

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL



**ESTUDO DO CONTROLO PÓS-TRANSCRICIONAL DA HEMOJUVELINA E SUA  
INFLUÊNCIA NA HOMEOSTASE DO FERRO**

**Filipa dos Santos Tomé**

**MESTRADO EM BIOLOGIA MOLECULAR E GENÉTICA**

2010

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL



**ESTUDO DO CONTROLO PÓS-TRANSCRICIONAL DA HEMOJUVELINA E SUA  
INFLUÊNCIA NA HOMEOSTASE DO FERRO**

Dissertação Orientada por:

Doutora Luísa Romão

(Departamento de Genética, Instituto Nacional de Saúde Dr. Ricardo Jorge)

Prof. Doutora Rita Zilhão

(Departamento de Biologia Vegetal, Faculdade de Ciências da Universidade de Lisboa)

**Filipa dos Santos Tomé**

MESTRADO EM BIOLOGIA MOLECULAR E GENÉTICA

2010

## **Agradecimentos**

Ao Doutor João Lavinha que na qualidade de Assessor do ex-Centro de Genética Humana e actual responsável da Unidade de Investigação e Desenvolvimento do Departamento de Genética, autorizou a realização deste trabalho, assim como ao Prof. Doutor Luís Nunes, então coordenador do Departamento de Genética e ao Prof. Doutor José Pereira Miguel, Presidente do Conselho Directivo do Instituto Nacional de Saúde Dr. Ricardo Jorge.

À Doutora Luísa Romão, pela oportunidade que me deu para realizar este trabalho, por toda a sua disponibilidade e orientação e por me ter acolhido tão bem no seu grupo.

À Prof. Doutora Rita Zilhão, pelo interesse demonstrado neste projecto e por toda a sua disponibilidade e ajuda.

Aos meus colegas de grupo Ana Luísa Silva, Alexandre Teixeira, Ana Morgado, Ana Ramos, Cristina Barbosa, Isabel Peixeiro e Rute Martins, e aos colegas também presentes Emília Fernandes, Bruno Silva e Andreia Coelho. Obrigada por toda a ajuda, por todos os óptimos momentos e por me terem recebido tão bem.

Aos meus amigos, sempre presentes. Ao Luís, por tudo.

À minha família, especialmente aos meus pais e irmão, a quem devo muito do sou hoje. Obrigada pelo amor incondicional. Ao Gaspar, pela boa companhia, doçura e simplicidade em todos os momentos.

Um agradecimento muito especial à Ana Luísa, por tão bem me ter acolhido e guiado no início da minha vida científica. Agradeço a confiança que me transmitiu, a infinita disponibilidade e a imprescindível contribuição dada a este projecto.

## Resumo

Uma classe importante das doenças humanas é representada por doenças hereditárias associadas à homeostase do ferro, entre as quais se destaca a Hemocromatose Hereditária (HH). Esta caracteriza-se por uma alteração dos mecanismos que regulam a absorção do ferro, levando ao aumento progressivo deste no organismo. Um dos tipos de HH é a Hemocromatose Juvenil (2A) causada por mutações no gene da hemojuvelina (*HJV*), envolvido na regulação do metabolismo do ferro através da modulação directa da transcrição do gene da hepcidina.

De maneira a perceber o papel do controlo pós-transcricional do gene *HJV* na homeostase do ferro foram analisados vários aspectos do metabolismo dos transcritos *HJV*, nomeadamente ao nível do *splicing*, do controlo de tradução por grelhas de leitura a montante da grelha de leitura principal (uORFs) e a sua degradação pelo mecanismo de decaimento do mRNA mediado por mutações *nonsense* (NMD).

Foi analisado o padrão de *splicing* em diferentes tecidos, transfectados e não transfectados, e os resultados indicam que o transcrito que codifica a proteína funcional não é produzido em quantidades elevadas, provavelmente fruto de um controlo rigoroso ao nível do *splicing* alternativo. Foi ainda detectada uma variante de *splicing* inicialmente referida no artigo de Papanikolaou e colaboradores (2004), mas cuja sequência não se encontra anotada nas bases de dados do transcriptoma.

Verificou-se também que a *HJV* tem duas uORFs funcionais que inibem consideravelmente a eficiência de tradução da ORF principal, tendo a uORF1 um papel mais preponderante nesta inibição. O reconhecimento do AUG principal é feito, pelo menos, através do mecanismo de *leaky scanning*.

Foi ainda demonstrado através de construções repórter que as uORFs da *HJV* são passíveis de induzir o NMD, sugerindo fortemente que na sua forma nativa o transcrito da *HJV* pode ser regulado por este mecanismo.

Palavras-chave: Hemojuvelina (*HJV*), Homeostase do ferro, *Splicing*, Grelhas de leitura a montante da grelha de leitura principal (uORFs), Decaimento do mRNA mediado por mutação *nonsense* (NMD)

## Abstract

An important class of human diseases is represented by inherited disorders related to iron homeostasis, among which is Hereditary Hemochromatosis (HH). It is characterized by an alteration of the mechanisms that regulate iron absorption, leading to its progressive increase in the body. Juvenile Hemochromatosis (2A) is a type of HH, caused by mutations in the hemojuvelin (HJV) gene, involved in the regulation of iron metabolism through direct modeling of the transcription of the hepcidin gene.

In order to understand the role of the post-transcriptional control of the *HJV* gene in iron homeostasis, several aspects of the transcript metabolism were analyzed, namely its splicing pattern, its translational control through upstream open reading frames (uORFs) and its degradation through the nonsense mediated mRNA decay (NMD) mechanism.

The splicing pattern was analyzed in different transfected and non transfected tissues, and the results reveal that the transcript that codifies for the functional protein is not easily detected, suggesting it is not present at high levels, probably due to a rigorous control at the alternative splicing level. Moreover, a splicing variant was detected, initially referred in Papanikolaou et al (2004), whose sequence, however, is not annotated in transcriptome databases.

Additionally, we have observed that HJV mRNA has two functional uORFs which considerably inhibit the translational efficiency of the main ORF, having the first uORF a more prominent role in the translational control. The recognition of the main ORF is achieved, at least, by leaky scanning.

It was also demonstrated through reporter constructions that HJV's uORFs are sensitive to NMD, strongly suggesting that the native HJV mRNA could be naturally regulated by this mechanism.

Keywords: Hemojuvelin (HJV), Iron Homeostasis, Splicing, upstream open reading frames (uORFs), nonsense mediated RNA decay (NMD)

## Abreviaturas

A	Adenina
BMP	<i>Bone Morphogenetic Proteins</i>
pb	pares de bases
C	Citosina
DH5 $\alpha$	Estirpe de bactérias <i>E. coli</i>
DMEM	<i>Dulbecco's Modified Eagle Medium</i>
DNA	Ácido desoxirribonucleico ( <i>Deoxyribonucleic acid</i> )
EJC	Complexo de junção exão-exão ( <i>Exon Junction Complex</i> )
FBS	Soro fetal bovino ( <i>Fetal Bovine Serum</i> )
G	Guanosina
GPI	glicosilfosfatidilinositol
h	Hora(s)
HAMP	Hepcidina ( <i>Hepcidin anti-Microbial Peptide</i> )
HeLa	Linha celular derivada de carcinoma cervical
HepG2	Linha celular derivada de carcinoma hepatocelular
HFE	<i>High Fe</i>
HH	Hemocromatose Hereditária ( <i>Hereditary Hemochromatosis</i> )
HJV	Hemojuvelina
HPR	Peroxidase ( <i>Horseradish Peroxidase</i> )
JH	Hemocromatose Juvenil ( <i>Juvenile Hemochromatosis</i> )
min	Minuto(s)
mRNA	RNA mensageiro
n	Qualquer nucleótido
N-terminal	Região terminal amínica
NMD	Decaimento do mRNA mediada por mutação <i>nonsense</i> ( <i>Nonsense Mediated mRNA Decay</i> )
ORF	Grelha de leitura ( <i>Open Reading Frame</i> )
PBS	Solução salina de tampão fosfato ( <i>Phosphate Buffered Saline</i> )
PCR	Reacção de polimerização em cadeia ( <i>Polymerase chain reaction</i> )
PLB	Tampão de lise passiva ( <i>Passive Lysis Buffer</i> )
PVDF	Fluoreto de polivinilideno ( <i>Polyvinylidene Fluoride</i> )
PTC	Codão de terminação prematuro ( <i>Premature Termination Codon</i> )

RGD	<i>Repulsive Guidance Domain</i>
RGM	<i>Repulsive Guidance Molecule</i>
RNA	Ácido ribonucleico ( <i>Ribonucleic acid</i> )
RPMI	Meio rico ( <i>Roswell Park Memorial Institute</i> )
seg	Segundo(s)
SDS	Dodecil sulfato de sódio ( <i>Sodium Dodecyl Sulfate</i> )
siRNA	<i>small interfering RNA</i>
snRNP	pequenas ribonucleoproteínas nucleares ( <i>small nuclear ribonucleoproteins</i> )
SS	<i>Splice Site</i>
T	Timina
TfR1	Receptor da Transferrina - 1
TfR2	Receptor da Transferrina - 2
U	Uracilo
uORF	Grelha de leitura a montante da grelha de leitura principal ( <i>upstream Open Reading Frame</i> )
UPF	<i>Up-Frameshift Protein</i>
UTR	Região não traduzida ( <i>Untranslated Region</i> )
vWf	<i>von Willebrand factor D</i>

# Índice

Índice .....	VIII
1 Introdução .....	1
1.1 Aspectos gerais sobre o ferro .....	1
1.2 Hemocromatose hereditária e hemocromatose juvenil.....	2
1.3 Expressão e envolvimento do gene <i>HJV</i> na homeostase do ferro.....	3
1.4 O processo de <i>splicing</i> e o transcrito <i>HJV</i> .....	4
1.5 Regulação da tradução por uORFs e o transcrito <i>HJV</i> .....	6
1.6 O mecanismo de NMD e o transcrito <i>HJV</i> .....	8
2 Metodologia Experimental .....	11
2.1 Clonagem dos vectores de expressão .....	11
2.1.1 Variante A.....	11
2.1.2 Construções para o estudo das uORFs .....	11
2.1.3 Construções para o estudo do NMD .....	12
2.2 Culturas celulares e transfecções .....	12
2.3 Extracção de RNA, síntese de cDNA e PCR.....	13
2.4 Transcrição <i>in vitro</i> .....	14
2.5 Ensaio de luminescência .....	14
2.6 RT-PCR quantitativo .....	14
2.7 Imuno-deteção por <i>Western-Blot</i> .....	15
3 Resultados.....	16
3.1 Análise do padrão de <i>splicing</i> do transcrito <i>HJV</i> .....	16
3.2 Controlo da tradução da <i>HJV</i> por uORFs.....	21
3.3 Controlo da degradação do transcrito <i>HJV</i> por NMD.....	25
4 Discussão final e perspectivas futuras.....	28
5 Bibliografia .....	30

# 1 Introdução

A capacidade de os organismos constantemente detectarem e se adaptarem a alterações ambientais é importante para que mantenham as suas funções celulares, tendo a modulação da expressão génica um papel central nesta flexibilidade e na manutenção da homeostase [1]. A regulação transcricional e a regulação pós-transcricional, incluindo ainda a regulação ao nível da estabilidade do RNA mensageiro (mRNA), da tradução, das modificações pós-traducionais e da degradação proteica têm um papel crucial nesta modulação [1] [2]. O controlo transcricional tem recebido muita atenção relativamente ao controlo pós-transcricional, começando, no entanto, a surgir uma maior consideração por este último [2] [3]. Os mecanismos de regulação pós-transcricional englobam vários processos, nomeadamente o processamento do mRNA (poli-adenilação, *capping e splicing*), exportação e localização do mRNA, e tradução ou degradação do mesmo. Todos estes têm em comum o facto de o seu objectivo ser controlar se, onde e quão eficientemente um dado mRNA é traduzido em proteína, o que reforça a importância do estudo destes mecanismos no contexto do controlo da expressão génica, e confere ao mRNA um papel central [3] [4].

O presente trabalho pretende estudar o controlo pós-transcricional do gene Hemojuvelina humana (*HJV*), envolvido na homeostase do ferro. Quando mutado, o gene *HJV* é responsável por uma condição denominada Hemocromatose Juvenil (JH). Este trabalho tem como objectivo avaliar o padrão de *splicing* do mRNA da *HJV*, o controlo da sua tradução por grelhas abertas de leitura localizadas a montante da grelha de leitura principal (*upstream open reading frames – uORFs*) e a sua degradação pelo mecanismo de decaimento do mRNA mediado por mutações *nonsense* (*nonsense mediated RNA decay – NMD*).

## 1.1 Aspectos gerais sobre o ferro

O ferro é um elemento crucial para a maioria das reacções biológicas realizadas pelos sistemas vivos, sendo um componente essencial dos transportadores de oxigénio tais como a hemoglobina e a mioglobina, dos citocromos e de vários enzimas envolvidos na oxidação ou redução de substratos [5]. No entanto, a sua capacidade de aceitar e doar electrões obriga a que esteja sempre associado a ligandos específicos, caso contrário os iões livres em solução formam radicais extremamente reactivos e potencialmente tóxicos, que levam a danos celulares [6] [7]. A homeostase do ferro é mantida essencialmente através do controlo rigoroso da sua absorção e libertação das células, uma vez que não há nenhum mecanismo

conhecido para o excretar do organismo [8]. Em mamíferos, o ferro é transportado no plasma ligado essencialmente à transferrina. O ferro ligado à transferrina é captado pelas células através do receptor da transferrina 1 (TfR1), e armazenado no citoplasma associado à ferritina. O movimento do ferro entre estes compartimentos é um processo finamente regulado, que é modulado de acordo com os requisitos do organismo [7]. O fígado tem um papel central no metabolismo do ferro, sendo os hepatócitos o maior local do seu armazenamento. É também no fígado que é expresso um complexo grupo de genes associados a doenças hereditárias relacionadas com o ferro [5]. A absorção a níveis apropriados de ferro pelas células é controlada através da alteração da expressão de TfR1 na superfície celular, pelo que quanto mais ferro uma célula necessite, maior a expressão de TfR1. No entanto, a libertação deste elemento é regulada por sinais sistémicos, e surgem agora evidências de que o efluxo celular de ferro é o maior ponto de regulação para a manutenção da sua homeostase [7]. A hepcidina tem um papel central nesta regulação. É um pequeno péptido solúvel sintetizado predominantemente nos hepatócitos [9], cuja expressão do mRNA é estimulada pelo excesso de ferro e inflamação, e diminui em situações de hipóxia e anemia [10]. Actua bloqueando a função do exportador de ferro ferroportina, induzindo a sua internalização e degradação. Consequentemente, a absorção do ferro pelo intestino diminui e a libertação de ferro das reservas nos hepatócitos e macrófagos também diminui [11].

