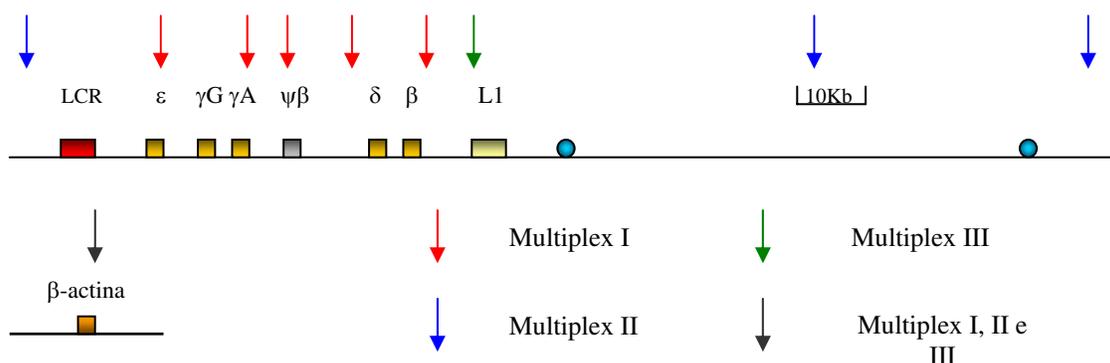


Figura 27- Posição aproximada dos pares de primer utilizados no rastreamento de deleções novas, por dosagem gênica baseada em Multiplex fluorescente (De Andrade 2002).

Análise

Os dados foram analisados em sequenciador automático de capilaridade (310 Genetic Analyser- ABI Prism- Applied Biosystems), com 2µl de cada reação e 1µl de marcador GS Rox 500, diluídos em 20µl de formamida, e em software apropriado (GeneScan-Applied Biosystems). Os dados da área de cada pico foram transferidos para uma planilha de Excel, onde foi calculado o Coeficiente de Dosagem Gênica [CDG = (área do pico x da amostra / área do pico y da amostra) / (área do pico x do controle/ área do pico y do controle)], tendo valores teóricos de 1 para região diplóide normal, 0,5 ou 2 para deleções e 1,5 ou 0,75 (a depender da posição dos valores na fração) para duplicações. Para o cálculo, foi utilizada a média de 3 amostras controles. Além disto, cada reação foi realizada em triplicata.

Amplificação do ponto de quebra

Para amplificarmos o ponto de quebra das deleções, realizamos uma reação contendo enzima para fragmentos longos (ELONGASE- Life Technologies), e primers que pudessem flanquear a deleção em estudo (Tabela 11). Os parâmetros para a reação foram: Desnaturação inicial a 94° C 1min. , seguidos de 30 ciclos de desnaturação a 94° C 30 segundos; anelamento a 60° C por 1 minuto e extensão a 68° C por 15 minutos. Cada reação teve um volume final de 25µl, com 200ng de DNA molde; 0.2 mmol/L de cada dNTP; 1 µl de ELONGASE (Life Technologies) e 5 pmoles de cada oligonucleotídeo, diluídos em tampões A e B (1:1 - 1.5 mM de MgCl₂), realizadas no termociclador Perkin Elmer Cetus.

Tabela 11- Iniciadores utilizados nas reações com Elongase.

Primer	Posição	Seqüência (5' → 3')
E1	5'- Entre ψβ e δ	CCGGGACAAGGTGCAGAAAG
HPF1D	3'- PHHF-1	CTTGGCTACTCATCCATCATTC

Quadros de valores de Coeficiente de Dosagem Gênica

As áreas em negrito representam as regiões detectadas com valores de deleções (0,5 ou 2 - a depender da posição na fração), segundo o cálculo exemplo para deleção Siciliana: $CDG = (\text{área do pico beta da amostra} / \text{área do pico actina da amostra}) / (\text{área do pico beta do controle} / \text{área do pico actina do controle}) = (9247/6867) / (25228,3/7774,3) = 0,41$

Deleção Siciliana

Multiplex I

Área do pico			Região					
Região	Controle	Amostra	EPSLON	GAMA-A	PSEUDO	5' DELTA	BETA	ACTINA
EPSLON	24849	18949	-	0,97	0,86	0,85	2,08	0,86
GAMA-A	21759,3	17023	1,03	-	0,89	0,87	2,13	0,89
PSEUDO	26315,6	23217	1,16	1,13	-	0,98	2,41	1,00
5' DELTA	28628,6	25780	1,18	1,15	1,02	-	2,46	1,02
BETA	25228,3	9247	0,48	0,47	0,42	0,41	-	0,41
ACTINA	7774,3	6867	1,16	1,13	1,00	0,98	2,41	-

Multiplex II

Área do pico			Região			
Região	Controle	Amostra	5' EPS.	3' PHHF-6	3' PHHF-1	ACTINA
5' EPS.	3449	821	-	1,17	1,15	1,21
3' PHHF-6	17318,5	3534	0,86	-	0,99	1,04
3' PHHF-1	17764	3669	0,87	1,01	-	1,05
ACTINA	3592	706	0,83	0,96	0,95	-

Multiplex III

Área do pico			Região	
Região	Controle	Amostra	3' BETA	ACTINA
3' BETA	53141,3	4775	-	0,56
ACTINA	64075	10227	1,78	-

PHHF-2*Multiplex I*

Área do pico			Região					
Região	Controle	Amostra	EPSLON	GAMA-A	PSEUDO	5' DELTA	BETA	ACTINA
EPSLON	13550	4662	-	0,76	0,88	1,37	1,85	0,76
GAMA-A	46442	21163	1,32	-	1,16	1,81	2,45	1,00
PSEUDO	28966	11331	1,14	0,86	-	1,55	2,10	0,86
5' DELTA	42950	10818	0,73	0,55	0,64	-	1,35	0,55
BETA	24517	4560	0,54	0,41	0,48	0,74	-	0,41
ACTINA	9757	4441	1,32	1,00	1,16	1,81	2,45	-

Multiplex II

Área do pico			Região			
Região	Controle	Amostra	5' EPS.	3' PHHF-6	3' PHHF-1	ACTINA
5' EPS.	3449	1343	-	2,71	1,28	1,18
3' PHHF-6	17318,5	2487	0,37	-	0,47	0,44
3' PHHF-1	17764	5384	0,78	2,11	-	0,92
ACTINA	3592	1182	0,85	2,29	1,09	-

Multiplex III

Área do pico			Região	
Região	Controle	Amostra	3' BETA	ACTINA
3' BETA	53141,3	12351	-	0,76
ACTINA	64075	19710	1,32	-

Os valores de deleções situaram-se nas regiões 5'delta, beta, 3'beta e 3'PHHF-6 (Figura 28 e 29). As regiões de pseudo-beta e 3' PHHF-1 apresentaram valores normais, indicando ausência de deleção. Assim, com estes valores, os pontos de quebra estão localizados na região entre pseudo beta e 5'delta (montante) e 3'PHHF-6 e 3'PHHF-1 (jusante). Realizamos reações contendo enzima para fragmentos longos (ELONGASE- Life technologies), com o objetivo de amplificar um fragmento contendo o ponto de quebra. A Figura 30 mostra os fragmentos amplificados (aproximadamente 10 Kb), que foram submetidos a sequenciamento direto.

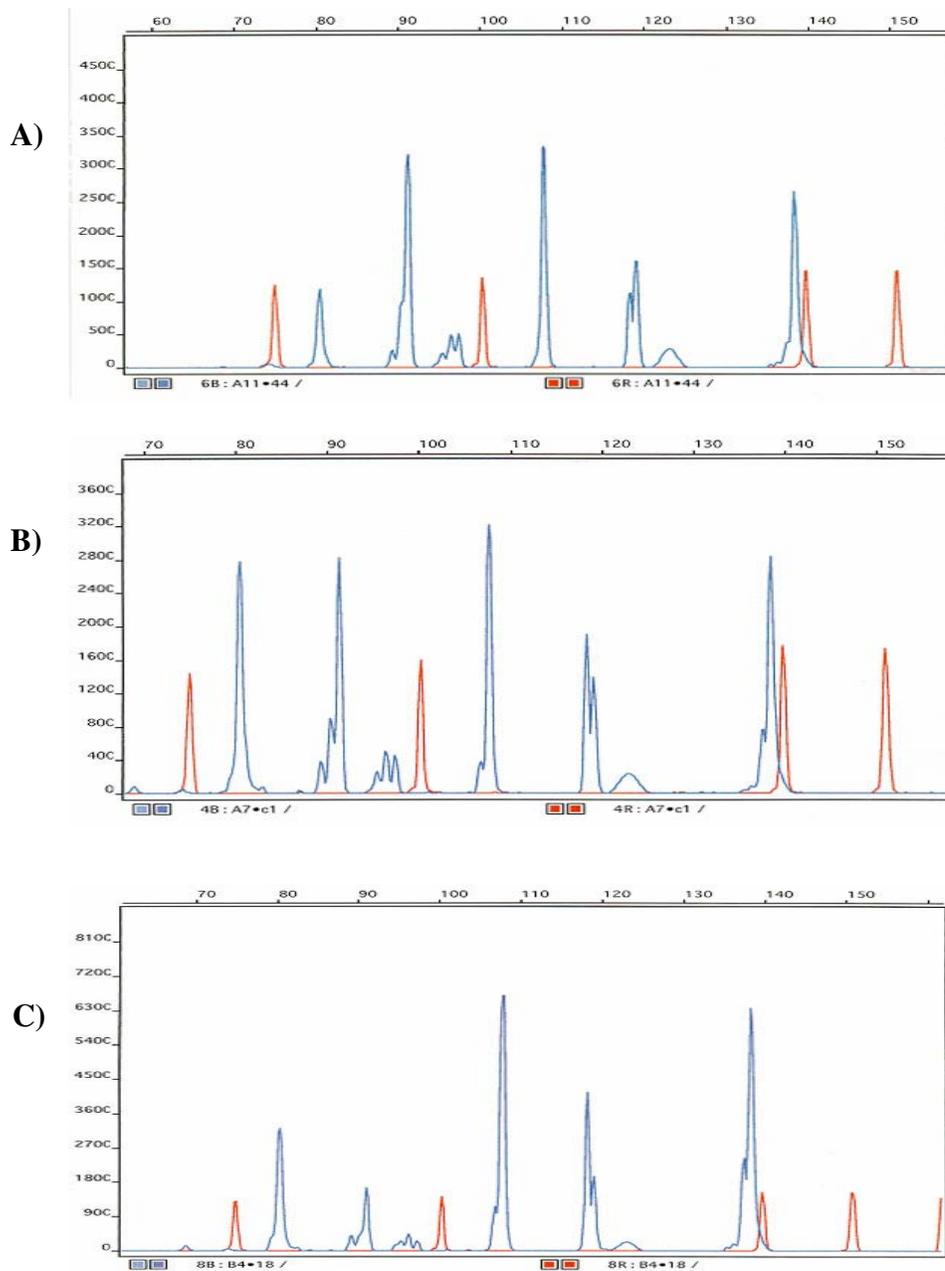


Figura 28- Eletroesferogramas do Multiplex I. Os picos vermelhos representam o marcador de peso molecular GS Rox 500. Os picos de cor azul representam os fragmentos amplificados da reação de multiplex I, que estão na ordem, da esquerda para a direita: beta; 5'delta; beta-actina; pseudo-beta; 3'gama a e epsilon. **A-** Deleção Siciliana; **B-** Controle; **C-** Amostra (PHHF-2). É possível identificar de forma visível a diferença na amplitude daqueles picos de apenas uma cópia (deleção), em comparação com o controle (De Andrade 2002).

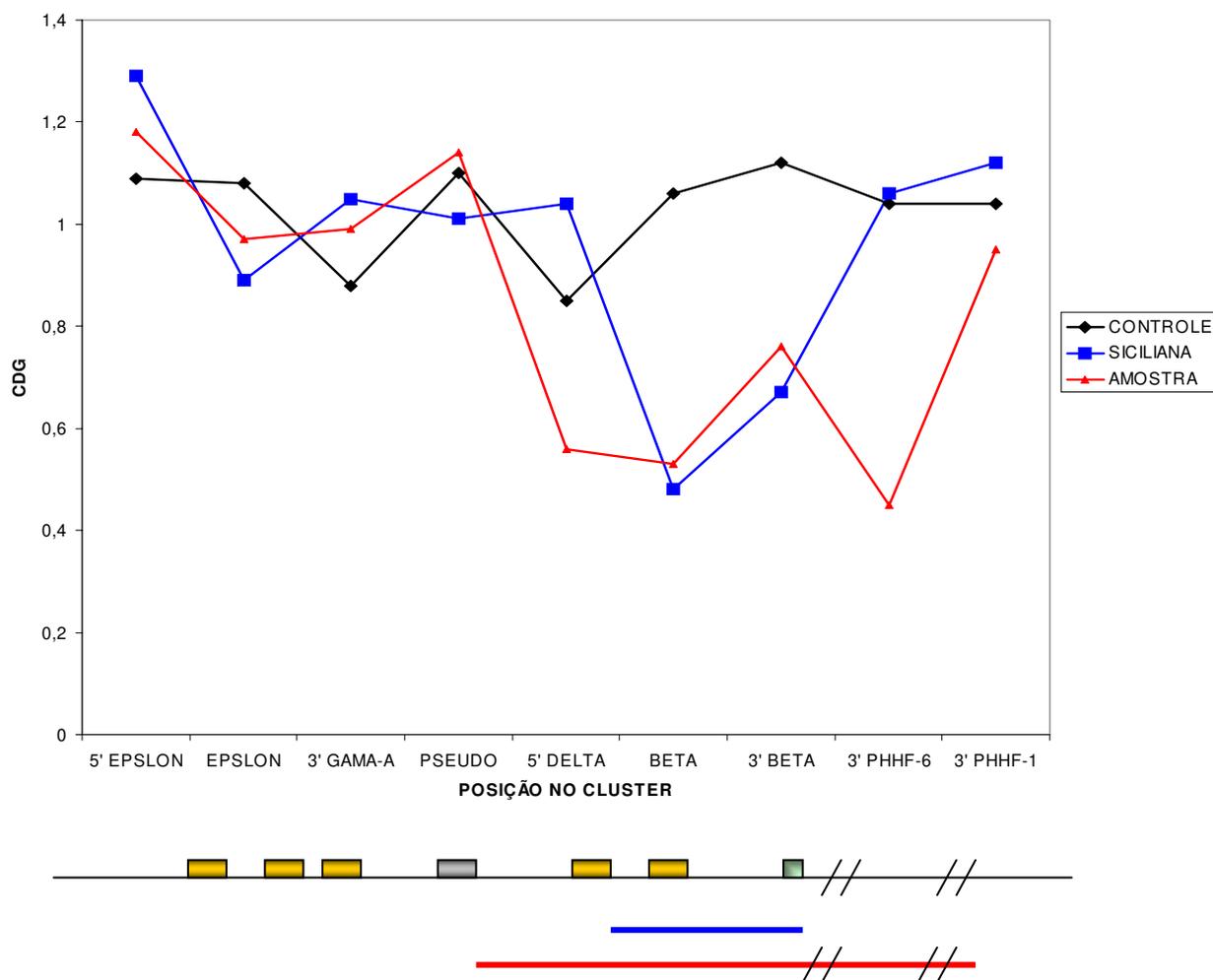


Figura 29- Gráfico comparando os dados do Coeficiente de Dosagem Gênica (referência-beta actina) de uma amostra controle, de uma deleção Siciliana e de um indivíduo estudado. Como esperado para a deleção Siciliana, as regiões beta e 3'beta apresentam valores baixos. A amostra analisada (PHHF-2) apresentou valores de deleção para 5'delta, beta, 3'beta e 3'PHHF-6. Também é mostrado, para efeito de ilustração, um esquema do tamanho e posição aproximada das deleções Siciliana (13 379 pb) e da amostra, de aproximadamente 105 Kb (De Andrade 2002).

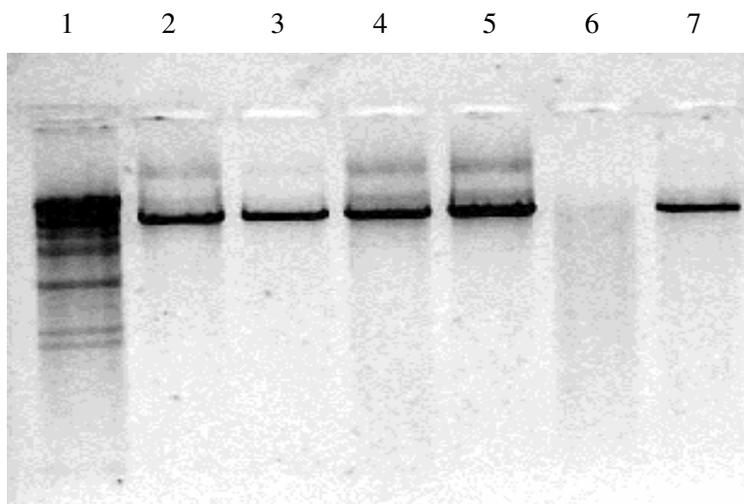


Figura 30- Imagem invertida do gel de agarose 0,9 %, corado com brometo de etídio, com fragmento contendo ponto de quebra. Os fragmentos, de aproximadamente 10Kb, foram amplificados com os primers E1 e HPF-1D e com enzima ELONGASE. Linha 1- marcador λ hind ; Linha 6- controle; Linhas 2-5 e 7- pacientes (De Andrade 2002).

A análise do sequenciamento revelou constituir-se da deleção encontrada em populações negras (PHHF-2), de aproximadamente 105 Kb. As seqüências obtidas foram comparadas com as do cluster da beta globina (GI-455025/GENE BANK) e da região correspondente ao ponto de quebra da PHHF-2 (GI-183848 e 32434/GENE BANK). A seqüência está representada na Figura 31, e apresenta as mesmas características já descritas para a região do ponto de quebra desta deleção, como a região de homologia (AGT), no sítio específico de recombinação.

**IDENTIFICATION OF NOVEL CANDIDATE GENES FOR GLOBIN
REGULATION IN ERYTHROID CELLS CONTAINING LARGE DELETIONS
OF THE HUMAN β -GLOBIN GENE CLUSTER**

Tiago Gomes de Andrade^a, Kenneth R. Peterson^b, Anderson F. Cunha^a, Luciana Sarmiento
Moreira^a, Sara T. O. Saad^a, Fernando Ferreira Costa^a.

^aHemocenter/Faculty of Medical Sciences - UNICAMP. Campinas, SP, Brazil

^bDepartment of Biochemistry and Molecular Biology - University of Kansas Medical
Center. Kansas City, KS, USA.

Corresponding author

Fernando Ferreira Costa

R. Carlos Chagas, 480 - Cidade Universitária Zeferino Vaz - Unicamp

CEP:13083-970; Campinas, SP - Brasil

Tel. + 55 - 19 - 3788 8734; Fax: 3289 1089

e-mail: ferreira@unicamp.br

Abstract

The genetic mechanisms underlying the continued expression of the γ -globin genes during the adult stage in deletional hereditary persistence of fetal hemoglobin (HPFH) and $\delta\beta$ -thalassemias are not completely understood. Herein, we investigated the possible involvement of transcription factors, using the SSH method to identify differentially expressed transcripts in reticulocytes from a normal and a HPFH-2 subject. Some of the detectable transcripts may participate in globin gene regulation. Quantitative RT-PCR experiments confirmed the down regulation of ZHX2, a transcriptional repressor, in two HPFH-2 subjects and in a carrier of the Sicilian $\delta\beta$ -thalassemia trait. The chromatin remodeling factors ARID1B and TSPYL1 had a very similar pattern of expression with an incremental increase in HPFH and decreased expression in $\delta\beta$ -thalassemia. These differences suggest a mechanism to explain the heterocellular and pancellular distribution of F cells in $\delta\beta$ -thalassemia and deletional HPFH, respectively. Interestingly, α -globin mRNA levels were decreased, similar to β -globin in all reticulocyte samples analyzed.