## **1.2 Hemocromatose hereditária e hemocromatose juvenil**

A importância do controlo rigoroso da absorção e libertação celular de ferro é evidente através de várias condições patológicas em humanos [7]. A Hemocromatose Hereditária (HH) é uma condição genética associada à alteração dos mecanismos que regulam a absorção do ferro, levando ao aumento progressivo deste no organismo [12] [13] [14]. O ferro em excesso é depositado em vários órgãos, comprometendo o seu funcionamento e podendo levar à sua falha, resultando em doenças graves como cirrose, diabetes, cardiomiopatias e artrite [12] [13] [15]. Há vários tipos de HH, definidos de acordo com o gene mutado: tipo 1, tipos 2A e 2B, tipo 3 e tipo 4, causados por mutações nos genes HFE (*High Fe*), HJV, HAMP (*Hepcidin*), TFR2 (receptor da transferrina – 2) e ferroportina, respectivamente [13] [14]. Com excepção do tipo 4, todas as condições apresentam características clínicas e histológicas semelhantes. Em todos eles, a absorção do ferro a nível intestinal é inapropriadamente elevada e ocorre uma rápida libertação de ferro pelos macrófagos (que armazenam o ferro sob a forma de ferritina), o que conduz a um aumento de transferrina saturada em circulação. Isto favorece a

absorção do ferro celular através do receptor da transferrina e causa deposições de ferro, especialmente nos hepatócitos [12].

A HH tipo 2, também chamada Hemocromatose Juvenil (JH), é caracterizada por ocorrer em indivíduos jovens, tipicamente com menos de trinta anos, afectar ambos os sexos igualmente e, em comparação com a HH tipo 1, ser mais severa, apresentando um aumento diário de absorção de ferro mais elevado e uma taxa de acumulação de ferro mais acelerada [12] [13]. Há duas formas de JH, uma devida a mutações no gene *HJV* (tipo 2A), que, como atrás referido, codifica uma proteína denominada hemojuvelina e outra no gene *HAMP* (tipo 2B), que codifica uma proteína denominada hepcidina. Ambas apresentam fenótipos clínicos e bioquímicos indistinguíveis e originam níveis baixos de hepcidina [12] [16].

### 1.3 Expressão e envolvimento do gene *HJV* na homeostase do ferro

O gene *HJV* tem cerca de 4.2kb e foi recentemente identificado por Papanikolaou e colaboradores (2004) [16]. É transcrito e processado num mRNA maduro contendo a totalidade dos exões (variante A) e em quatro variantes de *splicing* adicionais, que codificam três isoformas proteicas de 200, 314 e 426 aminoácidos. A *HJV* é essencialmente expressa no fígado adulto e fetal, no coração e no músculo [16] [17]. A proteína codificada pelo transcrito principal, a variante A, contém múltiplos domínios, incluindo um péptido sinal na zona N-terminal, um domínio de três aminoácidos *repulsive guidance domain* (RGD), um factor de *von Willebrand D* (vWf) e um domínio transmembranar na zona C-terminal característico de uma âncora de glicosilfosfatidilinositol (GPI-anchor) [16]. A remoção desta âncora poderá gerar uma forma solúvel da proteína [18]. A *HJV* revela uma considerável homologia de sequência de aminoácidos com os factores proteicos *Repulsive Guidance Molecules* (RGM) [16] que actuam como co-receptores de uma classe de ligandos chamados *Bone Morphogenetic Proteins* (BMPs). Estes são factores de crescimento com papéis importantes em várias actividades biológicas através da ligação a determinados receptores, desencadeando vias de sinalização que regulam a transcrição de genes específicos [19]. Foi demonstrado por Babbitt e colaboradores que a *HJV* é de facto um co-receptor para esses ligandos, actuando através da via de sinalização BMP para modular a transcrição do gene *HAMP*, tendo assim um papel importante no metabolismo do ferro através da regulação da expressão da hepcidina. [8] [20]. Recentemente, esta relação foi reforçada através de uma experiência em que usaram pequenos RNAs de interferência (siRNAs) para diminuir a concentração do mRNA da *HJV* e se observou um decréscimo nos níveis de mRNA da hepcidina, correlacionado com o

decréscimo da concentração de mRNA da *HJV* [18]. Pensa-se ainda que a forma solúvel da *HJV* (sem a âncora GPI) possa ter algum papel na regulação da expressão da hepcidina, mas este ainda não está totalmente esclarecido [18] [20]. Apesar de haver evidências que a sinalização mediada por BMP possa ser um mecanismo importante para a *HJV* regular a expressão da hepcidina e a homeostase do ferro [20], os detalhes e mecanismos moleculares deste sistema regulador ainda não estão inteiramente caracterizados [21]. Contudo, observou-se que enquanto os níveis de mRNA da hepcidina variam com diferentes estímulos de ferro e inflamação, os níveis de mRNA da *HJV* não são afectados [22]. Isto indica que a regulação da expressão da *HJV* não ocorre apenas a nível transcricional, e que a sua regulação pós-transcricional e traducional poderão ter um papel importante na homeostase do ferro.

#### 1.4 O processo de *splicing* e o transcrito *HJV*

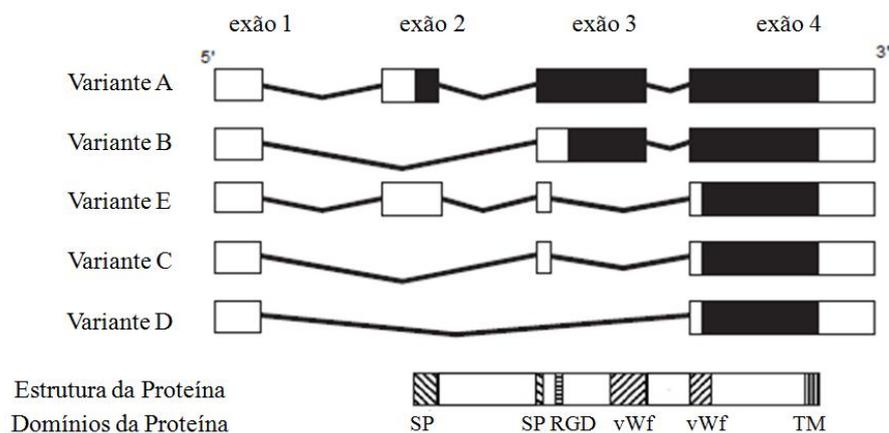
O *splicing* do pré-mRNA é um evento pós-transcricional que ocorre ao nível do processamento do mRNA, sendo potencialmente um ponto de controlo da expressão génica e crítico na expressão de praticamente todos os genes eucariotas [23]. Consiste na remoção dos intrões e junção dos exões adjacentes do pré-mRNA pelo spliceossoma, um complexo constituído por cinco pequenas ribonucleoproteínas nucleares (snRNPs) e cerca de 200 proteínas, e envolve duas reacções consecutivas de transesterificação [24] [25] [26]. Estas reacções dependem do reconhecimento do 5' *Splice Site* (SS) e do 3'SS. O primeiro marca a junção exão/intrão na extremidade 5' do intrão, que inclui um dinucleótido GU no intrão, seguida de uma sequência consensos não muito conservada. Do outro lado do intrão, a região 3'SS tem três elementos consensos conservados, que são o *branch point*, seguido de uma extensão de polipirimidina e uma sequência terminal AG. É o spliceossoma que reconhece estas sequências e catalisa os dois passos de transesterificação da reacção de *splicing*. [25]. O primeiro consiste na activação de um resíduo do *branchpoint*, normalmente uma adenosina, possibilitando o ataque nucleofílico à ligação fosfodiéster no 5'SS, e resultando na formação de uma estrutura intermediária denominada *lariat*. O segundo passo envolve um ataque nucleofílico por parte do 3'OH do exão a 5' à ligação fosfodiéster do 3'SS, resultando na ligação dos dois exões e excisão do intrão *lariat* [26].

Pode gerar-se mais do que um transcrito de mRNA a partir do mesmo mRNA precursor através de um processo denominado *splicing* alternativo [24], em que exões do transcrito primário podem sofrer *splicing* em diferentes arranjos, produzindo mRNAs diferentes. [27] Estes rearranjos permitem a produção de várias isoformas proteicas, contribuindo assim para

a diversidade do proteoma [24]. Contudo, a regulação do *splicing* alternativo nas células e tecidos não está ainda totalmente esclarecida [28].

Uma vez que as sequências consenso dos locais de *splicing* não são geralmente suficientes para determinar se um certo local vai ser reconhecido pelo spliceossoma, são necessários outros tipos de sinais de *splicing* para a sua activação [24] [25], tanto para o *splicing* constitutivo como para a regulação do *splicing* alternativo [24]. Estas incluem sequências regulatórias intrónicas ou exónicas (cis) às quais se ligam factores externos, influenciando assim a escolha do local de *splicing*. Estes elementos que actuam em cis são classificados como *enhancers* ou *silencers* exónicos ou intrónicos e podem estimular positivamente o recrutamento do spliceossoma (*enhancers* ou promotores), ou bloquear a ligação do spliceossoma a certos locais de *splicing* (*silencers* ou repressores) [24] [28].

O gene *HJV* é transcrito e processado num mRNA maduro contendo a totalidade dos exões e em 4 variantes de *splicing* adicionais (Figura 1) [16].



**Figura 1 - Representação dos transcritos da *HJV* e a estrutura da proteína** (adaptado de Papanikolaou *et al* [16]) A preto está representada a região codificante e a branco a região não codificante. Em baixo está representada a estrutura da proteína, com a respectiva correspondência dos motivos proteicos. **Legenda:** SP: péptido sinal; RGD: *repulsive guidance domain*; vWf: factor de *von Willebrand* D; TM: domínio transmembranar.

A variante A origina a proteína *HJV* funcional com 426 aminoácidos e todos os domínios acima descritos. De acordo com as previsões teóricas, a variante B origina uma proteína com 316 aminoácidos, truncada na região N-terminal (sem o péptido sinal e o domínio RGD) e as variantes C e D originam a mesma isoforma proteica de 200 aminoácidos que possui apenas o domínio transmembranar. A variante E potencialmente poderá ter como local de início de tradução o mesmo AUG da variante A, mas isso iria originar uma alteração da grelha de leitura da tradução (*frameshift*) alguns codões a jusante do codão de iniciação com conseqüente definição de um codão de terminação prematuro, formando-se uma proteína

truncada. Assim, e de acordo com a figura 1, o local de início da tradução da variante E seria o mesmo das variantes C e D, originando a mesma isoforma proteica. É, no entanto, curioso notar que a variante E não consta das bases de dados do transcriptoma, apesar de ter sido inicialmente descrita no artigo de Papanikolaou e colaboradores (2004) [16].

## 1.5 Regulação da tradução por uORFs e o transcrito *HJV*

Um dos passos mais importantes na regulação da expressão génica ocorre ao nível da tradução do mRNA. A própria tradução é regulada a vários níveis, tanto no decorrer da iniciação, como da alongação e terminação, sendo a iniciação o ponto mais crítico e limitante de todo este processo [29]. Para a maioria dos mRNAs eucariotas, a tradução ocorre pelo mecanismo de *scanning*, que requer a interacção da subunidade 40S do ribossoma e vários factores de iniciação com a região 5' do mRNA. Este complexo avança linearmente desde a estrutura *cap* ao longo da região 5' não traduzida (5'UTR) na direcção 5'-3' (*scanning*) até encontrar o primeiro codão AUG (regra do 1º AUG). O reconhecimento desse codão leva ao recrutamento da subunidade 60S para a formação do ribossoma 80S, de forma a iniciar-se a alongação [2] [30] [31].

As UTRs dos mRNAs têm um papel crucial na determinação da eficiência da tradução [32]. Em particular, há determinadas características da 5'UTR que influenciam a tradução, como o seu tamanho, a sua estrutura secundária e a presença de AUGs (e a sequência nucleotídica que os rodeiam) a montante do codão de iniciação da grelha de leitura (ORF) principal [33], denominados uAUGs. Um uAUG é definido como qualquer codão de iniciação localizado na região 5' líder de um mRNA, reconhecido pelo ribossoma, a montante do codão de iniciação usado pelo ribossoma para sintetizar o principal produto proteico. Ao uAUG pode estar associada uma grelha de leitura, denominada uORF [34]. Apesar de os codões de iniciação das uORFs se localizarem a montante da ORF principal, o seu codão de terminação pode localizar-se em diferentes posições relativamente a esta, isto é, a montante, ficando a uORF localizada na região 5' líder, dentro da própria região codificante, sobrepondo-se assim com a ORF principal, como pode ainda ser o mesmo da ORF principal, produzindo-se assim uma proteína maior em relação à produzida pelo AUG da ORF principal [35].