Key words: globin, HPFH, thalassemia, deletion, SSH, transcription, chromatin remodeling.

Introduction

The human globin genes comprise a multigenic family organized in two distinct clusters on chromosomes 16, for the α -like and chromosome 11 for the β -like globin genes. The first group includes the genes coding for the ζ , α and θ chains, while the second includes genes for ε , $A\gamma$ and $G\gamma$, δ and β globin chains. These genes are expressed during the different stages of erythroid development in the same linear order in which they are distributed on the chromosome [1]. These protein chains will assemble into hemoglobin Gower I ($\zeta_2\varepsilon_2$), Gower II ($\alpha_2\varepsilon_2$) and Portland ($\zeta_2\gamma_2$), in the embryo, Hb F ($\alpha_2\gamma_2$), in the fetal stage and Hb A1 ($\alpha_2\beta_2$) and Hb A2 ($\alpha_2\delta_2$), during adult life. It has been postulated that regulation of these genes depends upon the interaction of stage-specific transcription factors, gene-proximal *cis*-acting regions and an upstream, distal regulatory element known as the locus control region (LCR) [2].

Alterations in globin synthesis leads to globin chain imbalance and disorders known as thalassemias. Perturbations in the β -like globin genes, the β -thalassemias, result in an excess of α -like globin chains, that aggregate and precipitate causing injury to cells. $\delta\beta$ -thalassemias and hereditary persistence of fetal hemoglobin (HPFH) encompass two rare and heterogeneous groups of hemoglobin disorders with decreased or absent β -globin chains and a compensatory variable increase of γ -globin chains. $(\delta\beta)^\circ$ -thalassemias are characterized, in the heterozygous state, by increased levels of fetal hemoglobin ranging between 5% to 20% during adulthood, with a heterogeneous (heterocellular) pattern of HbF distribution [3]. This condition exhibits some of the thalassemic phenotype, such as microcytic and hypochromic red blood cells, and is caused by extensive variable deletions of the 3' end of the β -globin gene cluster. HPFH subjects have an abnormal switch from fetal to adult hemoglobin, resulting in high levels of Hb F in the adult stage, without clinical manifestations. They may arise from point mutations in the promoter region of the γ -globin genes or as a consequence of different 3' deletions of the β -like globin cluster [4]. The deletional forms of HPFH are associated with increased levels of γ -globin chains and a homogeneous (pancellular) distribution of HbF.

The genetic mechanisms underlying HPFH versus $\delta\beta$ -thalassemia are not completely understood. In fact, the discovery of modulators that regulate γ -globin gene transcription are the object of extensive research, since the occurrence of high levels of HbF concomitant with certain hemoglobinopathies, such as β -thalassemias and sickle cell disease, ameliorate these conditions [5,6]. For deletional HPFH, three main hypotheses were proposed to explain the relationship between these deletions and the non-suppression of γ - genes in the adult; 1- the removal of competitive regions that interact with the LCR [7]; 2- the juxtaposition of enhancer elements located downstream from the breakpoint region [8,9,10]; and 3- the removal of gene silencer elements [11,12,13]. However, despite evidences for these hypothesis, they are still not conclusive. Recently, Xiang et al. [14] developed a YAC construct containing the whole β -globin locus with the deletion responsible for the HPFH-2. Unexpectedly, the γ -globin gene was completely silenced in the adult transgenic mice [14]. This data suggest that other mechanisms may be responsible for the increased levels of HbF in these conditions. The authors speculate that other regions upstream to the β -globin cluster may harbor this activity.

Herein, we investigate the possible involvement of transcription factors, using a subtractive hybridization method [15] to identify differently expressed transcripts in reticulocytes from a normal and a HPFH-2 subject. This condition is caused by a large deletion of approximately 100 Kb, that completely removes the δ and β -globin genes, with the 5' breakpoint within the $\psi\beta$ region and the 3' breakpoint near an olfactory receptor gene downstream the β -globin cluster [9,16] (Figure 1). We also investigated the expression of identified genes by quantitative real-time PCR in a patient heterozygous for the Sicilian form of $\delta\beta$ -thalassemia [16] (Figure 1).

Results and Discussion

We utilized SSH to build two expression profiles for genes preferentially expressed in normal reticulocytes (NR) and those found with higher frequency in reticulocytes from a HPFH-2 subject (HR). This method allows the subtraction or exclusion

of commonly expressed genes and normalization of the libraries, favoring the detection of rare transcripts [15]. These characteristics were particularly useful in the analysis of gene expression in reticulocytes, since the majority of transcripts in these extremely specialized cells correspond to the globin genes.

We generated 541 sequences for the normal subtractive library (NL) and 349 sequences for the HPFH library (HL) with high homology scores (p value $>10^{-5}$) in BLAST alignments. Of these, 78 represented unique sequences in NL and 159 unique sequences in HL. We found 46 genes, 1 open reading frame (ORF), 8 hypothetical proteins and 7 expressed sequence tags (ESTs) differentially expressed in NL, which may represent transcripts down-regulated in HPFH cells. HL presented 89 annotated genes, 4 ORFs, 13 hypothetical proteins and 12 ESTs, corresponding to transcripts potentially up-regulated in HPFH-2. The identified genes correspond to many distinct functions, but those participating in transcription processes and potentially involved in globin gene regulation are highlighted in Table 2 for NL and Table 3 for HL, with chromosome location and absolute and normalized (clones per 1000) occurrences of each transcript. Supplementary Tables 1 and 2 display the complete list of transcripts represented in the Normal and HPFH libraries, respectively. Supplementary Table 3 shows genes found in both libraries and that may correspond to background, unsubtracted clones. We also detected 16 clones in NL and 42 clones in HL that match to genomic sequences (supplementary Tables 1 and 2). The lower redundancy found in HL compared to NL suggests that there is a larger proportion of genes activated or with higher expression in HPFH. The transcript most represented is SLC25A37, a member of the mitochondrial solute carrier family, appearing 356 and 78 times in NL and HL, respectively.

The genes corresponding to the transcripts found in both profiles are distributed widely in the genome, with the exception of the Y chromosome. Chromosomes 1 and 17 had the highest number of genes. However, Chromosome 17 is approximately three times smaller than chromosome 1, which is consistent with the fact that it has one of the highest genetic densities of the human chromosomes [17].

Some of the transcripts are described in Hembase (<http://hembase.niddk.nih.gov>), a repository of EST data developed from a SSH library representative of *in vitro* differentiated erythroid cells [18]. The detected genes are associated with diverse cellular functions. Erythroid specific genes as β - and γ -globins, glycophorin C, erythrocyte membrane protein band 4.1 and ferrochelatase were found in the libraries. In agreement with the loss of one β -globin allele in the persistence of HbF phenotype, two sequences corresponding to β -globin were detected in normal library (NL), while one clone from γ -globin was identified only in the HPFH library (HL). The low occurrence of globin genes indicates an adequate normalization of the libraries. Alpha globin was not detected.

Gene expression from globin genes (HBA, HBB and HBG), a transcription factor (ZHX2), three proteins with chromatin remodeling activities (ARID1B, TSPYL1 and NAP1L4), a redoxin related gene (SH3BGRL2) and a membrane protein (CD47) was evaluated by quantitative real time PCR (qRT-PCR) in reticulocytes. Data from Figure 2 are presented as fold changes from patient samples (two HPFH-2 individuals and a Sicilian $\delta\beta$ -thalassemia individual) relative to normal reticulocytes. Control (NR) and HR1 samples were the same used in the construction of the subtractive libraries. The expression of the studied genes was normalized using GAPDH and ACTB as constitutive, internal control genes.

Globin genes

Gene expression analysis of globins by qRT-PCR confirmed the diminution of β -globin in HPFH samples to approximately half the expression level of the normal reference and increased levels of γ -globin (up to 800 fold) (Figure 2). The levels of γ -globin in the Sicilian $\delta\beta$ -thalassemia sample were also elevated, although to a lesser degree (200-300 fold change) compared to the HPFH samples. This may be a consequence of the heterogeneous intercellular distribution of HbF instead of an actual difference of expression within individual cells. However, α -globin mRNA levels were decreased, similar to

β -globin in all reticulocyte samples analyzed. These data were confirmed by two independent experiments, each one performed with duplicate samples and expression levels normalized to two constitutive genes as controls. The decrease in α -globin expression may be the result of a common regulatory process affecting α -like and β -like globin synthesis. Since we evaluated gene expression in circulating mature reticulocytes and the primers used were designed to cross exon junctions, the RNA molecules studied correspond to mature, processed mRNAs in contrast to primary transcripts. Therefore, it is possible that the qRT-PCR data do not reflect an altered gene transcription, but rather post-transcriptional control mechanisms that affect the mRNA levels corresponding to both genes. Post-transcriptional mechanisms for globin gene regulation have been described [19], including a recent report on γ - and β -globin mRNA expression in Corfu $\delta\beta$ -thalassemia deletion cells [20]. In the Corfu deletion the β -globin gene is not removed and an alteration in α -globin was not detected. Therefore, the down-regulation of α -globin mRNA levels found in the HPFH and Sicilian $\delta\beta$ -thalassemia may be dependent upon the removal of the β -globin gene.