Dado que o mecanismo de *scanning* convencional prevê que a tradução se inicie no AUG mais perto da extremidade 5' do mRNA, as uORFs funcionam como uma barreira prevista para impedir o acesso do ribossoma ao AUG principal [34], e a tradução deste apenas é possível através de dois mecanismos que permitem o escape à regra do 1º AUG: o *leaky*

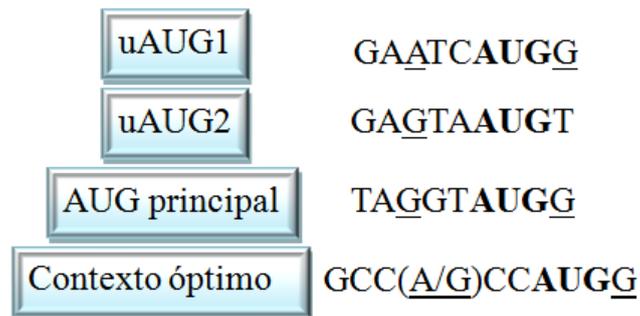
*scanning* e a reiniciação [2] [30]. Quando o reconhecimento da uORF é regulado por *leaky scanning*, os ribossomas tanto ignoram o uAUG e continuam o *scanning* até ao AUG subsequente, como podem reconhecê-lo e iniciar a tradução [2] [29]. A eficiência de reconhecimento do AUG pelo ribossoma está relacionada com parâmetros como os nucleótidos que ladeiam o AUG (que definem o contexto em que este se encontra), a proximidade do AUG ao local *cap* e a estrutura secundária no mRNA [29]. Assim, a ocorrência de *leaky scanning* é dependente do contexto nucleotídico do AUG iniciador. De facto, codões AUG com um contexto sub-ótimo não são, por vezes, reconhecidos pelas subunidades 40S, prosseguindo estas o *scanning* até ao AUG subsequente. O contexto considerado ótimo - GCC(A/G)CCAAUGG - foi analisado por Kozak [36], sendo as posições -3 e +4 (a negrito) as mais importantes. Um contexto considerado adequado é AnnAUGn ou GnnAUGGG e um contexto mais fraco é (C/T)nnAUG(A/C/T) em que n é qualquer nucleótido [35] [36]. A relação entre o contexto do AUG e o *leaky scanning* é reforçada pelo facto de ao melhorar-se o contexto de um uAUG, a iniciação no segundo AUG ser reduzida ou mesmo abolida. De notar que não se trata de o contexto do uAUG ser melhor do que o do AUG principal, uma vez que o ribossoma avança linearmente, e a sua decisão de parar ou não no uAUG não é influenciada pelo facto de existir um contexto preferível a jusante [30].

A reiniciação é um mecanismo mais raro. Neste caso a subunidade 40S continua ligada ao mRNA após a terminação da tradução de pequenas uORFs, continuando o *scanning*, e iniciando a tradução no AUG seguinte. O tempo que demora a tradução da primeira ORF parece ser a maior limitação para a ocorrência de reiniciação em eucariotas [30]. Uma possível explicação poderá ser o facto de certos factores de iniciação se dissociarem do ribossoma durante a elongação, e portanto se este processo for breve, os factores necessários para a reiniciação poderão ainda estar presentes quando o ribossoma atingir o codão de terminação da uORF. Isto poderá promover o recomeço do *scanning* por parte da subunidade 40S, não chegando esta a dissociar-se completamente do transcrito [29] [30] [37]. Por outro lado, estruturas secundárias fortes próximas do codão de terminação da uORF poderão impedir a ocorrência de reiniciação ao impedirem a deslocação da subunidade 40S em direcção ao AUG subsequente [30].

A presença de uAUGs na 5'UTR geralmente reduz a eficiência de iniciação no AUG da ORF principal [2], que em certos genes eucariotas deve ser baixa para impedir a produção em excesso de determinadas proteínas [38]. O facto de as uORFs serem frequentes em genes com funções biológicas críticas sugere que estas providenciam estratégias alternativas para a

coordenação da expressão génica, servindo assim como elementos mediadores importantes do controlo da tradução [29] [38].

O gene *HJV* tem na sua 5'UTR dois uAUGs e um codão STOP a montante do AUG principal, associados a duas uORFs. A primeira uORF tem 28 codões, e a segunda tem 19 codões. O contexto em que dos uAUGs e do AUG principal estão inseridos, assim como o contexto óptimo definido por Kozak [36], está evidenciado na figura 2.



**Figura 2 - Representação do contexto dos uAUGs e do AUG principal, assim como do contexto óptimo.** A negrito está evidenciado o codão de iniciação e os codões mais significativos para o contexto óptimo estão sublinhados.

O uAUG1 e o AUG principal têm um contexto favorável, na medida em que em relação à sequência consensos considerada óptima os nucleótidos mais importantes (sublinhados na figura 2) estão conservados. O uAUG2 não apresenta um contexto tão favorável, uma vez que apenas um dos nucleótidos mais importantes está conservado. Nenhum estudo foi feito até agora de maneira a elucidar o papel destas uORFs na regulação da expressão da *HJV*.

## 1.6 O mecanismo de NMD e o transcrito *HJV*

Está bem estabelecido que o controlo da expressão génica ocorre não só ao nível transcricional, mas também em vários eventos pós-transcricionais [39] em que a qualidade e fidelidade do metabolismo do mRNA têm de ser asseguradas [4]. Tal controlo é conseguido através de mecanismos de vigilância que actuam em vários passos da biogénese do mRNA, detectando e degradando transcritos defeituosos e inadequadamente processados [40]. O mecanismo de vigilância mais bem caracterizado designa-se mecanismo de decaimento do mRNA mediado por mutações *nonsense* (*nonsense-mediated mRNA decay* - NMD) que detecta mRNAs com codões de terminação prematuros (PTC) e os conduz à degradação [4]. Um dos objectivos fisiológicos deste mecanismo é limitar a síntese de proteínas truncadas na região C-terminal, protegendo a célula dos seus efeitos deletérios dominantes negativos ou efeitos de ganho de função [4].

Mutações *nonsense* são alterações que introduzem prematuramente na região codificante do mRNA codões de terminação da tradução; podem ser originadas tanto por mutações ao nível do DNA, deleções, inserções ou substituições, entre outros [4]. No entanto, podem também surgir como consequência de processos não associados a erros, envolvidos na regulação no metabolismo do mRNA, tais como o *splicing* alternativo ou utilização de codões de iniciação alternativos. Por exemplo, transcritos que contêm uORFs podem ser reconhecidos como transcritos portadores de PTCs [4]. Mais ainda, uma estimativa recente indica que de cerca dos 70% dos pré-mRNAs que sofrem *splicing* alternativo, 45% geram pelo menos uma forma de *splicing* que se prevê ser um substrato para NMD [41]. Este parece assim ter um papel importante não só na degradação de transcritos anómalos, mas também na regulação do nível basal de mRNAs envolvidos em diferentes processos celulares [41].

Uma questão central é perceber como o NMD distingue um PTC de um codão de terminação natural [41]. Sabe-se que, de uma maneira geral, o NMD depende, em mamíferos, do *splicing* [4] [42]. Durante o *splicing* do pré-mRNA, é depositado a cerca de 20 - 24 nucleótidos a montante de cada junção exão-exão um complexo proteico designado complexo de junção exão-exão (*Exon Junction Complex* – EJC) [43], que actua no NMD como marca para o reconhecimento de um PTC [44]. A terminação da tradução parece ser também um evento crítico no reconhecimento de um codão de terminação como prematuro [45]. A ponte mecanicista entre estes dois processos é feita através de três factores denominados proteínas *up-frameshift* (UPF), a UPF1, UPF2 e UPF3, que constituem o centro da maquinaria do NMD [4] [41] [42] [46]. O modelo actual para o mecanismo de NMD postula que durante o 1º ciclo de tradução o ribossoma remove (ou altera) os EJCs à medida que percorre o transcrito. Na presença de um PTC localizado a mais de 50 - 54 nucleótidos a montante da última junção exão-exão, o ribossoma termina a tradução antes de o último EJC ter sido alterado ou removido. Isto permite a interacção entre os factores UPF2 e UPF3 (associados ao EJC) e o factor UPF1 (associado ao complexo de terminação da tradução), levando à activação do NMD e à degradação do transcrito. Na ausência de PTCs ou no caso de estes se localizarem depois ou demasiado perto do último EJC, o ribossoma atinge o codão de terminação sem qualquer EJC funcional associado ao transcrito, não sendo possível a activação do NMD, e o transcrito é sujeito a múltiplos ciclos de tradução [4] [47].

No caso de transcritos com uORFs, em que codão de terminação é reconhecido pelo ribossoma, este pode também ser reconhecido pela maquinaria de NMD como prematuro, mas tal não é imperativo, ou seja, nem todos os mRNAs que contêm uORFs funcionais são alvos de NMD [48] [49] [50]. Um exemplo é o do gene que codifica a trombopoetina (TPO), que

contém sete uORFs, mas cuja regulação é independente do NMD. No entanto, o aumento do tamanho de uma das uORFs levou a um decréscimo da expressão do mRNA da TPO, revertido parcialmente pelo *knockdown* da UPF1, o que sugere um papel do tamanho da uORF na activação deste mecanismo [49]. Foi também descrito que mRNAs da  $\beta$ -globina com mutações *nonsense* e que obedecem à regra dos 50-54 nucleótidos são resistentes ao NMD, e que a inserção de um fragmento entre o codão de iniciação e a mutação *nonsense*, aumentando a ORF de 15 para 39 codões, torna o mRNA sensível ao NMD [51]. Foi demonstrado que o parâmetro crítico para esta inibição é de facto a proximidade do codão *nonsense* ao codão de iniciação da tradução [4] [52]. Em plantas, já está descrito que a activação do NMD está dependente do tamanho da uORF, propondo que uORFs com cerca de 50 aminoácidos despoletam o NMD eficientemente, mas uORFs mais pequenas (31 a 15 aminoácidos) falham na mesma resposta [53]. No entanto, esta dependência não está totalmente esclarecida em mamíferos.

Tal como referido no ponto anterior, o transcrito da HJV humana possui duas uORFs na região 5' líder, às quais está associado um codão de terminação. Caso as uORFs sejam funcionais, este codão pode ser reconhecido pela maquinaria de NMD como prematuro, tornando o transcrito *HJV* num alvo de regulação por este mecanismo. O codão de terminação das uORFs está localizado a mais de 54 nucleótidos a montante da última junção exão-exão. Nenhum estudo foi feito até agora de maneira a esclarecer se o NMD poderia ter algum papel na regulação da expressão génica da HJV.

O objectivo desta dissertação é estudar o controlo pós-transcricional do gene *HJV*. Para tal, avaliar-se-á o padrão de *splicing*, usando tecidos transfectados e não transfectados com o gene *HJV*, de forma a determinar por RT-PCR quais as variantes maioritárias. Pretende também avaliar-se o controlo da tradução por uORFs, usando para tal o gene repórter da luciferase, do qual se medirá a actividade, o que permitirá quantificar a proteína produzida. Será ainda efectuada a detecção das proteínas produzidas por *Western-Blot*. Pretende ainda investigar-se se o transcrito *HJV* é sensível ao mecanismo de NMD, através da quantificação dos seus níveis de RNA por PCR em tempo real. A identificação de mecanismos que controlem a expressão deste gene poderá ajudar a melhor compreender qual o seu papel na homeostase do ferro.

## 2 Metodologia Experimental

### 2.1 Clonagem dos vectores de expressão

#### 2.1.1 Variante A

Para clonar a região codificante da *HJV* correspondente à variante A no vector de expressão pcDNA3/hygro+ (*Invitrogen*), amplificou-se por PCR a partir de cDNA da *HJV* o fragmento *Nhe I / Xho I*, que abrange o 1º e 3º exões, usando os oligonucleótidos iniciadores 5'UTR *Nhe* e *Xho I* mRNA AS II (Quadro I, Anexo), e o fragmento *BstXI / Bam HI*, que abrange o 3º e 4º exões usando os oligonucleótidos iniciadores ex3S II e ex4 AS (Quadro I, Anexo). Os fragmentos foram clonados nos locais de restrição correspondentes no vector de expressão pcDNA3/hygro+ (*Invitrogen*) que continha o gene *HJV*, previamente clonado, usando os enzimas de restrição *Nhe I* (*Amersham*), *Xho I* (*Fermentas*), *BstXI* (*Roche*) e *Bam HI* (*Amersham*) e a ligase de DNA do fago T4 (*Fermentas*). Usou-se este produto para transformar bactérias quimicamente competentes DH5α (*Invitrogen*), das quais se extraiu o DNA plasmídico, utilizando o *kit* de extracção *JETquick Plasmid MiniPrep SpinII Kit/250 Genomed* para isolamento de clones recombinantes. A construção molecular obtida denominou-se VarianteA clonada, e a sua correcta construção foi confirmada por sequenciação automática.

#### 2.1.2 Construções para o estudo das uORFs

Para estudar o efeito das uORFs na expressão da *HJV* usou-se um gene repórter, o gene da luciferase, de maneira a quantificar a proteína produzida através de ensaios de luminescência. Foram utilizados dois vectores de expressão: o pGL2-*enhancer* (*Promega*), que codifica para a luciferase do pirilampo (*Firefly*), em que previamente foi clonado o promotor do citomegalovírus humano (hCMV) presente no vector pcDNA3/hygro+ (*Invitrogen*), e o pGL4.70 (*Promega*), que codifica para a luciferase da *Renilla*. A região 5' líder do gene *HJV* foi clonada a montante do cDNA da luciferase do pirilampo, criando-se uma fusão em que o AUG da ORF da luciferase se situa exactamente no mesmo local do AUG da ORF da *HJV*. Para tal amplificou-se por PCR a região 5' líder da *HJV* a partir de cDNA sintetizado a partir de fígado humano, com os oligonucleótidos iniciadores link *Hind III* (com um *linker* para o local de restrição do enzima *Hind III*) e *HJV luc*, e amplificou-se a partir do vector pGL2-*enhancer* a região entre os locais de restrição *Hind III* e *Xba I*, com os oligonucleótidos iniciadores luc *HJV* e pGL2 AS (Quadro I, Anexo). Os produtos dos dois

PCRs foram ligados através de uma terceira amplificação usando como DNA molde os produtos de amplificação anteriores e o oligonucleótido iniciador directo do 1º PCR e o inverso do 2º PCR, respectivamente link *Hind* III e pGL2 AS. Este produto foi clonado no pGL2-*enhancer* usando os enzimas *Hind* III (*Promega*) e *Xba* I (*Promega*), e a ligase de DNA do fago T4. Esta construção foi denominada pGL2+5'UTR HJV e a sua sequência confirmada por sequenciação automática. Foram realizadas alterações neste construção através de mutações pontuais nos seus AUGs e codão STOP, ou no AUG principal. Estas mutações foram realizadas por mutagénesse dirigida através do kit *QuikChange Site-Directed Mutagenesis* (*Stratagene*), de acordo com as instruções do fabricante, e usando oligonucleótidos iniciadores específicos para cada mutação desejada (Quadro I, Anexo). O programa utilizado consistiu numa desnaturação inicial durante 30s a 95°C, seguida de 12 ciclos de: desnaturação 30seg a 95°C, hibridação durante 1min a 55°C e extensão durante 16min a 68°C. Fez-se também um passo de extensão adicional durante 10min a 68°C. Digeriu-se o produto de PCR com *Dpn* I (*Fermentas*) durante 1h a 37°C para remover as cadeias-mãe, não alteradas, e usou-se este produto para transformar bactérias quimicamente competentes DH5 $\alpha$ . Após extracção do DNA plasmídico, a sequência de todos as construções foi confirmada por sequenciação automática.