Alternately, gene transcription may be altered in these cells, as a result of perturbations of chromosome interactions, promoted by the removal of important regulatory regions located in the deleted area. Recent studies have demonstrated that, in addition to the fact that chromosomes are located in the nucleus in a non-random pattern, chromatin loops may extend beyond these territories and possibly interact with other regions of the genome, including different chromosomes [21]. Osborne et al. [22] demonstrated that transcriptionally active genes co-localize to sub-regions of genome, called transcription factories, at a frequency higher than would be expected by chance (some genes have a 60% co-localization frequency). Their study, using murine erythroid cells, demonstrates that even genes situated on different chromosomes may be found in these transcriptionally active regions. This is the case for the genes encoding α - and β -globin, localized in murine chromosomes 11 and 7, respectively.

On the short arm of human chromosome 16, where the α -globin cluster is located, we found three transcripts expressed in HL for genes potentially increased in these cells. Of particular interest is C16ORF35 (also denominated conserved gene telomeric to

α -globin cluster or -14 gene), an ORF of unknown function which has within intron 5, the HS 40 regulatory element, responsible for the regulation of the α -like globin genes [23]. Future studies confirming alterations in this gene and demonstrating associations between human chromosome 11 and 16 may elucidate how these regions are able to interact for the regulation of globin genes and the correct assembly of hemoglobin tetramers.

Transcription and Chromatin Remodeling Factors

The main goal of our study was to identify genes differentially expressed in cells heterozygous for large deletions of the human β -like globin cluster, possibly associated with the regulation of globin genes and the increased levels of γ -globin in HPFH. We found several transcription and chromatin remodeling factors; mainly in the HL profile (Tables 2 and 3). From these, we selected the chromatin remodeling factors ARID1B, TSPYL1 and NAP1L4, and the zinc finger and homeobox gene ZHX2, for qRT-PCR assay. ARID1B (also known as BAF250b) is a SWI-like protein, involved in ATP-dependent chromatin modification, and is potentially a subunit of the PYR complex, which is associated with a γ - δ intergenic region in the β -like globin cluster [24,25] (Figure 1). This region has been implicated in γ -globin silencing during the adult stage of development; its removal in deletional forms of HPFH may contribute to the reactivation of fetal globin [26]. The PYR complex contains Ikaros as its DNA binding subunit, proteins from the SWI/SNF transcription activator chromatin remodeling family and a member of the NuRD transcription repressor family, Mi-2 [27]. These features suggest that PYR maybe an important complex involved in globin gene switching [2]. The qRT-PCR experiments show an increase of ARID1B in two subjects heterozygous for HPFH-2, albeit at different levels (approximately a 5 and 2 fold change, respectively), confirming the data found in the subtracted libraries (Figure 2).

TSPYL1, a member of the NAP (Nucleosome Assembly Protein) family, is also associated with chromatin remodeling and is located in a chromosomal region (6q22-q23) frequently related to unlinked HPFH, although apparently adjacent to the polymorphic markers that delimit the 1.5 Mb region responsible for HPFH [28].

Interestingly, the expression of this gene followed a very similar pattern to ARID1B, with the same fold change in all samples analyzed, suggesting a common regulatory mechanism for these genes in erythroid cells. The other NAP protein, NAP1L4, is located in the same chromosome region (11p15.5) affected by the deletions under study in this work, approximately 2 Mb 3' to the β -globin cluster. NAP1L4 is a histone chaperone and alteration of its expression may disrupt expression of other genes [29]. However, our qRT-PCR experiments did not corroborate the diminution of NAP1L4 in HPFH-2 or Sicilian $\delta\beta$ -thalassemia reticulocytes (data not shown).

The increased expression of ARID1B and TSPYL1 could help in the maintenance of an open chromatin structure in the region comprising the fetal globin, leading to high levels of expression in the adult stage. However, their expression in Sicilian thalassemia, is decreased in qRT-PCR experiments approximately 0.15 fold, which may be a consequence of the fragments that are removed in HPFH-2 and maintained in Sicilian thalassemia, such as the PYR region, for example (Figure 1).

ZHX2 has zinc finger and homeobox domains and was categorized as a transcriptional repressor [30]. Interestingly, it is located in a genomic region (8q24.13) in which a quantitative trait locus (QTL) associated with a heterocellular HPFH, unlinked to chromosome 11, was recognized [31]. Moreover, a genetic alteration in the murine homologue *zhx2*, which is normally upregulated after birth, is responsible for the hereditary persistence of α -fetoprotein and H19 genes [32]. These proteins are normally expressed in the fetal liver, which is also the site of erythropoiesis during this stage of development. Two independent qRT-PCR experiments confirmed the downregulation of ZHX2 in HPFH and $\delta\beta$ -thalassemia (Figure 2). Moreover, its expression seems to be inversely correlated to γ -globin in the HPFH samples analyzed.

The candidate genes we studied by qRT-PCR and others we identified in the subtracted libraries may act synergistically to produce the HPFH phenotype. ZHX2 may be responsible, in part, for the silencing of the γ -globin genes at the adult stage of development and its suppression during this period may reactivate fetal globin. However, to consistently maintain this reactivation, the presence of other factors, such as ARID1B and TSPYL1,

may be necessary. Such a mechanism may provide the basis to explain the heterocellular and pancellular distribution patterns of F cells, found in $\delta\beta$ -thalassemia and deletional HPFH, respectively. If functional studies confirm the participation of ZHX2 in HbF repression, and ARID1B and TSPYL in the incremental of γ -globin, they become potential targets in strategies to reactivate γ -globin in hemoglobinopathies where higher levels of HbF would have beneficial effects.

Other transcription factors detected in our study include members of the Krüppel like zinc finger protein family (ZNF135, ZNF165 and ZNF286) and HOXB13, all found in the HL profile. Other members of the Krüppel family, such as KLF1 (EKLF) and KLF11 (FKLF), have been found to be involved in globin gene regulation [33,34]. HOXB13 encodes a transcription factor that belongs to the homeobox gene family, and is essential for vertebrate embryonic development. This protein was implicated in fetal skin development and cutaneous regeneration [35]. The authors demonstrated that there was substantially more HOXB13 expression detected in cultured fetal fibroblasts compared with cultured adult fibroblasts, which may imply that its function is more important at this stage of development. There are no reports of expression of these genes in hematopoietic tissues.

We also detected, in the HL profile, SMARCA4 (BRG1), a member of the SWI/SNF family of proteins that was described as one subunit of the PYR complex [27]. Notably, we also identified genes involved in DNA repair and recombination that potentially associate with SMARCA4 in protein complexes. These include FANCC, deficiency of which causes Fanconi anemia (FA), a genetic disorder that predisposes patients to hematopoietic failure, birth defects and cancer [36,37]; BLM, related to the RecQ subset of DexH box-containing DNA helicases and the determinant of Bloom syndrome, which is characterized by marked genomic instability [38]; and TOP3B, a DNA topoisomerase that controls and alters the topologic states of DNA during transcription and likely interacts with helicases of the RecQ family [39]. It would be interesting to investigate whether these genes have altered expression as a consequence of the deleted PYR binding site in β -globin cluster and/or if there is a relationship between altered expression and response mechanisms to DNA damage in these cells.

Genes related to other functions

Expression of SH3BGRL2 (downregulated in HPFH) and CD47 (upregulated in HPFH) related to other cellular process were also investigated by qRT-PCR. SH3BGRL2 is present at relatively high frequency in NL (supplementary Table 1). Although its function is not yet completely known, SH3BGRL2 may have redoxin activities [40], which may indicate that oxygen homeostasis is altered in HPFH, perhaps as a consequence of high concentrations of HbF in these cells. qRT-PCR confirmed the decreased expression of SH3BGRL2 in HPFH-2 and Sicilian $\delta\beta$ -thalassemia reticulocytes (Figure 2). This is the first description of the expression of these genes in erythroid cells suggesting that it may be important in the normal physiology of erythrocytes to protect against the harmful effects of oxygen. CD47 expression has been correlated with a prolonged survival of erythroid cells in the circulation [41]. It is a membrane protein and interacts with the SIRP α receptors in spleen macrophages. This association activates signaling pathways in these cells that inhibits phagocytosis. We confirmed its upregulation in HPFH-2 and an even higher level in Sicilian $\delta\beta$ -thalassemia (Figure 2 and supplementary Table 2). Perhaps the study of different forms of thalassemia may elucidate whether this protein has an important function in red cell survival with clinical consequences and establish its usefulness as a marker for disease severity. Some pro-apoptotic and anti-apoptotic genes were detected in the profiles, mainly in HL (supplementary Tables 1 and 2). These may reflect other cellular alterations promoted by the HPFH deletions. Future investigations might clarify whether these genes are related to the clinical characteristics and distinct hematological parameters found in HPFH and thalassemias.

Conclusions

Although transgenic animal models to characterize *cis*-regulatory regions, including the complete β -globin locus in yeast artificial chromosomes (YACs) [42], have been fundamentally important to elucidate the mechanism of gene regulation, they do not reproduce the entire genomic context in which the implicated genes evolved. The data available from YAC transgenic mice containing the HPFH-2 deletion [14] clearly demonstrate that simply removing the deleted region in the context of the whole locus

construct and juxtapositioning the 3' enhancer were not sufficient to reproduce the HPFH phenotype. The presence of transcription and chromatin remodeling factors with altered expression in HPFH may indicate that the involvement of trans-acting factors in the mechanisms of increased γ -globin in these cells are more important than previously supposed and have a decisive role in the consequent phenotype. Obviously, these trans-acting factors must interact with specific regions in the β -globin gene cluster, and possibly, under normal conditions, with those regions removed by the HPFH and thalassemia deletions. In addition, the removal of these regions may interfere with the expression of other genes and possibly is the cause of the observed alterations. Our data suggest an integrated model to explain the persistent, high levels of γ -globin in deletional HPFH and $\delta\beta$ -thalassemias that unifies the role of *cis* -regulatory elements, such as promoters, silencers, enhancers and the LCR, with transcription and chromatin remodeling factors in perturbing expression in cells containing these genetic deficiencies (Figure 3).