### 2.1.3 Construções para o estudo do NMD

Para estudar a possível degradação por NMD usou-se um gene repórter, a  $\beta$ -globina, no qual se substituiu a sua região 5' pela região 5' líder da *HJV*, bem como por esta mesma região mas apresentando mutações nos diferentes uAUGs. Usou-se as construções já efectuadas para o estudo das uORFs para amplificar a 5' UTR da *HJV*, intacta ou com mutações nos uAUGs, usando oligonucleótidos iniciadores contendo *linkers* com locais de restrição para os enzimas *Nhe* I e *Nco* I (Quadro I, Anexo). Estes produtos foram clonados no vector PTRE- $\beta$ N, em substituição da região 5' do gene  $\beta$ -globina entre os mesmos locais de restrição. Usou-se este produto para transformar bactérias quimicamente competentes DH5 $\alpha$ . Após extracção do DNA plasmídico, a sequência de todos as construções foi confirmada por sequenciação automática.

## 2.2 Culturas celulares e transfecções

A linha celular HeLa com expressão estável do trans-activador tet (HeLa/*tTA*) foi mantida em cultura em meio DMEM + Glutamax-I (*Gibco*) suplementado com 10% de soro fetal bovino (FBS), numa estufa a 37°C com atmosfera de CO<sub>2</sub> a 5%. A linha celular HepG2

foi mantida em meio RPMI 1640 + Glutamax-I (*Gibco*), suplementado com 10% FBS e numa estufa a 37°C com atmosfera de CO<sub>2</sub> de 5%. As linhas celulares referidas foram semeadas em placas de 35mm com 2mL de meio, usando uma quantidade de células de modo a que no dia seguinte se encontrassem a uma confluência de cerca de 90% no caso das HeLa e de 70% no caso das HepG2. As células foram incubadas durante cerca de 24h numa estufa a 37°C com atmosfera de CO<sub>2</sub> a 5%. Diluiu-se 4μL de lipofectamina<sup>TM</sup> 2000 (*Invitrogen*) em 125μL de Opti-MEM (*Gibco*), seguido de incubação 5min à temperatura ambiente (TA). A esta preparação juntou-se o DNA diluído em 125μL de Opti-MEM. No caso dos vectores de expressão com a variante A clonada e o gene *HJV* usou-se 1500ng de cada plasmídeo, para os vectores de expressão para o estudo das uORFs 5ng do vector de expressão pGL4.70+ hCMV e 2μg do vector de expressão em questão, de forma a ter uma razão entre os dois vectores de 1:400, e no caso dos vectores de expressão para o estudo do NMD usou-se 300ng do vector de expressão em questão e 2700ng do vector de expressão vazio. Juntou-se as duas preparações, incubou-se 20min à TA e adicionou-se ao meio de cultura de cada placa, previamente substituído. As placas foram de novo colocadas na estufa a 37°C, durante cerca de 24h.

### **2.3 Extracção de RNA, síntese de cDNA e PCR**

Obteve-se o RNA através de lise directa usando o *kit* de extracção *NucleoSpin RNA II* (*Macherey-Nagel*), e procedeu-se de acordo com o manual do fabricante. Os cDNAs foram sintetizados através do *kit SuperScript<sup>TM</sup> II Reverse Transcriptase*, de acordo com o manual do fabricante. As diferentes reacções de PCR para a detecção das variantes de *splicing* foram efectuadas usando conjuntos de oligonucleótidos iniciadores específicos (*vide* resultados e Quadro I, Anexo). As condições de PCR utilizadas foram equivalentes para a amplificação das diferentes variantes, sendo constituídas por um passo inicial de desnaturação a 95°C durante 5min, e depois 30 ciclos de amplificação, constituídos por uma desnaturação a 95°C durante 30seg, hibridação a 55°C durante 45seg e um passo de extensão a 72°C durante 1min. Fez-se também um passo final de extensão a 72°C durante 10min. A mistura de reacção foi feita de acordo com o manual do fabricante (*Promega*), à qual se adicionou o DNA molde, variável para cada caso. Os fragmentos obtidos foram analisados em gel de agarose (1%-2%) e a sua sequência confirmada por sequenciação automática.

## 2.4 Transcrição *in vitro*

O RNA foi sintetizado usando o kit *RiboMAX<sup>TM</sup> Large Scale RNA Production System – T7*, de acordo com o manual do fabricante. Após digestão do template com DNase durante 15min a 37°C, purificou-se e extraiu-se o RNA através do kit de extracção *NucleoSpin RNA II (Macherey-Nagel)*.

## 2.5 Ensaio de luminescência

Este ensaio permite quantificar a produção da proteína luciferase através da medição da sua actividade enzimática na presença de um substrato, um derivado do luminol, que ao ser metabolizado emite fótons que podem detectados usando um luminómetro. A actividade da luciferase foi medida usando o *Dual Luciferase Reporter Assay System (Promega)* e o luminómetro *Lucy 2 (Anthos)*. Após 24h de transfecção, procedeu-se à obtenção de extractos proteicos através da lise das células, fazendo inicialmente uma lavagem com 500µL de solução salina de tampão fosfato (PBS), seguida da adição de 50µL de *passive lysis buffer 1 x (PLB, Promega)* por poço, seguida de raspagem. Pipetou-se o conteúdo da placa, que foi de seguida congelado a -80°C, descongelado a 37°C e centrifugado a 4°C durante 30seg, descartando-se o *pellet* e guardando o lisado. A análise da actividade da luciferase consiste na medição sequencial da actividade da luciferase do pirilampo (neste caso, clonada no vector de expressão em estudo) e da *Renilla* (clonada no vector de expressão pGL4.70) na mesma reacção. Ao lisado (10µL) é adicionado o substrato da luciferase do pirilampo (40µL) e a sua actividade é medida; de seguida é adicionado o substrato da *Renilla* (40µL), que pára esta reacção, e estimula simultaneamente a actividade da luciferase da *Renilla*, permitindo depois a sua medição. A luciferase da *Renilla* funciona como controlo interno da reacção, permitindo eliminar variabilidade técnica como, por exemplo, diferentes eficiências de transfecção ou número de células por placa.

## 2.6 RT-PCR quantitativo

Após retirar-se a alíquota para medição dos níveis de luminescência, ressuspendeu-se o *pellet* e adicionou-se o tampão de lise do kit de extracção *NucleoSpin RNA II (Macherey-Nagel)*, procedendo-se depois à extracção do RNA de acordo com o manual do fabricante. O cDNA foi sintetizado recorrendo-se às esferas comerciais *Ready-To-Go You-Prime First-Strand Beads, GE Healthcare*, e a reacção de PCR quantitativo efectuada usando a mistura de reacção *SYBR green PCR master mix, Applied Biosystems*, ambos de acordo com as

instruções do fabricante. Relativamente à quantificação dos transcritos correspondentes ao estudo das uORFs, usou-se oligonucleótidos iniciadores específicos para do gene da luciferase do pirilampo e da *Renilla*. (Quadro I, Anexo). No caso dos transcritos usados para o estudo do NMD, usou-se oligonucleótidos iniciadores específicos para o gene  $\beta$ N e a Puromicina, aqui usada como controlo interno. A Puromicina corresponde a uma unidade transcricional independente, clonada no vector de expressão onde se encontra clonado o gene repórter  $\beta$  (Quadro I, Anexo). Para ambos os casos, o programa consistiu num passo inicial de desnaturação a 95°C durante 10min, seguido de quarenta ciclos que incluem um passo de desnaturação a 95°C por 15seg e um passo de extensão a 60°C por 30seg. A reacção foi quantificada usando o aparelho de PCR em tempo real *ABI PRISM 7000 (Applied Biosystems)* e o sistema de detecção *SYBR GREEN*.

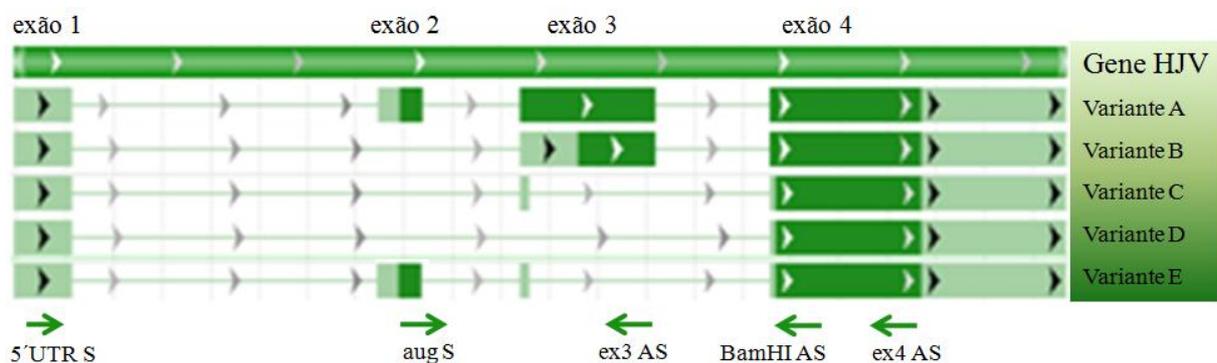
## **2.7 Imuno-deteccção por *Western-Blot***

Após 24h de transfecção, as células HeLa foram lisadas como referido para os ensaios de luminescência. A 20 $\mu$ L do lisado foi adicionado 5 $\mu$  de dodecil sulfato de sódio (SDS) 5x, desnaturou-se a mistura a 95°C durante 10min e armazenou-se a -20°C. Os lisados proteicos foram resolvidos num gel de SDS-PAGE a 10% e transferidos para uma membrana de PVDF (*Polyvinylidene Fluoride*) (*Bio-Rad*). Esta foi marcada com o anticorpo primário de rato anti-Luciferase (*Invitrogen*) a uma diluição de 1:200 e com o anticorpo primário de rato anti- $\alpha$  Tubulina (*Sigma*), numa diluição de 1:5000. Após lavagem da membrana, a deteccção foi feita usando o anticorpo secundário de cabra anti-IgG de rato (*Abcam*) a uma diluição de 1:4000, conjugado com HPR (*horseradish-peroxidase*), seguido de quimioluminescência. A revelação foi efectuada no processador Kodak X-OMAT 1000.

### 3 Resultados

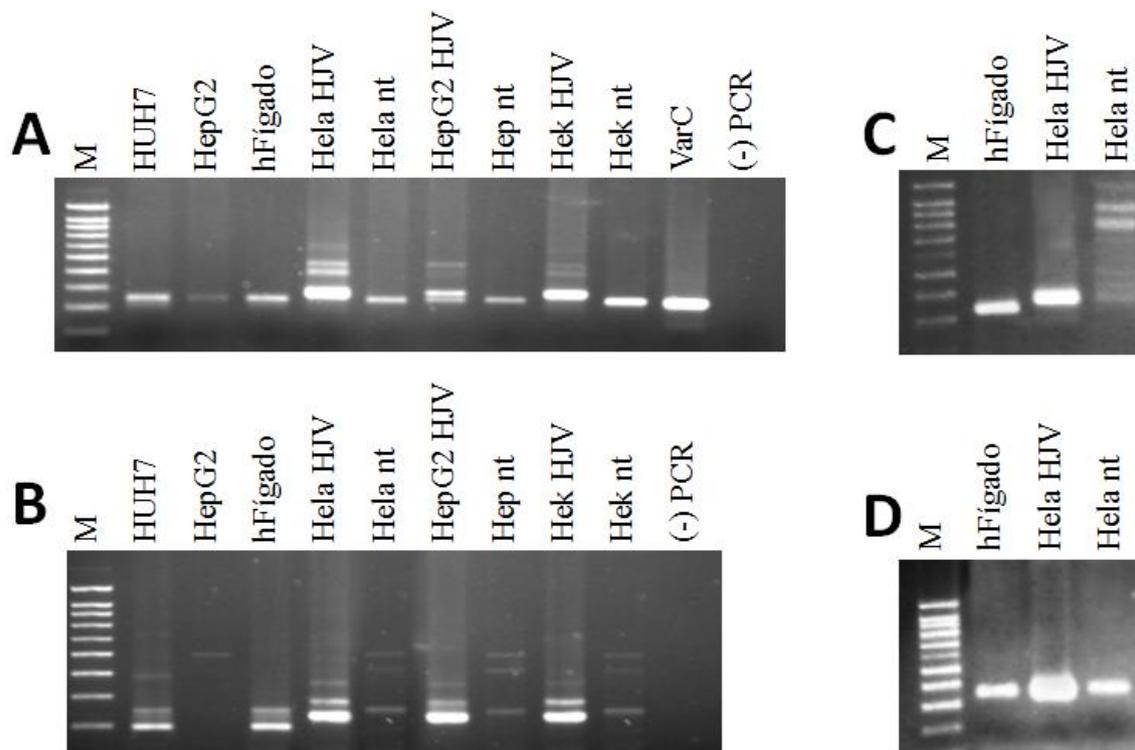
#### 3.1 Análise do padrão de *splicing* do transcrito *HJV*

Tal como foi detalhadamente descrito na introdução desta dissertação, o gene *HJV* é transcrito e processado num mRNA maduro, contendo a totalidade dos exões (variante A) e em quatro variantes de *splicing* adicionais (variantes B, C, D e E) [16]. Usando oligonucleótidos específicos para cada exão, é possível discriminar estas variantes através de amplificação por reacção de polimerização em cadeia (PCR) e visualização em gel de agarose, permitindo a análise do seu padrão de *splicing* (Figura 3).