Part of the altered gene expression detected within the subtracted libraries may be an indirect effect due to the increased concentration of HbF in the cells. However, genetic alterations similar to those presented in this work may be implicated in the perturbation of chromosome interactions and/or disruption of the expression of regulatory non-coding transcripts [43] with functional consequences for cellular gene expression. Additionally, other genes may have modifications of their expression, but were not detected in our experiments. To our knowledge, this is the first description of wide-ranging gene expression alterations in cells containing deletional HPFH and thalassemia. Finally, the dissection of the differential expression data presented in this study may help elucidate important pathways of erythroid differentiation and maintenance of red cells in peripheral circulation, as well as aid in the understanding of genetic mechanisms involved in erythrocyte pathologies such as sickle cell disease and thalassemias.

Materials and Methods

Subjects, DNA analysis and hematological data

Molecular identification of the HPFH-2 and Sicilian $\delta\beta$ -thalassemia deletion was carried out using a combination of GAP PCR [44], a fluorescent gene dosage method, and sequencing, as previously described [45,46]. We selected two non-related subjects heterozygous for HPFH-2 and a patient carrier of the Sicilian form of $\delta\beta$ -thalassemia, with high levels of HbF. Hematological parameters from subjects are described in Table 1. As control, we used samples from an individual with normal hematological indices. Hematological data and red blood cell (RBC) levels were determined electronically (Coulter Counter S. Sr.). The quantitation of HbA₂ and HbF levels was performed using the Variant II Hemoglobin Testing System (Bio-Rad, CA-USA). Hemoglobin electrophoresis was performed on cellulose acetate with Tris-boric acid-EDTA buffer (TBE) at pH 8.9. The HbF distribution in RBCs was evaluated by the Kleihauer technique [47], and globin analysis was performed on UREA-TRITON-polyacrylamide gels [48].

RNA isolation

Reticulocytes were isolated as previously reported [49]. Peripheral blood cells were centrifuged to remove the plasma. Erythrocytes were then lysed with red blood cell lysis solution (0.144 M NH₄Cl, 0.01 M NH₄HCO₃) and the samples centrifuged. The remaining pellet containing leukocytes was discarded and the supernatant homogenized with a 1/10 dilution of a Sucrose/KCl solution (1.5 M C₁₂H₂₂O₁₁, 0.15 M KCl). After centrifugation, supernatant containing reticulocytes was treated with 800 μ l of 10% acetic acid and centrifuged again. The pellet was then resuspended in 1 ml of Trizol[®] (Invitrogen, USA) reagent for each 5-10 ml of initial blood. RNA extraction was carried out according to manufacturer's instructions and resuspended in small amounts (5-15 μ l) of DEPC-treated water. All centrifugations were performed in a Sorvall Centrifuge (rotor SLA600TC), 5.000 g, at 4°C.

Construction of Suppression Subtractive Hybridization (SSH) Libraries

cDNA was amplified using the BD SMART PCR cDNA Synthesis kit (BD Biosciences). Synthesis of the second strand was performed with Platinum High Fidelity Polymerase (Invitrogen, USA) and 5' PCR Primer IIA (BD Biosciences, USA). PCR products were detected after 25 cycles and 33 cycles was the optimum determined for reticulocytes in our experiments. Double-stranded cDNA was digested with *RsaI* and used for the construction of SSH [15]. We prepared the libraries in forward and reverse directions using cDNA from normal and HPFH-2 reticulocytes, following the manufacturer's recommendations for the PCR-Select cDNA Subtraction kit (Clontech, CA, USA). Nested PCR products were ligated into the pGEMT vector (Promega, USA) and transformed into *E. coli* DH5 α competent cells. Individual clones were transferred to 96 well plates containing selective LB medium and incubated overnight. Two μ l were subjected to PCR using M13 primers; the products of amplification were sequenced using MegaBACE (GE Healthcare). The sequences were analyzed with BLASTn software (available at www.ncbi.nlm.nih.gov/BLAST).

Real Time PCR

Quantitative real time PCR (qRT-PCR) was performed using an ABI 5700 Sequence Detector System $\text{\textcircled{R}}$ (Applied Biosystems, USA), samples were analyzed in duplicate for each experiment. Ten ng of reticulocytes's cDNA samples were used with SYBRGreen PCR Master Mix $\text{\textcircled{R}}$ (Applied Biosystems, USA) for detection of amplified products. PCR primers were designed using Primer Express (Applied Biosystems, USA) software to specifically amplify mRNA from the gene of interest, as follows: α -globin (HBA) forward (F) 5'-TGGTCCCCACAGACTCAGAGA - 3', reverse (R) 5'-CGGCCTTGACGTTGGTCTT-3'; β -globin (HBB) F 5'-AGGCACCGAGCACTTTCTTG-3', R 5'-ATCTGTCCACTCCTGATCCAGTT-3'; γ -globin (HBG) F 5'-CATGGCAAGAAGGTGCTGACT-3', R 5'-GCAAAGGTGCCCTTGAGATC-3'; ARID1B F 5'-TCCAAAGATATCGTTACTCCTGAGG-3', R 5'-

TTAATAGTGTCCAAAGCCCACGTAC-3'; TSPYL1 F 5'-
AAAGGTGCCCTAAGCAGGAGA-3', R 5'-GCTCTTAACGACAGCACAGTCTTC-3';
NAP1L4 F 5'-GGAGGGAGAAGACGAGGATGA-3', R 5'-
GCTGGCTGGGTTCTTCTAGA-3'; ZHX2 F 5'-GAACTGGCTGAATCAGACTCCG-
3', R 5'-CACAGCAGTTCTAACAGACTTCCCT-3'; SH3BGRL2 F 5'-
GCGACCGATACTGTGGAGATTAT-3', R 5'-
CATCTTCCACTCTTCTTCTCTAAGGTTTC-3'; β -actin (ACTB) F 5'-
AGGCCAACCGCGAGAAG-3', R 5'-ACAGCCTGGATAGCAACGTACA-3';
glyceraldehyde-3-phosphate dehydrogenase (GAPDH) F 5'-
GCACCGTCAAGGCTGAGAAC-3', R 5'-CCACTTGATTTTGGAGGGATCT-3'.

Gene expression was calculated according to Vandensompele et al. [50]. Basically, the mean threshold cycle (Ct) is subtracted from the minimal mean value between samples for each gene analyzed, obtaining the ΔCt . Next, the Q value ($2^{\Delta Ct}$) is calculated based on the amplification efficiency for each primer set. Then, the Normalization Factor (NF) is calculated, which represents the geometric mean between the Q values from the studied gene and control genes (ACTB and GAPD) for each sample. Normalized expression values are obtained by dividing Q for the studied gene by NF. The final data are presented as the ratio (fold change) between these normalized values for each sample (HPFH reticulocytes or $\delta\beta$ -thalassemia reticulocytes) and the reference sample (normal reticulocytes).

Acknowledgments

This research was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Hemocenter (Unicamp). We thank Dulcinéia Albuquerque for analysis of real-time PCR, Tiago Ferraz and Genomic and Gene Expression Lab (Unicamp) for bioinformatics and Mateus Andrade for helping with figures.

References

- [1] G. Stamatoyannopoulos, Molecular and Cellular Basis of Hemoglobin Switching, in: M. H. Steinberg, B. G. Forget, D. R. Higgs, R. L. Nagel (Eds.), Disorders of Hemoglobin, Cambridge University Press, Cambridge, 2001, pp 131-145.
- [2] A. Bank, Regulation of human fetal hemoglobin: new players, new complexities, Blood (2005)
- [3] D.J. Weatherall, J.B. Clegg, The Thalassemia Syndromes, Oxford, Blackwell Science, Oxford, 2001.
- [4] B. G. Forget, Molecular basis of hereditary persistence of fetal hemoglobin. Ann N Y Acad. Sci. 850 (1998) 38-44.
- [5] H.F BUNN, Pathogenesis and treatment of sickle cell disease, New England Journal of Medicine, 377 (1997) 762-769.
- [6] S.M JANE, Understanding fetal globin gene expression: a step towards effective Hb F reactivation in haemoglobinopathies, Bri. J. Haematology. 102 (1998) 415-422.
- [7] T.M. Townes, R.R. Behringer, Human globin locus activating region: role in temporal control, Trends Genet. 6 (1990) 219-223.
- [8] M.O. Arcasoy, High levels of human γ -globin gene expression in adult mice carrying a transgene of deletion-type hereditary persistence of fetal hemoglobin, Mol. Cell Biol. 17 (1997) 2076-2089.
- [9] E.A Feingold, L.A Penny, A.W. Nienhuis, B.G. Forget, An olfactory receptor gene is located in the extended human β -globin gene cluster and is expressed in erythroid cells, Genomics 61 (1999) 15-23.
- [10] E.Z. Katsantoni et al., Persistent gamma-globin expression in adult transgenic mice is mediated by HPFH-2, HPFH-3, and HPFH-6 breakpoint sequences, Blood. 102 (2003) 3412-9.

- [11] T. Huisman et al. The present status of the heterogeneity of fetal hemoglobin in β -thalassemia: an attempt to unify some observations in thalassemia and related conditions, *Ann. N Y Acad Sci* 232 (1974) 107-124.
- [12] M. Vitale, Evidence for a globin promoter-specific silencer element located upstream of the human δ -globin gene, *Biochemical and Biophysical Research Communications* 204 (1994) 413-418.
- [13] R. Calzolari, T. Mcmorrow, N. Yannoutsos, A. Langeveld, F. Grosveld, Deletion of a region that is a candidate for the difference between the deletion forms of hereditary persistence of fetal hemoglobin and $\delta\beta$ -thalassemia affects β - but not γ -globin gene expression, *EMBO J.* 18 (1999) 949-958.
- [14] P. Xiang et al, Juxtaposition of the HPFH2 enhancer is not sufficient to reactivate the gamma-globin gene in adult erythropoiesis, *Hum Mol Genet.* 2005 (in press).
- [15] L. Diatchenko et al, Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries, *Proc Natl Acad Sci U S A* 93 (1996) 6025-30.
- [16] P. Henthorn, O. Smithies, D.L. Mager, Molecular analysis of deletions in the human β -globin gene cluster: deletion junctions and locations of breakpoints, *Genomics.* 6 (1990) 226-237.
- [17] J.C. Venter et al., The sequence of the human genome, *Science* 291 (2001) 1304-51.
- [18] A. Gubin, J.M. Njoroge, G.G. Bouffard, J.L. Miller, Gene expression in proliferating human erythroid cells, *Genomics* 59 (1999) 168-77.
- [19] J. E. Russel, J. Morales, S.A. Liebhaber, The role of mRNA stability in the control of globin gene expression, *Prog. Nucleic Acids Res. Mol. Biol* 57 (1997) 249-287.
- [20] L. Chakalova et al, The Corfu deltabeta thalassemia deletion disrupts gamma-globin gene silencing and reveals post-transcriptional regulation of HbF expression. *Blood.* 105 (2005) 2154-60.

- [21] C.G. Spilianakis, M.D. Lalioti, T. Town, G.R. Lee, R.A. Flavell, Interchromosomal associations between alternatively expressed loci, *Nature*. 35 (2005) 637-45.
- [22] C.S. Osborne, Active genes dynamically colocalize to shared sites of ongoing transcription, *Nat Genet*. 36 (2004) 1065-71.
- [23] P. Vyas, M.A. Vickers, D.J. Picketts, D.R. Higgs, Conservation of position and sequence of a novel, widely expressed gene containing the major human alpha-globin regulatory element, *Genomics* 29 (1995) 679-89.
- [24] X. Wang, et al. Two related ARID family proteins are alternative subunits of human SWI/SNF complexes, *Biochem J*. 383 (2004) 319-25.
- [25] Z. Nie et al, A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex, *Mol Cell Biol*. 20 (2000) 8879-88.
- [26] D. O'Neill et al, Tissue-specific and developmental stage-specific DNA binding by a mammalian SWI/SNF complex associated with human fetal-to-adult globin gene switching, *Proc. Natl. Acad. Sci. USA*. 96 (1999) 349-354.
- [27] D. O'Neill et al. An Ikaros-containing chromatin-remodeling complex in adult-type erythroid cells. *Mol. Cell. Biol.*, 20 (2000) 7572-7582.
- [28] J. Close et al., Genome annotation of a 1.5 Mb region of human chromosome 6q23 encompassing a quantitative trait locus for fetal hemoglobin expression in adults, *BMC Genomics*. 5 (2004) 33.
- [29] P. Rodriguez et al., Functional characterization of human nucleosome assembly protein-2 (NAP1L4) suggests a role as a histone chaperone, *Genomics*. 44 (1997) 253-65.
- [30] H. Kawata et al., Zinc-fingers and homeoboxes (ZHX) 2, a novel member of the ZHX family, functions as a transcriptional repressor, *Biochem J*. 373 (2003) 747-57.

- [31] C. Garner et al., Quantitative trait locus on chromosome 8q influences the switch from fetal to adult hemoglobin, *Blood*. 104 (2004) 2184-6.
- [32] S. Perincheri, R.W.C. Dingle, M.L. Peterson, B.T. Spear, Hereditary persistence of alpha-fetoprotein and H19 expression in liver of BALB/cJ mice is due to a retrovirus insertion in the *Zhx2* gene, *Proc Natl Acad Sci U S A*. 102 (2005) 396-401.
- [33] A.C. Perkins, K.M. Gaensler, S.H. Orkins, Silencing of human fetal globin expression is impaired in the absence of the adult beta-globin gene activator protein EKLF, *Proc Natl Acad Sci U S A*. 93 (1996) 12267-71.
- [34] G. Suske, E. Bruford, S. Philipsen, Mammalian SP/KLF transcription factors: bring in the family, *Genomics*. 85 (2005) 551-6.
- [35] E.J. Stelnicki, J. Arbeit, D.L. Cass, C. Saner, Harrison, M.; Largman, C.: Modulation of the human homeobox genes PRX-2 and HOXB13 in scarless fetal wounds, *J. Invest. Derm.* 111 (1998) 57-63.
- [36] M. Chen et al., Inactivation of *Fac* in mice produces inducible chromosomal instability and reduced fertility reminiscent of Fanconi anaemia, *Nature Genet.* 12 (1996) 448-451.
- [37] T. Otsuki et al., Fanconi anemia protein, FANCA, associates with BRG1, a component of the human SWI/SNF complex, *Hum. Molec. Genet.* 10 (2001): 2651-2660.
- [38] A.R. Meetei et al., A multiprotein nuclear complex connects Fanconi anemia and Bloom syndrome, *Molec. Cell. Biol.* 23 (2003): 3417-3426.
- [39] L. Wu, I.D. Hickson, The Bloom's syndrome helicase stimulates the activity of human topoisomerase III alpha, *Nucleic Acids Res.* 30 (2002) 4823-9.
- [40] M. Mazzocco et al., The identification of a novel human homologue of the SH3 binding glutamic acid-rich (SH3BGR) gene establishes a new family of highly conserved small proteins related to Thioredoxin Superfamily, *Gene*. 291 (2002) 233-9.

- [41] P.A. Oldenburg et al, Role of CD47 as a marker of self on red blood cells, *Science* 288 (2000) 2051-4.
- [42] K.R. Peterson, Use of yeast artificial chromosomes (YACs) in studies of mammalian development: production of β -globin locus YAC mice carrying human globin developmental mutants, *Proc. Natl. Acad. Sci. USA.* 92 (1995) 5655-5659.
- [43] K.E. Plant, S.J. Routledge, N.J. Proudfoot, Intergenic transcription in the human beta-globin gene cluster, *Mol Cell Biol.* 21 (2001) 6507-14.
- [44] J.E. Craig, R.A. Barnetson, J. Prior, J.L. Raven, S.L. Thein, Rapid detection of deletions causing $\delta\beta$ Thalassemia and Hereditary Persistence of Fetal Hemoglobin by enzymatic amplification, *Blood.* 86 (1994) 1673-1682.
- [45] T.G. De Andrade, A. Fattori, S.T.O. Saad, M.F. Sonatti, F.F. Costa, Molecular identification of Sicilian ($\delta\beta$)thalassemia associated with β -thalassemia and hemoglobin S in Brazil, *Braz J Med Biol Res.* 35 (2002) 873-6.
- [46] T.G. De Andrade, S.T.O. Saad, M.F. Sonatti, F.F. Costa. Simple fluorescent PCR method for detection of large deletions in the beta-globin gene cluster, *Am J. Hematol.* 72 (2003) 225-7.
- [47] J.V. Dacie, S.M. Lewis, *Practical haematology*, Churchill Livingstone, Edinburgh, London, Mulbourne, New York, 1984.
- [48] B.P. Alter, S.C. Goll, G.D. Frenov, E. Gravely, T.H.J. Huisman, Globin chain electrophoresis: a new approach to the determination of the $G\gamma/ A\gamma$ ratio of globin synthesis, *Br. J. Haematology* 44 (1980) 525-534.
- [49] M. Goossens, Y.M. Kan, DNA analysis of hemoglobin disorders, *Methods Enzymol* 76 (1981) 805.
- [50] J. Vandesompele et al., Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, *Genome Biol.* 3 (2002) 34.

Figure 1- Representation of the human β -globin gene cluster. Arrows indicate DNase I hypersensitive sites. Genes are shown as thick arrows in the line. Black bars represent the extent of the described deletions. PYR binding region is also indicated.

Figure 2- Gene expression pattern of selected genes in reticulocyte samples. NR: normal reticulocytes; HR: HPFH-2 reticulocytes; TR: $\delta\beta$ -thalassemia reticulocytes. Bars depict fold change relative to reference sample (NR). Relative gene expression was obtained and normalized using two constitutive genes as internal controls (ACTB and GAPDH). Data presented correspond to the geometric mean between these controls. Gene expression data from HBA and ZHX2 are presented as the average from two independent experiments.

Figure 3- Model to explain the effect of different deletions within the β -globin cluster. Under normal conditions (A), the LCR interacts with the δ and β -globin genes during the adult stage of development, perhaps involving the PYR complex. Various HPFH and thalassemia deletions may interfere with the expression of trans-acting factors, via disruption of interchromosomal interactions or decreased expression of regulatory non-coding RNAs. Intrachromosomal interactions are also perturbed in this model. These cis and trans alterations may alter the pattern of LCR, promoter, enhancer and transcription factor interactions that regulate gene expression in the β -globin cluster. In addition, different transcription and/or chromatin remodeling factors may synergistically interact with the remaining regions in the loci to produce a random ($\delta\beta$ -thalassemia) (B) or stable (HPFH) (C) patterns of γ -globin expression.

Expression of new red cell related genes in erythroid differentiation

Tiago Gomes de Andrade*, Luciana Sarmiento Moreira, Adriana dos Santos Duarte, Carolina Lanaro, Dulcinéia Martins de Albuquerque, Sara T. O. Saad, Fernando Ferreira Costa.

Hemocenter/Faculty of Medical Sciences - UNICAMP. Campinas, SP, Brazil

Adress for correspondence

R. Carlos Chagas, 480 - Cidade Universitária Zeferino Vaz - Unicamp

CEP: 13083-970 - Campinas, SP - Brasil.