**Figura 3 - Representação do gene *HJV* e das suas variantes de *splicing*.** Abaixo estão representados, no local correspondente de hibridação, os oligonucleótidos iniciadores usados nas amplificações por PCR. (Imagem adaptada de <http://www.ncbi.nlm.nih.gov/gene/148738>).

Para tal, fez-se inicialmente PCR usando um oligonucleótido iniciador directo no primeiro exão (5'UTR S – Quadro I, Anexo) e um oligonucleótido iniciador reverso no quarto exão (BamHI AS – Quadro I, Anexo), comuns a todas as variantes. Assim, com estes oligonucleótidos iniciadores, os produtos de PCR resultariam em bandas de peso molecular de 1046, 861, 339, 302 e 488 pares de bases (pb), correspondentes respectivamente às variantes A, B, C, D e E. No caso das variantes obtidas a partir do gene *HJV* transfectado, os respectivos pesos moleculares seriam acrescidos de 36 pb correspondentes a um *tag* inserido no início do exão 3, comum a todas as variantes com excepção da variante D. Como DNA molde foi usado cDNA produzido a partir de RNA de fígado humano, de RNA obtido a partir das linhas celulares HUH7, HepG2, HeLa e HEK não transfectadas ou das linhas celulares HeLa, HepG2 e HEK transfectadas com o plasmídeo de expressão contendo o gene *HJV*.

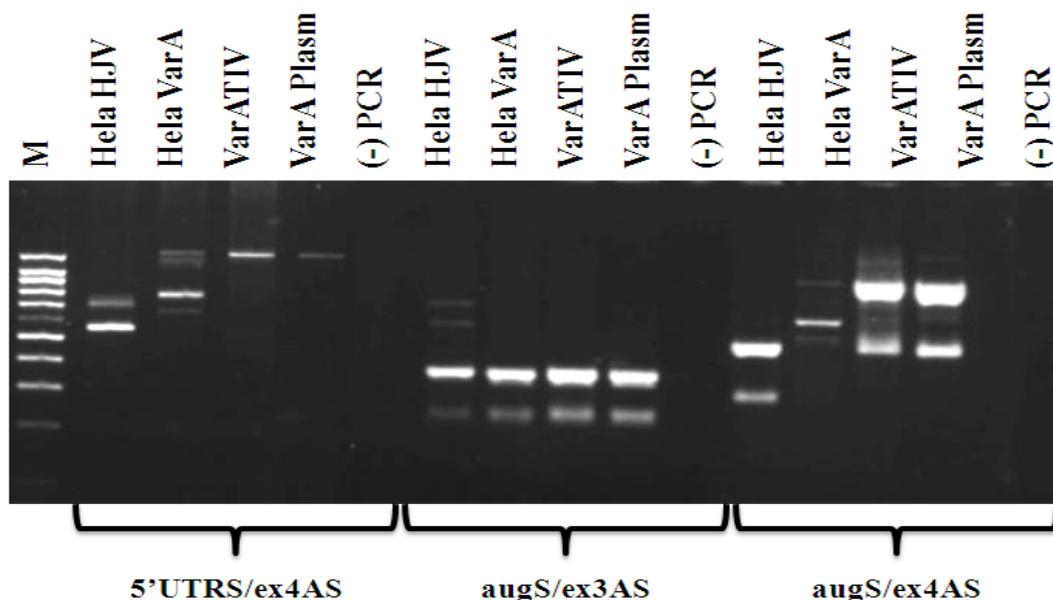


**Figura 4 - Amplificação de cDNA *HJV* em diferentes tecidos.** Os oligonucleótidos iniciadores usados são **A)** 5'UTR S e BamHI AS, comuns a todas variantes; **B)** aug S e BamHI AS, que permite a amplificação das variantes A e E; **C)** 5'UTR S e ex3 AS, que permite a amplificação das variantes A e B; **D)** aug S e ex3AS, que permitem a amplificação exclusiva da variante A. O controlo positivo var C usado na figura A corresponde à variante C clonada num vector de expressão. O M corresponde ao marcador de peso molecular *100bp ladder*; a banda com maior peso molecular é de 1200 pb, a seguinte de 1000 pb, e as seguintes têm uma diferença de 100 pb. É visível em todos os casos a diferença de peso molecular entre as bandas relativas a linhas transfectadas e não transfectadas, correspondentes à presença do *tag* (36 pb) na *HJV* clonada no vector de expressão transfectado. O peso molecular da variante A, que não é detectada nos casos A, B e C, seria de, respectivamente, 1082, 844 e 407 pares de bases (com *tag*). Na figura D o seu peso molecular é de 356 pb. A sequência de todos os produtos de PCR foi confirmada por sequenciação. Confirmou-se igualmente por sequenciação que as bandas visíveis de menor intensidade correspondem a produtos de PCR inespecíficos. Foram efectuados três ensaios independentes com resultados muito semelhantes, sendo esta figura exemplo de um dos ensaios.

A análise em gel de agarose (figura 4A) permite a observação de uma banda de peso molecular correspondente à variante C e, nalguns casos, a variante D. Pode comparar-se o peso molecular directamente com o controlo positivo, a variante C clonada. Surpreendentemente, a variante A, cuja isoforma representa a proteína funcional, não é detectada usando este conjunto de oligonucleótidos. Fez-se também PCR usando um oligonucleótido iniciador directo no exão 2 (aug S) e um inverso no exão 4 (BamHI AS). Este conjunto de oligonucleótidos permite a amplificação selectiva das variantes A e E, com pesos moleculares respectivamente de 808 e 286 pb, acrescidos de 36 pb correspondentes ao *tag*, no caso dos tecidos transfectados. Mais uma vez, este PCR apenas permitiu a visualização de um produto correspondente à variante E (figura 4B) não sendo novamente detectada a variante A.

Igualmente, ao usar oligonucleótidos iniciadores para o primeiro e terceiro exões (respectivamente, 5'UTR S e ex3 AS), que permitem a amplificação selectiva das variantes A e B, excluindo a amplificação das variantes C e E, foi apenas possível a detecção da variante B (figura 4C). Os pesos moleculares das variantes A e B são respectivamente de 558 e 373 pb, acrescidos de 36 pb correspondentes ao *tag*. Finalmente, apenas fazendo um PCR usando oligonucleótidos iniciadores específicos para a variante A (aug S e ex3 AS – Quadro I, Anexo), que excluem a amplificação de todas as outras variantes, é possível a visualização da banda correspondente à variante A em gel de agarose (figura 4D), com peso molecular de 320 pb, acrescidos de 36 pb. Estes resultados levam-nos a levantar a hipótese de que as variantes A, B e D estão presentes nos tecidos e são produzidas quando transfectadas mas em quantidades muito baixas, e que as variantes maioritárias parecem ser a C e a E.

Existem no entanto dois aspectos a ter em consideração no que diz respeito à variante E. Primeiro, é curioso notar que embora esta variante seja inicialmente referida no artigo de Papanikolaou e colaboradores (2004) [16] a sua sequência não se encontra anotada nas bases de dados. Segundo, o facto de a variante E diferir da variante A apenas na ocorrência de um *skipping* parcial do exão 3 leva a equacionar a hipótese da sua detecção por RT-PCR ser apenas um artefacto técnico, seja na síntese dos cDNAs ou na própria reacção de PCR, iludindo-nos na interpretação dos resultados. Para testar esta hipótese o mRNA correspondente à variante A foi produzido através de transcrição *in vitro* (VarA TIV) a partir do respectivo cDNA clonado num vector de expressão. O produto desta reacção foi utilizado para RT-PCR usando três conjuntos de oligonucleótidos iniciadores, de maneira a ser tecnicamente possível amplificar todas as variantes (5'UTR S/ex4 AS), apenas a variante A (aug S/ex3 AS) e as variantes A e E (aug S/ex4 AS). Paralelamente, foi também utilizado RNA da linha celular HeLa transfectada com o gene *HJV* (HeLa HJV) ou transfectada com a variante A clonada num vector de expressão (HeLa VarA) ou, como DNA molde, o próprio plasmídeo que contém a variante A clonada (VarA Plasm). O resultado obtido encontra-se na figura 5.



**Figura 5 – Amplificação de HeLa HJV, HeLa VarA, VarA TIV e VarA Plasm usando três conjuntos de oligonucleótidos iniciadores.** O M corresponde ao marcador de peso molecular *100bp ladder*; a banda com maior peso molecular é de 1200 pb, a seguinte de 1000 pb, e as seguintes têm uma diferença de 100 pb. O peso molecular da variante A usando estes oligonucleótidos é de: 5'UTR S/ex4AS – 1260; aug S / ex3AS - 356; aug S / ex4AS - 1022 pares de bases (com *tag*). A sequência de todos os produtos de PCR foi confirmada por sequenciação. Confirmou-se igualmente por sequenciação que as bandas visíveis de menor intensidade correspondem a produtos de PCR inespecíficos. Foram efectuados três ensaios independentes com resultados muito semelhantes, sendo esta figura exemplo de um dos ensaios.

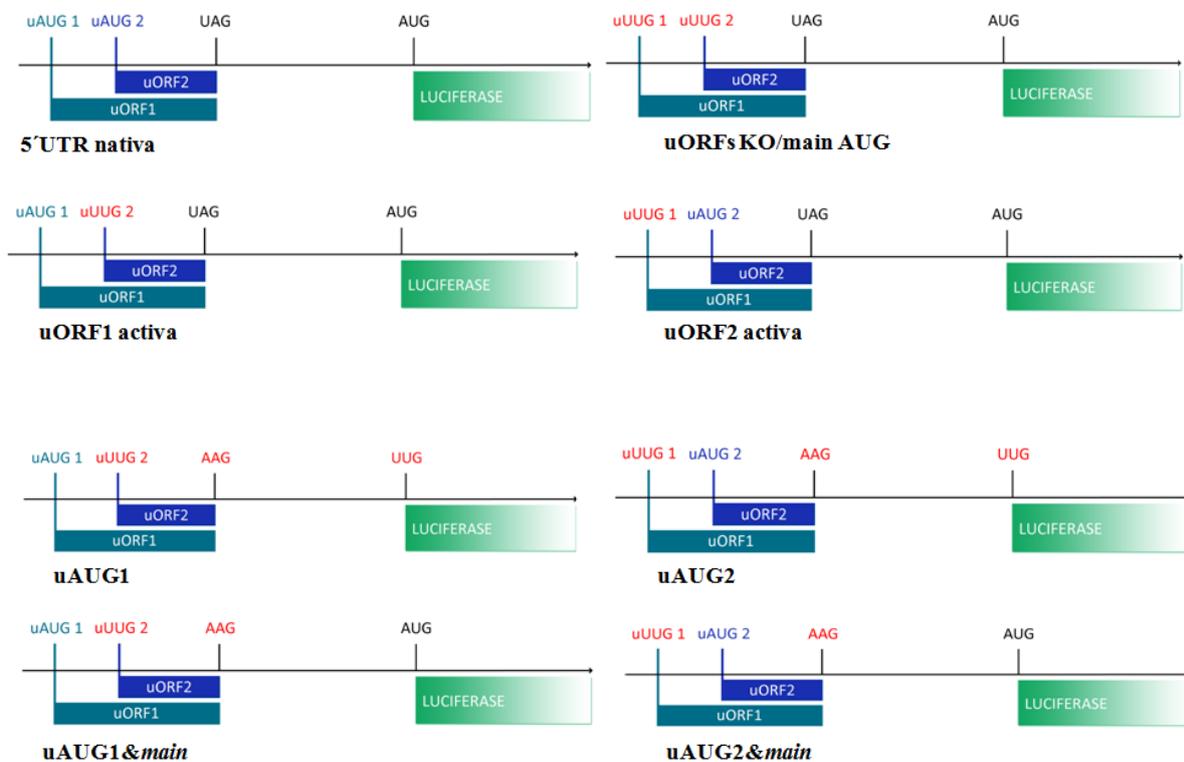
Assim, e de acordo com os resultados anteriores, considerando o conjunto de oligonucleótidos iniciadores que amplifica todas as variantes (5'UTR S/ex 4AS), observa-se que a variante produzida maioritariamente a partir do gene *HJV* transfectado é a variante C, com peso molecular de 553 pb. Surpreendentemente, o produto de PCR maioritariamente detectado com este conjunto de oligonucleótidos, produzido a partir da transfecção do cDNA correspondente à variante A, corresponde à variante E (738 pb) e não à variante A (1260 pb), tal como seria previsível tendo em conta que não seria expectável que este cDNA fosse sujeito a qualquer evento de *splicing* durante o seu processamento intracelular. É de notar, contudo, uma banda de baixa intensidade com peso molecular semelhante ao da variante A, e que poderá eventualmente corresponder a esta. No entanto, não foi possível confirmar por sequenciação a sua identidade. Em contrapartida, com este mesmo conjunto de oligonucleótidos, já é possível a detecção de uma banda correspondente à variante A ao ser usado como DNA molde o cDNA produzido a partir de RNA transcrito *in vitro* ou o plasmídeo que contém o respectivo cDNA desta variante. Usando o conjunto de oligonucleótidos iniciadores específico para a variante A (aug S/ex3AS), é visível uma banda em todas as amostras de tamanho correspondente à mesma, 356 pb, logo, a variante A é produzida mesmo nos dois primeiros casos em que não é detectada com o anterior conjunto

de oligonucleótidos. No caso do conjunto de oligonucleótidos iniciadores que permite, com base no peso molecular, a distinção entre a variante A e a E (aug S/ex4AS), é visível, no caso do gene transfectado, uma banda correspondente à variante E (500 pb) e, nos casos do cDNA produzido a partir de RNA transcrito *in vitro* e do plasmídeo que contém clonado o cDNA da variante A, bandas correspondentes à variante A, de 1022 pb.

Estes resultados sugerem que a não detecção da variante A nas situações acima descritas não parece ser devida a questões técnicas inerentes às próprias reacções de RT e/ou PCR, uma vez que ocorre amplificação e detecção desta quando se utiliza como substratos RNA produzido *in vitro* a partir do cDNA da variante A ou plasmídeo contendo o mesmo cDNA. Aparentemente, a detecção da variante A, quando se utilizam oligonucleótidos que não a amplificam exclusivamente, apenas é possível quando o RNA a partir do qual é feito o cDNA não é produzido na célula. Parece estar a ocorrer um evento celular ainda não esclarecido que influencia o *splicing* do transcrito *HJV*, favorecendo a produção das variantes C e E. A variante C, aparentemente maioritária, contém o codão de iniciação no quarto exão, produzindo-se uma proteína apenas com o domínio transmembranar, cuja função não é conhecida mas que, apesar de não conter o péptido sinal característico da isoforma A, potencialmente poderá chegar à membrana e exercer algum efeito na via de sinalização *HJV/BMP*. [16] A variante E contém um codão de iniciação no exão 2 (o mesmo da variante A), que ao ser utilizado originaria um *frameshift* de 31 codões a jusante que introduz um codão de terminação prematuro, produzindo uma proteína truncada com um total de 70 codões, sem qualquer função conhecida. Alternativamente, poderá ser usado o codão de iniciação no exão 4, produzindo-se uma isoforma igual à das variantes C e D. Em ambos os casos, a *HJV* funcional com o péptido sinal, RGD (motivo de três aminoácidos) e vWf não é produzida. A produção das isoformas C, D e talvez mesmo a E, pode estar relacionada com a função reguladora da *HJV*, cuja concentração deve portanto ser finamente controlada. Seria um desperdício celular a *HJV* funcional ser produzida em grandes quantidades quando não seria necessária a sua utilização. O *splicing* alternativo pode assim contribuir para a regulação da expressão da *HJV*, favorecendo a produção de proteínas com ausência de domínios específicos no caso de não haver necessidade de proteína funcional, mantendo simultaneamente a transcrição activa para assegurar uma resposta rápida caso a concentração celular de ferro se altere subitamente. Conjuntamente, ocorreria igualmente alguma produção a nível basal da *HJV*, o que explica a detecção da variante A em quantidades muito baixas.

### 3.2 Controlo da tradução da HJV por uORFs

Como referido na introdução desta dissertação, a região 5' líder do transcrito da HJV contém dois codões de iniciação a montante do codão de iniciação principal, sendo assim definidas duas uORFs de 28 e 19 codões, que partilham o mesmo codão STOP. De forma a avaliar se estas uORFs exercem algum efeito na expressão da HJV foram realizadas construções em que a sequência 5' líder do transcrito luciferase (gene repórter) foi substituída pela sequência 5' líder do transcrito da HJV. Esta construção foi denominada 5'UTRHJV nativa e a sua sequência confirmada por sequenciação automática. A partir desta construção foram criadas novas construções tendo sido introduzidas mutações alternadas nos 3 codões de iniciação e no codão STOP. As construções efectuadas encontram-se representadas na figura 6, com a respectiva designação.

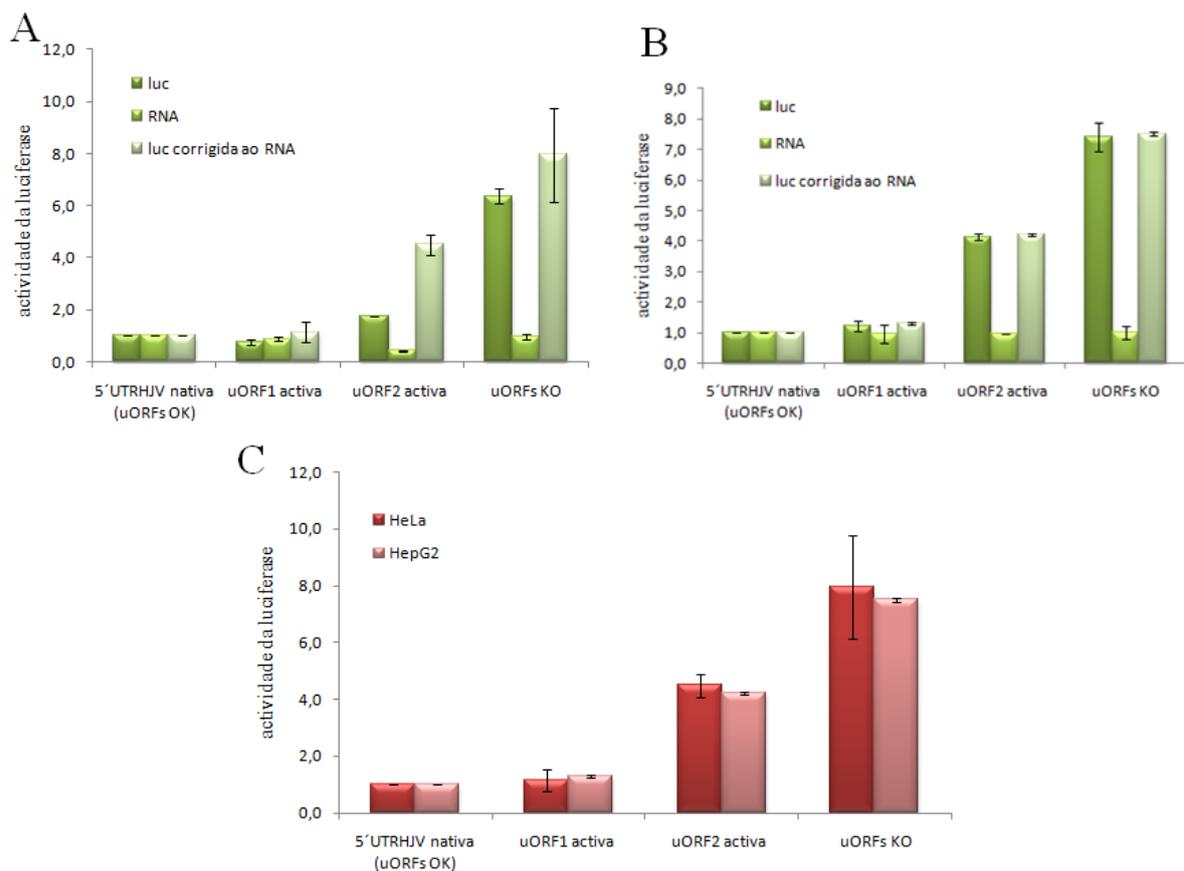


**Figura 6** – Representação, com a respectiva designação, das oito construções efectuadas. Os codões a vermelho representam os codões mutados.

Estas construções foram usadas para transfectar as linhas celulares HeLa e HepG2, permitindo a produção de mRNAs e tradução nas proteínas correspondentes. A quantidade de proteína produzida foi avaliada por ensaios de luminescência, através da medição da

actividade enzimática da luciferase na presença do seu substrato, um derivado do luminol, que ao ser metabolizado emite fótons que podem ser detectados usando um luminómetro.

Inicialmente averiguou-se se as uORFs eram funcionais nos dois tecidos. Para tal, transfectaram-se as construções 5'UTR HJV nativa, uORFs KO, uORF1 activa e uORF2 activa e usaram-se os extractos proteicos para determinar os níveis de luminescência. Em paralelo, determinaram-se os níveis de RNA correspondentes por RT-PCR quantitativo. A figura 7 representa os resultados obtidos para HeLa (figura 7A) e HepG2 (figura 7B), nomeadamente os valores de luminescência, os níveis de RNA e os valores de luminescência corrigidos aos níveis de RNA. É importante fazer a correcção dos níveis de luminescência aos níveis de RNA por forma a eliminar qualquer eventual variabilidade ao nível, por exemplo, da eficiência da transfecção, garantindo assim que estamos apenas a avaliar um efeito da eficiência da tradução dos transcritos repórter. A figura 7C representa a comparação entre os níveis de luminescência corrigidos aos níveis de RNA em células HeLa e HepG2.

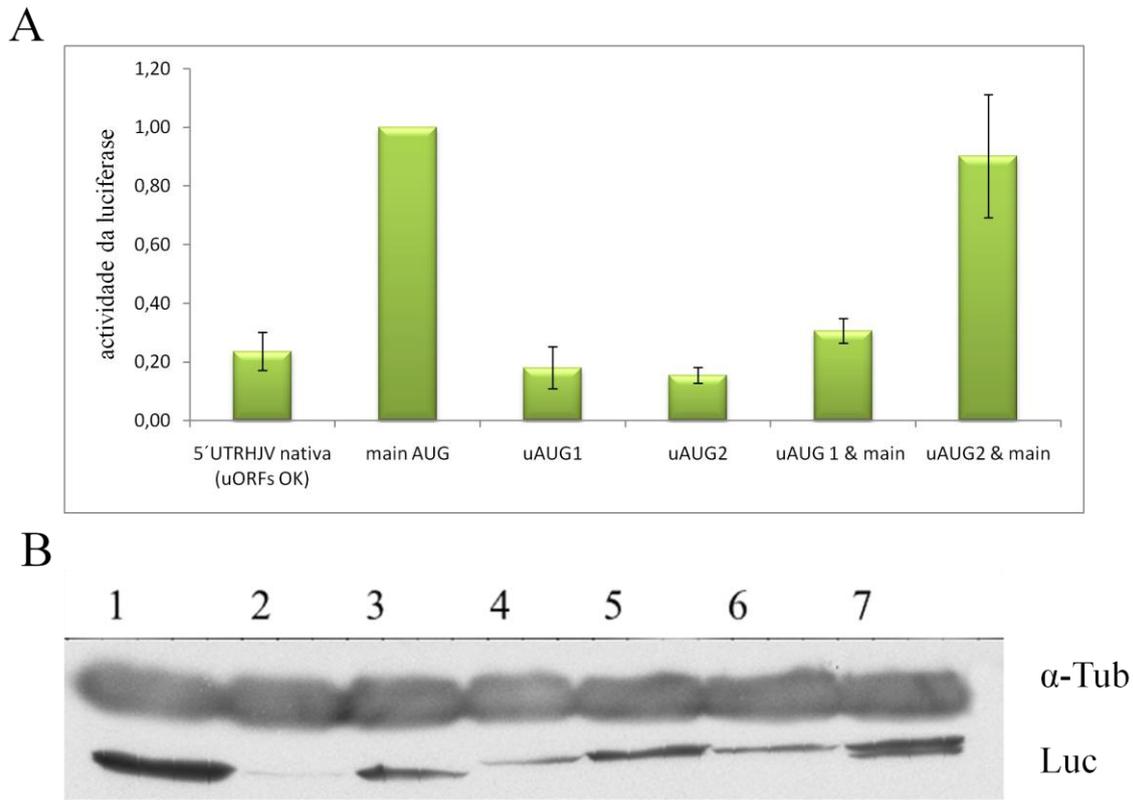


**Figura 7 – Estudo da funcionalidade das uORFs presentes no transcrito *HJV*.** **A**- Representação dos níveis de luminescência, dos níveis de RNA e da correcção dos níveis de luminescência aos níveis de RNA em HeLa. **B** – O mesmo que A, mas para HepG2. **C** – Comparação dos níveis de luminescência corrigidos aos de RNA em ambos os tecidos. A normalização é feita em relação às uORFs OK., que representa a situação em que as uORFs estão activas. Foram efectuados três ensaios independentes tanto para a quantificação da proteína como do mRNA.

Pode observar-se que no caso das uORFs KO, em que não pode ocorrer inibição da expressão da ORF principal porque ambas uORFs têm o seu codão de iniciação mutado, a quantidade de proteína é cerca de 8 vezes superior ao caso em que ambas as uORFs se encontram activas (uORFs OK). Isto significa que quando as uORFs estão presentes e activas, os níveis de proteína diminuem consideravelmente, o que sugere um papel destas na inibição da expressão proteica. No caso em que apenas a uORF1 está activa, os níveis de proteína são semelhantes ao caso em que ambas estão activas (uORFs OK), o que sugere que a uORF1 é responsável pela maior parte da inibição. Esta análise é reforçada pelo caso em que apenas a uORF2 está activa, em que os níveis de proteína são cerca de 5 vezes superiores ao caso em que ambas estão activas, sugerindo que esta não tem um papel tão determinante como a uORF1 na inibição da tradução da uORF principal. No entanto, é curioso notar que quando ambas uORFs estão inactivas a expressão é superior a qualquer um dos casos em que cada uma está inactiva individualmente, e mesmo superior à soma da expressão destas, sugerindo um efeito sinérgico entre as duas uORFs. Parece ocorrer algum tipo de cooperação entre elas de maneira a potenciar a inibição da expressão proteica quando ambas estão activas e não apenas cada uma exercer o seu papel individualmente.

Apesar da função das uORFs ser controlar o nível de proteína produzida inibindo a tradução da ORF principal, é necessário que ocorra alguma síntese da proteína para que esta possa exercer a sua função na célula [2] [30]. Segundo o modelo convencional de iniciação da tradução, o complexo de iniciação contendo a subunidade ribossomal 40S percorre a região 5' líder do transcrito até ao reconhecimento do primeiro AUG, sendo então iniciada a síntese proteica. No caso de existir alguma uORF, o primeiro AUG não corresponde ao AUG da ORF principal, pelo que é necessário existirem mecanismos que permitam o seu reconhecimento. Os dois mecanismos principais que permitem a tradução da ORF principal na presença de uORFs são o *leaky scanning* e a reiniciação [29]. De maneira a esclarecer se neste caso o AUG principal poderia estar a ser reconhecido por *leaky scanning*, transfectou-se a linha celular HeLa com as construções 5'UTRHJV nativa, uORFs KO (neste caso, denominado *main* AUG, uma vez que é o único que está activo), uAUG1 (segundo o mesmo raciocínio, por ser o único activo), uAUG2, AUG1&*main* e AUG2&*main*. Os extractos proteicos foram quantificados por luminescência e analisados por *Western-Blot*, e os níveis de RNA determinados por RT-PCR quantitativo para, mais uma vez, assegurar que qualquer variação observada se deve exclusivamente a efeitos na eficiência da tradução. As construções uAUG1 e uAUG2 apenas possuem um codão de iniciação, portanto a tradução tem obrigatoriamente de ser iniciada nesse codão, formando-se uma proteína de fusão, semelhante à proteína da

luciferase, mas com uma porção N-terminal mais longa, correspondente à uORF e à região inter-cistrónica. Já no caso das construções *uAUG1&main* e *uAUG2&main*, a iniciação da tradução pode ser feita no codão de iniciação da uORF ou no codão de iniciação principal. Caso ocorra *leaky scanning*, o ribossoma inicia a tradução ora num ora noutro codão, produzindo-se tanto proteína de fusão como a proteína da luciferase. Os níveis de luminescência (já corrigidos aos níveis de RNA) e representação da imuno-deteccção por *Western-Blot* estão resumidos na figura 8.