Tel.: + 55 - 19 – 3788 8734; Fax: 3289 1089

e-mail: tiago@unicamp.br

Abstract

Red Blood Cells (RBCs) are extremely differentiated cells containing hemoglobin tetrameres responsible for the delivery of oxygen to the different tissues of the body. In humans, globin genes are organized in two distinct clusters on chromosomes 16, for the α -like and chromosome 11 for the β -like globin genes. Using suppression subtractive hybridization (SSH) method, we have previously identified genes differentially expressed in erythroid cells heterozygous for large deletions in β -like globin cluster. Herein, we investigated the expression of four new red cell related genes in erythroid differentiation. The expression of the chromatin remodeling factors *ARID1B* and *TSPYL1* presented similar patterns, with an upregulation after Erythropoietin (EPO) addition, similar to previous data found in reticulocytes, which suggest a common regulatory mechanism for these genes. *ZHX2*, a transcriptional repressor, was downregulated, while a redoxin related gene, *SH3BGRL2*, had higher levels of expression upon differentiation. These are the first investigation of the mentioned newly described genes in erythroid differentiation and demonstrate that the expression of these genes is affected by EPO stimuli. These genes may participate in globin regulation and may be important in the normal physiology of erythrocytes.

Key words: globin, transcription, erythroid differentiation, chromatin remodeling, oxidative stress.

1- Introduction

Red Blood Cells (RBCs) or erythrocytes are the circulating blood cells responsible for the delivery of oxygen to the different tissues of the body and removal of carbon dioxide. These are extremely differentiated cells in which the majority of proteins synthesized comprise hemoglobin, heterotetrameres composed by two subunits related to the α -globin gene subfamily and two subunits related to the β -globin gene subfamily. Globin polypeptides bind heme, which in turn allows hemoglobin in RBC to bind oxygen reversibly. In humans, globin genes are organized in two distinct clusters on chromosomes 16, for the α -like and chromosome 11 for the β -like globin genes. The first group includes the genes coding for the ζ , α and θ chains, while the second includes genes for ϵ , $A\gamma$ and $G\gamma$, δ and β globin chains. These genes are expressed during the different stages of erythroid development in the same linear order in which they are distributed on the chromosome. These protein chains will assemble into hemoglobin Gower I ($\zeta_2\epsilon_2$), Gower II ($\alpha_2\epsilon_2$) and Portland ($\zeta_2\gamma_2$), in the embryo, Hb F ($\alpha_2\gamma_2$), in the fetal stage and Hb A1 ($\alpha_2\beta_2$) and Hb A2 ($\alpha_2\delta_2$), during adult life (Stamatoyannopoulos et al, 2001).

By suppression subtractive hybridization (SSH) method (Diatchenko et al, 1996), we have previously identified genes differentially expressed in reticulocytes, morphologically identifiable cells immediately prior to RBCs during erythroid differentiation, from subjects heterozygous for $\delta\beta$ -thalassemia and deletional hereditary persistence of fetal hemoglobin (HPFH). These conditions are characterized by large deletions of variable extents of the human β -like globin cluster, with decreased or absent β -globin chains and a compensatory variable increase of γ -globin chains (Weatherall and Clegg, 2001; Forget, 1998). Some of the detectable transcripts may participate in globin gene regulation and RBC function. Herein, we investigated the expression of four new red cell related genes in erythroid differentiation. The selected genes encode two proteins with chromatin remodeling activities (ARID1B and TSPYL1), a transcription factor (ZHX2), and a redoxin related protein (SH3BGRL2).

2- Materials and Methods

2.1- Cell culture

Blood from normal volunteers was cultured using a two-phase liquid culture procedure, as previously described (Pope et al. 2000). Mononuclear cells were isolated from the peripheral blood samples by centrifugation in a Ficoll-Hypaque gradient and cultured for 7 days (Phase I) in IMDM medium (Invitrogen), supplemented with 20% fetal calf serum (Invitrogen), 1ug/ml cyclosporin A (Sandoz), and 10% conditioned medium collected from the culture of human bladder carcinoma 5637 cell line. Cells were incubated at 37°C, in an atmosphere of 5% CO₂ and 92% extra humidity. After 7 days, the non-adherent cells were harvested and re-cultured in phase II medium, IMDM supplemented with 30% fetal calf serum (Invitrogen), 1% deionized bovine serum albumin (BSA-Sigma), 10⁻⁵M 2-mercaptoethanol (Sigma), 1.5nmol/l glutamine (Invitrogen), 300µg iron saturated transferrin (Sigma) 10⁻⁶M dexamethasone, 5ng/ml human stem cell factor (SCF-Calbiochem), 1U/ml human recombinant erythropoietin (EPO - Cilag), 2.5ug/ml funzigone (Invitrogen), 50ug/ml streptomycin (Invitrogen) and 25ug/ml glutamicin (Invitrogen). Cell samples were collected from phase II cultures at 0h (before EPO) and 264hs (11 days after EPO). Cell numbers and viability were determined by Trypan Blue exclusion. Samples of 5x10⁶ cells were pelleted, resuspended in Trizol (Invitrogen) and stored at -80°C for total RNA extraction and cDNA synthesis. For morphological cell analyses, cytopsin slides were prepared and were stained with Leishman before examining with an Eclipse E-600 microscope (Nikon) using the Image Pro-Express 4.0 software (Media Cybernetic, LP) to morphologically analyze the cell differentiation stage.

2.2- Flow Cytometry analysis

Expression of HbF by flow cytometry was performed using anti-HbF antibody. The cells were washed with PBS and suspended in 1% paraformaldehyde prior to analysis.

Data from 10,000 events were acquired for analysis using Cell Quest Software (Becton Dickinson). Non-specific staining was established from the corresponding isotype control and subtracted from the corresponding positive population percentage.

2.3- RNA extraction and cDNA synthesis

Total RNAs were isolated using Trizol Reagent (Invitrogen) according to manufacturer's instructions. Two μg RNA samples were incubated with 1 U DNaseI (Invitrogen) for 15 min at room temperature and EDTA was added to a final concentration of 2 mM to stop the reaction. The enzyme was subsequently inactivated for 10 min at 65 °C. The DNaseI-treated RNA samples were reverse transcribed with SuperScript III (Invitrogen) for 50 min at 42 °C. Two U of RNaseH (Invitrogen) were subsequently added and the samples were incubated at 37 °C for 20 min. The cDNA samples were quantified using the UV spectrophotometer, GeneQuant (Pharmacia).

2.4- Real Time PCR

Quantitative real time PCR (qRT-PCR) was performed using an ABI 5700 Sequence Detector System ® (Applied Biosystems, USA), samples were analyzed in duplicate for each experiment. Five ng of cDNA from cell culture samples were used with SYBRGreen PCR Master Mix ® (Applied Biosystems, USA) for detection of amplified products. PCR primers were designed using Primer Express (Applied Biosystems, USA) software to specifically amplify mRNA from the gene of interest, as follows: β -globin (*HBB*) F 5'-AGGCACCGAGCACTTTCTTG-3', R 5'-ATCTGTCCACTCCTGATCCAGTT-3'; γ -globin (*HBG*) F 5'-CATGGCAAGAAGGTGCTGACT-3', R 5'-GCAAAGGTGCCCTTGAGATC-3'; *ARID1B* F 5'-TCCAAAGATATCGTTACTCCTGAGG-3', R 5'-TTAATAGTGTCCAAAGCCCACGTAC-3'; *TSPYL1* F 5'-AAAGGTGCCCTAAGCAGGAGA-3', R 5'-GCTCTTAACGACAGCACAGTCTTC-3';

ZHX2 F 5'-GAACTGGCTGAATCAGACTCCG-3', R 5'-
CACAGCAGTTCTAACAGACTTCCCT-3'; *SH3BGRL2* F 5'-
GCGACCGATACTGTGGAGATTAT-3', R 5'-
CATCTTCCACTCTTCTTCTCTAAGGTTC-3'; β -actin (*ACTB*) F 5'-
AGGCCAACCGCGAGAAG-3', R 5'-ACAGCCTGGATAGCAACGTACA-3';
glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) F 5'-
GCACCGTCAAGGCTGAGAAC-3', R 5'-CCACTTGATTTTGGAGGGATCT-3'.

Gene expression was calculated according to Vandensompele et al, 2002. Basically, the mean threshold cycle (Ct) is subtracted from the minimal mean value between samples for each gene analyzed, obtaining the ΔCt . Next, the Q value ($2^{\Delta Ct}$) is calculated based on the amplification efficiency for each primer set. Then, the Normalization Factor (NF) is calculated, which represents the geometric mean between the Q values from the studied gene and control genes (*ACTB* and *GAPDH*) for each sample. Normalized expression values are obtained by dividing Q for the studied gene by NF. The final data are presented as the ratio (fold change) between these normalized values for each sample (cells 11 days after EPO) and the reference sample (cells before EPO).

3- Results and Discussion

3.1- Gene expression of globins and new red cell related genes in erythroid differentiation

We have previously identified genes differentially expressed in reticulocyte cells heterozygous for large deletions of the human β -like globin cluster, possibly associated with the regulation of globin genes and the normal physiology of RBCs. Some of the detected genes were not described in this tissue before and were selected here for investigation in an *in vitro* model for erythroid differentiation (Fibach et al, 1989; Pope et al, 2000). We studied gene expression from two genes with

chromatin remodeling activities (*ARID1B* and *TSPYL1*), a transcription factor (*ZHX2*), and a redoxin related gene (*SH3BGRL2*). Gene expression from globin genes (*HBA*, *HBB* and *HBG*) was also evaluated. We cultured peripheral blood mononuclear cells from a normal individual, and collected cells for RNA extraction before (0h) and 11 days after (216hs) Erythropoietin (EPO) stimulus. At this point in culture, we found cells at different stages of maturation (Pope et al, 2000) as revealed by cytospin (data not shown), with a large proportion of cells expressing high levels of Hb F, as shown by flow cytometry analysis (Figure 1).

Expression data from Figure 2 are presented as fold changes from cells under erythroid differentiation relative to undifferentiated cells. The expression of the studied genes was normalized using *GAPDH* and *ACTB* as constitutive, internal control genes. Gene expression analysis of globins by qRT-PCR confirmed the increased expression of β -globin and γ -globin after EPO addition. The expression of γ -globin was moderately higher, which may reflect a predominant proportion of cells at early stages of maturation, although in this system there is no real switch between fetal and adult globins (Pope et al, 2000).