**Figura 8. Mecanismo pelo qual o AUG da ORF principal é reconhecido.** **A** – Representação dos níveis de luminescência corrigidos aos níveis de RNA. A normalização foi feita à construção *main AUG*, que representa a situação em que a tradução apenas pode iniciar no AUG principal. **B** – Representação da imuno-deteccção por *Western-Blot*. **Legenda:** 1- luciferase; 2- 5'UTRHJV nativa; 3- *main AUG*; 4- uAUG1; 5- uAUG2; 6- uAUG1&*main*; 7- uAUG2&*main*. Foram efectuados três ensaios independentes tanto para a quantificação e deteção da proteína como para a quantificação do mRNA.

Como é visível, detecta-se apenas proteína de fusão nos casos uAUG1, uAUG2 e uAUG1&*main*, e na construção uAUG2&*main* observam-se duas bandas correspondentes às duas proteínas produzidas, de fusão e da proteína da luciferase, compatíveis com a ocorrência de *leaky scanning*. Faz sentido este ser maioritário neste caso uma vez que o AUG da uORF1 possui um contexto de iniciação mais favorável do que o AUG da uORF2, podendo assim ser mais fácil para o ribossoma “falhar” o reconhecimento deste último. Não é visível a banda correspondente à proteína da luciferase no caso uAUG1&*main*, mas tal não implica que não

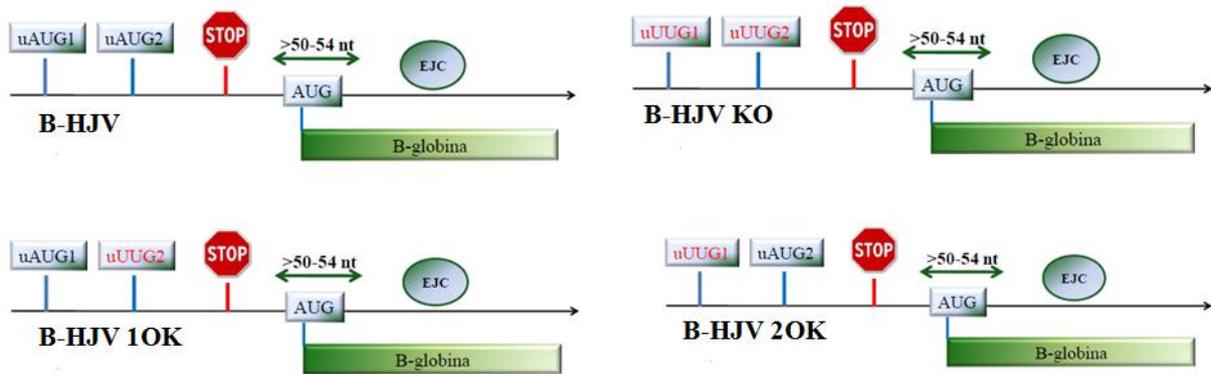
está a ocorrer *leaky scanning*. Este poderá estar a ocorrer a um menor nível, não detectável através desta técnica. Estes resultados indicam que o ribossoma reconhece o codão de iniciação da ORF principal, pelo menos quando apenas a uORF2 se encontra activa, por *leaky scanning*. Usando estas construções não é possível estudar a existência de reiniciação uma vez que o codão STOP das uORFs está mutado, não permitindo a terminação da tradução e subsequente reiniciação. É ainda evidente uma relação entre os níveis de luminescência e as bandas proteicas detectadas. Nos casos em que apenas se produz proteína de fusão, ou seja, no caso uAUG1, uAUG2 e uAUG1&main, os níveis de luminescência são semelhantes entre si, com um ligeiro aumento no caso uAUG1&main. No caso em que se observa ocorrência de *leaky scanning*, no uAUG2&main, em que se produz proteína de fusão e a proteína da luciferase, os níveis de luminescência são semelhantes ao caso em que apenas se pode produzir proteína da luciferase por apenas estar activo o AUG principal (*main* AUG). Assim, o ligeiro aumento do valor de luminescência no caso uAUG1&main face aos casos uAUG1 e uAUG2 pode dever-se à produção de proteína a partir do AUG principal em pequenas quantidades, compatível com a ocorrência de *leaky scanning* a níveis não detectáveis por *Western-Blot*.

É ainda curioso observar que as bandas proteicas do *Western-Blot* não têm diferenças de intensidade suficientemente significativas que justifiquem a diferença considerável nos níveis entre 5'UTRHJV nativa, uAUG1, uAUG2 e uAUG1&main e entre *main* AUG e uAUG2&main. Estes resultados indicam que a diminuição dos níveis de luminescência não se deve exclusivamente a diferentes quantidades de proteína, sugerindo que a proteína de fusão que é produzida no caso em que o codão STOP e o AUG principal estão mutados pode ter uma actividade enzimática diferente da proteína produzida a partir apenas do AUG principal.

### **3.3 Controlo da degradação do transcrito *HJV* por NMD**

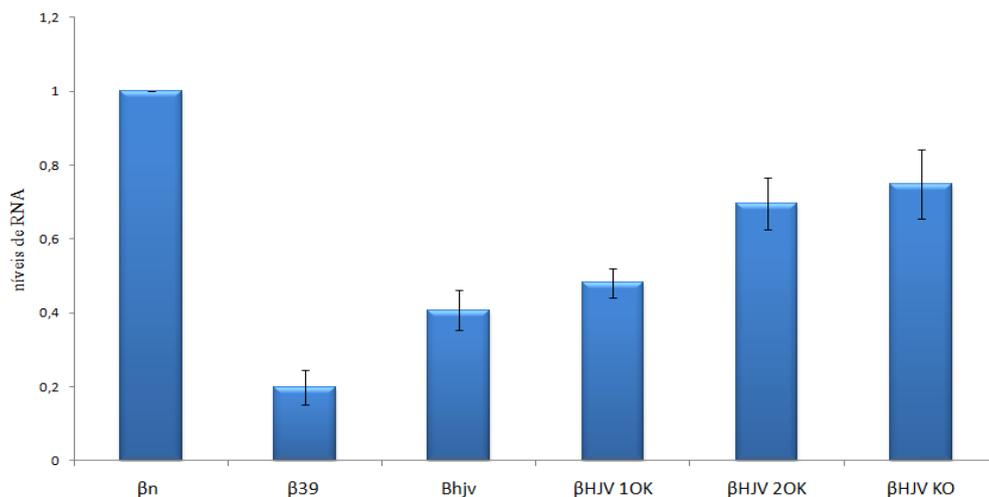
Como determinado no ponto anterior, o transcrito *HJV* possui na região 5' líder duas uORFs funcionais, às quais está associado um codão de terminação da tradução. Este poderá ser reconhecido como prematuro pela maquinaria de NMD, uma vez que se localiza a mais de 54 nucleótidos da última junção exão-exão [4]. Tendo em conta a dificuldade inerente à detecção do transcrito principal da *HJV*, o desenvolvimento de um método quantitativo que permitisse a sua análise directa de forma a avaliar a regulação da *HJV* por NMD seria praticamente inexecutável. Isto levou ao desenvolvimento de uma estratégia alternativa que possibilita avaliar de que modo a presença de uORFs funcionais na região 5' líder do transcrito

*HJV* está envolvida na activação do mecanismo de NMD. Para tal, usou-se um gene repórter, a  $\beta$ -globina humana, cuja relação com o NMD está já bastante bem estudada. [52] Assim, a região 5' líder do gene da  $\beta$ -globina foi substituída pela região 5' líder nativa da *HJV*, bem como por esta região apresentando mutações nos diferentes uAUGs. As construções produzidas estão resumidas na Figura 9.



**Figura 9 – Representação das construções efectuadas, e respectiva designação.** Os codões a vermelho representam os codões mutados. Quando a distância entre o codão STOP e o EJC é superior a 50-54 nucleótidos, é permitida a activação do NMD.

Estas construções foram transfectadas em células HeLa e quantificou-se o mRNA produzido por RT-PCR quantitativo. A utilização das construções em que se anulam as uORFs separadamente é importante na medida em que permite avaliar se ambas são responsáveis pela indução do NMD. Usou-se como controlos o gene  $\beta$ -globina normal ( $\beta$ -N) e o gene  $\beta$ -globina portador de uma mutação *nonsense* no codão 39 ( $\beta$ -39), que induz o NMD [52]. Os resultados estão esquematizados na figura 10.



**Figura 10 – Quantificação por RT-PCR quantitativo dos níveis de RNA.** A normalização foi feita à construção  $\beta$ N, que representa o controlo negativo, em que não há degradação do RNA pelo mecanismo de NMD. Foram efectuados três ensaios independentes para a quantificação do mRNA.

Pode observar-se que os níveis de mRNA correspondentes ao  $\beta$ -39 são bastante inferiores em comparação ao  $\beta$ -N, indicando que os controlos estão a comportar-se de maneira apropriada, validando o mecanismo de NMD neste sistema. Os níveis de mRNA da  $\beta$ -HJV são apenas ligeiramente superiores aos do  $\beta$ -39, indicando que a presença das uORFs da região 5' líder nativa da *HJV* tem a capacidade de induzir o NMD nos transcritos repórter. Pode observar-se ainda que os níveis de mRNA da  $\beta$ -HJV KO são apenas ligeiramente inferiores aos de  $\beta$ -N, ou seja, quando as uORFs estão inactivas os níveis de mRNA são semelhantes a uma situação em que o NMD não é activado, reforçando a hipótese de ser de facto o codão STOP das uORFs a ser reconhecido como prematuro, resultando na activação do NMD. Isto sugere que, na sua forma nativa, os transcritos *HJV* poderão ser igualmente substratos para a maquinaria de NMD. O facto de os níveis de mRNA da  $\beta$ -HJV serem semelhantes aos da  $\beta$ -HJV 1OK, e os da  $\beta$ -HJV 2OK semelhantes aos da  $\beta$ -HJV KO sugere que a uORF1 possa ser a mais determinante na activação do NMD. Estes dados estão de acordo com os resultados anteriores, em que se observou que a uORF1 tem um papel mais preponderante na inibição da tradução que a uORF2. Como o NMD depende do evento de terminação da tradução, se a uORF1 é traduzida mais vezes que a uORF2 faz sentido desencadear mais vezes o NMD. Pode ainda considerar-se o facto de a uORF1 ser maior que a uORF2 (28 *versus* 19 codões) o que, de acordo com exemplos anteriormente descritos na introdução [49] [50], pode ser um factor determinante na activação do NMD. Para testar esta hipótese poder-se-ia diminuir o número de codões da uORF1, mantendo o contexto do seu uAUG, e observar se mantinha a capacidade de indução de NMD.

## 4 Discussão final e perspectivas futuras

A HJV foi inicialmente descrita em 2004, e desde então os estudos realizados foram essencialmente funcionais. Tem sido atribuída particular importância ao estudo de mutações que alteram a funcionalidade da proteína, e a sua consequência na via de sinalização mediada por BMPs. No entanto, estudos recentes evidenciam que a HJV pode ser regulada não apenas a nível transcricional, mas essencialmente ao nível pós-transcricional, tornando-se assim fundamental perceber de que modo este tipo de regulação pode influenciar a expressão da HJV e que consequências advirão para a homeostase do ferro [18].

A análise do padrão de *splicing* em diferentes tecidos, transfectados e não transfectados, permite observar que a variante A, que codifica a proteína funcional, não é produzido em quantidades elevadas, que poderá estar relacionado com o papel regulador da proteína HJV. Foi também detectada em quantidades consideráveis uma variante de *splicing* inicialmente referida por Papanikolaou e colaboradores (2004), mas cuja sequência não se encontra anotada nas bases de dados do transcriptoma. A variante maioritariamente detectada parece ser a variante C, cuja função na célula não está ainda esclarecida. Pode assim concluir-se que o controlo do gene *HJV* através do *splicing* alternativo parece existir através do favorecimento da produção de proteínas sem alguns domínios funcionais em detrimento da proteína funcional. Não se sabe, contudo, qual o papel dessas isoformas proteicas na célula.

Foram também identificados dois mecanismos de regulação, um ao nível do controlo da tradução e outro ao nível da degradação do mRNA. Verificou-se que o transcrito *HJV* tem duas uORFs funcionais que inibem consideravelmente a eficiência de tradução da ORF principal. O AUG principal é reconhecido, pelo menos, por *leaky scanning*. De futuro, seria pertinente investigar se o AUG principal também é reconhecido através do mecanismo de reiniciação.

A HJV é também regulada ao nível da degradação do mRNA, nomeadamente através do mecanismo de NMD. Neste caso, as dificuldades técnicas inerentes à detecção da variante A no estudo do seu padrão de *splicing* fez questionar a viabilidade da sua detecção e quantificação. Isto obrigou a uma estratégia alternativa usando construções em que a região 5' líder de um gene repórter, a  $\beta$ -globina, cuja relação com o NMD está já bem estudada [52], foi substituída pela região 5' líder da *HJV*. Os resultados indicam que ocorre degradação do mRNA nestas construções, o que sugere fortemente que o mesmo acontecerá no transcrito *HJV* nativo. É no entanto necessário confirmar que esta degradação é devida ao NMD

mediante a observação de um aumento dos níveis destes transcritos em consequência da inibição deste mecanismo (através, por exemplo, da inibição por siRNA da expressão de uma das proteínas cruciais para este mecanismo, a UPF1). Seria ainda importante fazer estudos usando o transcrito *HJV* nativo, de maneira a concluir com toda a certeza que este mRNA é sensível ao NMD.