3.2- Chromatin Remodeling factors

We showed previously that the expression of the chromatin remodeling factors *ARID1B* and *TSPYL1* followed a very similar pattern, with the same fold change in all reticulocyte samples analyzed. This pattern was retained in the primary erythroid culture, with an increment of approximately 5 fold after 216 hours of differentiation, suggesting a common regulatory mechanism for these genes in cells (Figure 2).

ARID1B is a SWI-like protein, involved in ATP-dependent chromatin modification, and is potentially a subunit of the PYR complex, which is associated with a γ - δ intergenic region in the β -like globin cluster (Wang et al, 2004; Nie et al, 2000). This region has been implicated in γ -globin silencing during the adult stage of development; its removal in deletional forms of HPFH may contribute to the reactivation of

fetal globin (O'Neill et al 1999). The PYR complex contains Ikaros as its DNA binding subunit, proteins from the SWI/SNF transcription activator chromatin remodeling family and a member of the NuRD transcription repressor family, Mi-2 (O'Neill et al, 2000). These features suggest that PYR may be an important complex involved in globin gene switching (Bank, 2005).

TSPYL1, a member of the NAP (Nucleosome Assembly Protein) family, is also associated with chromatin remodeling and is located in a chromosomal region (6q22-q23) frequently related to unlinked HPFH, although apparently adjacent to the polymorphic markers that delimit the 1.5 Mb region responsible for HPFH (Close et al, 2004). The increased expression of *ARID1B* and *TSPYL1* could help in the maintenance of an open chromatin structure in the region comprising the globin genes and may participate in their transcriptional regulation.

3.3- Expression of the transcriptional repressor *ZHX2*

We also investigated the expression of the zinc finger and homeobox gene *ZHX2*. This transcription factor was present at lower levels in reticulocytes from two HPFH subjects and a $\delta\beta$ -thalassemia subject analyzed. *ZHX2* has zinc finger and homeobox domains and was categorized as a transcriptional repressor (Kawata et al, 2003). Interestingly, it is located in a genomic region (8q24.13) in which a quantitative trait locus (QTL) associated with a heterocellular HPFH, unlinked to chromosome 11, was recognized (Garner et al, 2004). Moreover, a genetic alteration in the murine homologue *zhx2*, which is normally upregulated after birth, is responsible for the hereditary persistence of α -fetoprotein and *H19* genes (Perincheri et al, 2005). These proteins are normally expressed in the fetal liver, which is also the site of erythropoiesis during this stage of development. In cultured cells, the level of *ZHX2* mRNA was also diminished (Figure 2), which may implicate that this gene is involved in the repression of globin genes, particularly γ -globin.

3.4- Increased expression of a redoxin related gene

Mature red blood cells are more exposed to oxidative damage than other tissues due to a constant exposure to higher oxygen concentrations. Oxidative stress may damage RBC membrane and result in anemia (Forget, 2001). Therefore, under normal circumstances, RBCs are protected against peroxides and oxidative damage by antioxidant proteins. Herein, we investigated the expression of SH3BGRL2 gene, detected in relatively high frequency in our normal subtractive cDNA library. Although its function is not yet completely known, SH3BGRL2 may have redoxin activities. SH3BGRL2 is homologue to SH3BGR gene and belongs to the Thioredoxin-like protein superfamily, redox active proteins including thioredoxin and glutaredoxin [40]. qRT-PCR showed an incremental increase of approximately 27 fold (Figure 2) in cultured cells undergoing erythroid differentiation, which suggests that this protein may participate in mature red cell function and may be important in the normal physiology of erythrocytes to protect against the harmful effects of oxygen.

3.5- Conclusions

These are the first investigation of the mentioned newly described genes in erythroid differentiation and demonstrate that the expression of these genes is affected by EPO stimuli. Future studies regarding functional properties of these genes in erythrocytes may help elucidate important pathways of erythroid differentiation and maintenance of red cells in peripheral circulation, as well as aid in the understanding of genetic mechanisms involved in erythrocyte pathologies such as sickle cell disease and thalassemias.

Acknowledgments

This research was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP - Proc. N. 02/03396-8) and Hemocenter (Unicamp). We thank Mateus Andrade for helping with figures.

4- References

A. Bank, Regulation of human fetal hemoglobin: new players, new complexities, *Blood* **18** (2005).

J. Close, L. Game, B. Clark, J. Bergounioux, A. Gerovassili and S. L. Thein, Genome annotation of a 1.5 Mb region of human chromosome 6q23 encompassing a quantitative trait locus for fetal hemoglobin expression in adults, *BMC Genomics*. **5** (2004), pp. 33.

L. Diatchenko Y.F. Lau, A.P. Campbell, A. Chenchik, F. Moqadam, B. Huang, S. Lukyanov, K. Lukyanov, N. Gurskaya, E.D. Sverdlov and P.D. Siebert, Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries, *Proc Natl Acad Sci U S A* **93** (1996), pp. 6025-30.

E. Fibach, D. Manor, A. Oppenheim, E. A. Rachmilewitz, Proliferation and maturation of human erythroid progenitors in liquid culture. *Blood*. **73** (1989), pp.100-3.

B. G. Forget, Molecular basis of hereditary persistence of fetal hemoglobin. *Ann N Y Acad. Sci.* **850** (1998), pp. 38-44.

B. G. Forget, The β -thalassemias. In: M. H. Steinberg, B. G. Forget, D. R. Higgs, R. L. Nagel, Editors, *Disorders of Hemoglobin*, Cambridge University Press, Cambridge (2001), pp. 131-145.

C. Garner, N. Silver, S. Best, S. Menzel, C. Martin, T. D. Spector and S. L. Thein, Quantitative trait locus on chromosome 8q influences the switch from fetal to adult hemoglobin, *Blood* **104** (2004), pp. 2184-6.

H. Kawata, K. Yamada, Z. Shou, T. Mizutani, T. Yazawa, M. Yoshino, T. Sekiguchi, T. Kajitani and K. Miyamoto, Zinc-fingers and homeoboxes (ZHX) 2, a novel member of the ZHX family, functions as a transcriptional repressor, *Biochem J.* **373** (2003), pp. 747-57.

M. Mazzocco, M. Maffei, A. Egeo, A. Vergano, P. Arrigo, R. Di Lisi, F. Ghiotto and P. Scartezzini, The identification of a novel human homologue of the SH3 binding glutamic acid-rich (SH3BGR) gene establishes a new family of highly conserved small proteins related to Thioredoxin Superfamily, *Gene* **291** (2002), pp. 233-9.

Z. Nie, Y. Xue, D. Yang, S. Zhou, B. J. Derro, T. K. Archer and W. A Wang, A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex, *Mol Cell Biol.* **20** (2000), pp. 8879-88.

D. O'Neill, J. Yang, H. Erdjument-Bromage, K. Bornschlegel, P. Tempst and A. Bank, Tissue-specific and developmental stage-specific DNA binding by a mammalian SWI/SNF complex associated with human fetal-to-adult globin gene switching, *Proc. Natl. Acad. Sci. USA.* **96** (1999), pp. 349-354.

D. O'Neill, S.S. Shoetz, R.A. Lopez, M. Castle, L. Rabinowitz, E. Shor, D. Krawchuk, M.G. Goll, M. Renz, H. Seelig, S. Han, R.H. Seong, S. D. Park, T. N. Agaliot, Munshi, D. Thanos, H. Erdjument, P. Tempst, and A. Bank, An Ikaros-containing chromatin-remodeling complex in adult-type erythroid cells, *Mol. Cell. Biol.* **20** (2000), pp. 7572-7582.

S. Perincheri, R.W.C. Dingle, M.L. Peterson and B.T. Spear, Hereditary persistence of alpha-fetoprotein and H19 expression in liver of BALB/cJ mice is due to a retrovirus insertion in the *Zhx2* gene, *Proc Natl Acad Sci U S A* **102** (2005), pp. 396-401.

S. H. Pope, E. Fibach, J. Sun, K. Chin, G. P. Rodgers, Two-phase liquid culture system models normal human adult erythropoiesis at the molecular level. *Eur J Haematol.* **64** (2000), pp. 292-303.

G. Stamatoyannopoulos, Molecular and Cellular Basis of Hemoglobin Switching. In: M. H. Steinberg, B. G. Forget, D. R. Higgs, R. L. Nagel, Editors, *Disorders of Hemoglobin*, Cambridge University Press, Cambridge (2001), pp. 131-145.

J. Vandesompele, K. De Preter, F. Pattyn, B. Poppe, N. V. Roy, A. De Paepe and F. Speleman, Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, *Genome Biol.* **3** (2002), pp. 34.

X. Wang, N.G. Nagl, D. Wilskerd, M. Van Scoy, S. Pacchione, P. Yaciuk, P. B. Dallas and E. Moran, Two related ARID family proteins are alternative subunits of human SWI/SNF complexes, *Biochem J.* **383** (2004), pp. 319-25.

D.J. Weatherall and J.B. Clegg, *The Thalassemia Syndromes*, Oxford, Blackwell Science, Oxford (2001).

Figure 1- Detection of differentiated erythroid cells expressing HbF in primary culture. Flow cytometric histogram shows cells stained with anti-HbF before (A) and 11 days after (B) EPO addition.

Figure 2- Gene expression pattern of *TSPYL1*, *ARID1B*, *ZHX2*, *SH3BGRL2*, *HBB* and *HBG* in cultured cells from a normal subject. Cells were collected before (0h) and 11 days (216hs) after the addition of EPO. Bars depict fold change relative to reference sample (0h). Relative gene expression was calculated and normalized using two constitutive genes as internal controls (*ACTB* and *GAPDH*). Data presented correspond to the geometric mean between these controls.