Pode assim concluir-se que existe um controlo rigoroso da *HJV* que não permite a sua expressão elevada, controlo este efectuado a vários níveis. Por um lado, o processamento do seu mRNA não favorece a produção da proteína funcional. Além disso, o próprio mRNA sofre degradação em níveis consideráveis, limitando fortemente a abundância de mRNAs para tradução. Mais ainda, há uma forte barreira à tradução pelo AUG principal através da existência de duas uORFs.

Seria ainda fundamental complementar estes estudos com experiências para avaliar a resposta destes mecanismos de regulação à exposição ao ferro. Por exemplo, seria interessante investigar se na presença de ferro a barreira imposta pelas uORFs à tradução da ORF principal era diminuída. Apesar de não terem sido observadas diferenças de expressão nem de controlo da tradução entre tecidos, é possível equacionar a hipótese destes mecanismos funcionarem em resposta a diferentes concentrações de ferro. Uma vez que a *HJV* estimula a produção de hepcidina, cujos níveis aumentam em resposta a concentrações elevadas de ferro, é possível que em situações de acumulação de ferro no organismo ocorra uma repressão destes mecanismos pós-transcricionais de maneira a aumentar a expressão da *HJV*. Um aumento da expressão da *HJV* traduzir-se-ia assim num aumento de expressão de hepcidina, que ao induzir a internalização das moléculas de ferroportina promove o decréscimo da libertação de ferro para o plasma.

## 5 Bibliografia

1. **López-Mauri, L, Marguerat, S and Bahler, J.** Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews Genetics*. 2008, Vol. 9, pp. 583-593.
2. **Meijer, HA and Thomas, AAM.** Control of eukaryotic protein synthesis by upstream open reading frames in the 5'-untranslated region of an mRNA. *Biochem. J.* 2002, Vol. 367, pp. 1-11.
3. **Bahler, J and Lackner, DH.** Translational Control of Gene Expression: from transcripts to transcriptomes. *International Review of Cell and Molecular Biology*. United Kingdom : Elsevier Inc., 2008.
4. **Silva, AL and Romao, L.** The mammalian nonsense-mediated mRNA decay pathway: To decay or not to decay! Which players make the decision? *FEBS Letters*. 2009, Vol. 583, pp. 499-505.
5. **Ganz, T.** Molecular Control of Iron Transport. *Frontiers in Nephrology*. 2007, Vol. 18, pp. 394-400.
6. **Graham, RM, et al.** Liver Iron Transport. *World J Gastroenterol*. 2007, Vol. 13(35), pp. 4725-4736.
7. **Anderson, GJ, et al.** Regulation of systemic iron homeostasis: how the body responds to changes in iron demand. *Biometals*. 2007, Vol. 20, pp. 665-674.
8. **Babitt, JL, et al.** Bone Morphogenetic protein signaling by hemojuvelin regulates hepcidin expression. *Nature Genetics*. 2006, Vol. 38, pp. 531-539.
9. **Hentze, MW, Muckenthaler, MU and Andrews, NC.** Balancing Acts: Molecular Control of Mammalian Iron Metabolism. *Cell*. 2004, Vol. 117, pp. 285-297.
10. **Nicolas, G, et al.** The gene encoding the iron regulatory peptide hepcidin is regulated by anemia, hypoxia, and inflammation. *J. Clin. Invest.* 2002, Vol. 110, pp. 1037-1044.
11. **Nemeth, E, et al.** Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization. *Science*. 2004, Vol. 306, pp. 2051-2053.
12. **Roetto, A and Camaschella, C.** New insights into iron homeostasis through the study of non-HFE hereditary haemochromatosis. *Best Practice & Research Clinical Haematology*. 2005, Vols. 18, No2, pp. 235-250.
13. **Pietrangelo, A.** Hereditary Hemochromatosis - A New Look at an Old Disease. *The New England Journal of Medicine*. 2004, Vol. 350, pp. 2383-2397.
14. **Griffiths, WJH.** The genetic basis of haemochromatosis. *Aliment Pharmacol Ther.* 2007, Vol. 26, pp. 331-342.
15. **Viatte, L and Vaulont, S.** Hepcidin, the iron watcher. *Biochimie*. 2009, Vol. 91, pp. 1223-1228.
16. **Papanikolaou, G, et al.** Mutations in HFE2 cause iron overload in chromosome 1q- linked juvenile hemochromatosis. *Nature Genetics*. 2004, Vol. 36, pp. 77-82.
17. **Martinez, AR, Niemela, O and Parkkila, S.** Hepatic and extrahepatic expression of the new iron regulatory protein hemojuvelin. *Haematologica*. 2004, Vol. 89, pp. 1441-1445.
18. **Lin, L, Goldberg, P and Ganz, T.** Competitive regulation of hepcidin mRNA by soluble and cell-associated hemojuvelin. *Blood*. 2005, Vol. 106 Number 8, pp. 2884-2889.
19. **Chen, D, Zhao, M and Mundy, GR.** Bone morphogenetic proteins. *Growth Factors*. 2004, Vol. 22(4), pp. 233-241.
20. **Anderson, GJ and Frazer, DM.** Iron metabolism meets signal transduction. *Nature*. 2006, Vol. 38 Number 5, pp. 503-504.
21. **Malyszko, J.** Hemojuvelin: The Hepcidin Story Continues. *Kidney Blood Press Res*. 2009, Vol. 32, pp. 71-76.
22. **J, Krijt.** Expression of Rgmc, the murine ortholog of hemojuvelin gene, is modulated by development and inflammation, but not by iron status or erythropoietin. *Blood*. 2004, Vol. 104(13), pp. 4308-4310.
23. **House, AM and Lynch, KW.** Regulation of Alternative Splicing: More than Just the ABCs. *Journal of Biological Chemistry*. 2008, Vol. 283 No 3, pp. 1217-1221.
24. **Kim, E, Goren, A and Ast, G.** Alternative splicing: current perspectives. *BioEssays*. 2008, Vol. 30, pp. 38-47.
25. **Black, DL.** Mechanisms of Alternative Pre-messenger RNA Splicing. *Annu. Rev. Biochem.* 2003, Vol. 72, pp. 291-336.
26. **Wachtel, C and Manley, JL.** Splicing of mRNA precursors: the role of RNAs and proteins in catalysis. *Mol. BioSyst.* 2009, Vol. 5, pp. 311-316.
27. **Blencowe, BJ.** Alternative Splicing: New Insights from Global Analyses. *Cell*. Vol. 126, pp. 37-47.
28. **Blaustein, M, Pelisch, F and Srebrow, A.** Signals, pathways and splicing regulation. *International Journal of Biochemistry & Cell Biology*. 2007, Vol. 39, pp. 2031-2048.
29. **Morris, DR and Geballe, AP.** Upstream Open Reading Frames as Regulators of mRNA Translation. *Molecular and Cellular Biology*. 2000, Vols. 20, No 23, pp. 8635-8642.

30. **Kozak, M.** Pushing the limits of the scanning mechanism of initiation of translation. *Gene*. 2002, Vol. 299, pp. 1-34.
31. **Wang, XQ and Rothnagel, JA.** 5'- Untranslated regions with multiple upstream AUG codons can support low-level translation via leaky scanning and reinitiation. *Nucleic Acids Research*. 2004, Vols. 32, No. 4, pp. 1382-1391.
32. **Nagao, I and Obokata, J.** In vitro selection of translational regulatory elements. *Analytical Biochemistry*. 2006, Vol. 354, pp. 1-7.
33. **Churbanov, A, et al.** Evolutionary conservation suggests a regulatory function of AUG triplets in 5'-UTRs of eukaryotic genes. *Nucleic Acids Research*. 2005, Vols. 33, No. 17, pp. 5512-5520.
34. **Sachs, MS and Geballe, AP.** Downstream control of upstream open reading frames. *Genes Dev*. 2006, Vol. 20, pp. 915-921.
35. **Kochetov, AV.** Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. *BioEssays*. 2008, Vol. 30, pp. 683-691.
36. **Kozak, M.** An analysis of 5'- noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research*. Vols. 15, No 20, pp. 8125-8148.
37. —. Constraints on reinitiation of translation in mammals. *Nucleic Acids Research*. 2001, Vol. 29 No 24, pp. 5226-5232.
38. **Kochetov, AV, et al.** uORFs, reinitiation and alternative translation start sites in human mRNAs. *FEBS Letters*. 2008, Vol. 582, pp. 1293-1297.
39. **Arraiano, CM and Maquat, LE.** Post-transcriptional control of gene expression: effectors of mRNA decay. *Molecular Microbiology*. 2003, Vol. 49 (1), pp. 267-276.
40. **Fasken, MB and Corbett, AH.** Process or perish: quality control in mRNA biogenesis. *Nature Structural & Molecular Biology*. Vols. 12, No 6, pp. 482-488.
41. **Stalder, L and Muhlemann, O.** The meaning of nonsense. *Cell*. 2008, pp. 315-321.
42. **Chang, YF, Imam, JS and Wilkinson, MF.** The Nonsense-Mediated Decay RNA Surveillance Pathway. *Annu. Rev. Biochem.* 2007, Vol. 76, pp. 51-74.
43. **Hir, HL, et al.** The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO Journal*. 2000, Vol. 19 No. 24, pp. 6860-6869.
44. **Hir, HL, et al.** The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO Journal*. 2001, Vol. 20 No. 17, pp. 4987-4997.
45. **Wang, W, et al.** The role of Upf proteins in modulating the translation read-through of nonsense-containing transcripts. *EMBO Journal*. 2001, Vol. 20 No. 4, pp. 880-890.
46. **Lykke-Andersen, J, Shu, MD and Steitz, JA.** Human Upf Proteins Target an mRNA for Nonsense-Mediated Decay When Bound Downstream of a Termination Codon. *Cell*. 2000, Vol. 103, pp. 1121-1131.
47. **Silva, AL and Romão, L.** Controle de Qualidade do RNA Mensageiro em Mamíferos: mecanismos de degradação. [book auth.] CM Arraiano. *O Mundo do RNA*. s.l. : Lidel.
48. **Mendell, JT, et al.** Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nature Genetics*. 2004, Vol. 36 No 10, pp. 1073-1078.
49. **Stockklausner, C, et al.** The uORF-containing thrombopoietin mRNA escapes nonsense-mediated decay (NMD). *Nucleic Acids Research*. 2006, Vols. 34, No 8, pp. 2355-2363.
50. **Zhao, C, et al.** Stress-sensitive Regulation of IFRD1 mRNA Decay Is Mediated by an Upstream Open Reading Frame. *Journal of Biological Chemistry*. 2010, Vol. 285 No. 12, pp. 8552-8562.
51. **Inacio, A, et al.** Nonsense Mutations in Close Proximity to the Initiation Codon Fail to Trigger Full Nonsense-mediated mRNA Decay. *Journal of Biological Chemistry*. 2004, Vols. 279, No 31, pp. 32170-32180.
52. **Silva, AL, et al.** The canonical UPF1-dependent nonsense-mediated mRNA decay is inhibited in transcripts carrying a short open reading frame independent of sequence context. *RNA*. 2006, Vol. 12, pp. 2160-217.
53. **Nyiko, T, et al.** Plant upstream ORFs can trigger nonsense-mediated mRNA decay in a size-dependent manner. *Plant Mol Biol*. 2009, Vol. 71, pp. 367-378.

## ANEXOS

Quadro I- Sequência dos oligonucleótidos iniciadores utilizados. A vermelho está assinalada a mutação efectuada nas construções para o estudo do efeito das uORFs.

Sequência 5' - 3'	
<b>Clonagem Variante A</b>	
5'UTR Nhe I	GCT AGC CTT CTC TGG TTC CCT GAC C
Xho I mRNA AS II	CCG CTC GAG GCA TTG AGA ATG AGC ATG TCC
ex3S II	TCA CTT TCA CAC ATG CCG
ex4 AS	GAT CGA GAG AGT CGC TGA C
<b>Clonagem 5'UTR HJV no pGI2</b>	
link Hind III	CCC AAG CTT ACC GTC AAC TCA GTA GCC
luc HJV	TTT GGC GTC TTC CAT ACC TAT CCA GCC AGG
HJV luc	GGC TGG ATA GGT ATG GAA GAC GCC AAA AAC
pGL2 AS	CGT ACG TGA TGT TCA CCT C
<b>Clonagem 5'UTR HJV no PTRE-βN</b>	
5'UTR Nhe I	GCT AGC CTT CTC TGG TTC CCT GAC C
HJV aug AS Nco I	CCC TTT GGA CCG ACC TAT GGT CCA TGG
<b>Mutagénesis dirigida</b>	
mut aug uORF1 S	CAC AGA GTA GGG AAT <b>CTT</b> GGC TGG AGA ATT GGA TAG
mut aug uORF1 AS	CTA TCC AAT TCT CCA GCC <b>AG</b> ATT CCC TAC TCT GTG
mut aug uORF2 S	GAA TTG GAT AGC AGA GTA <b>TTG</b> TTT GAC CTC TGG AAA C
mut aug uORF2 AS	GTT TCC AGA GGT CAA ACA <b>ATA</b> CTC TGC TAT CCA ATT
mut main aug S	CCT GGC TGG ATA GGT <b>TTG</b> GAA GAC GCC AAA AAC
mut main aug AS	GTT TTT GGC GTC TTC <b>CAA</b> ACC TAT CCA GCC AGG
mut STOP uORF S	GGT CAA AAT TCA CTA <b>GGA</b> AGG AGG GTC ATC AGC TGG
mut STOP uORF AS	CCA GCT GAT GAC CCT CCT <b>TCC</b> TAG TGA ATT TTG ACC
<b>PCR em tempo real</b>	
Firefly S	CGC TGG AGA GCA ACT GCA T
Firefly AS	CCA GGA ACC AGG GCG TAT CT
Renilla S	GAG AAC GCC GTG ATT TTT CTG
Renilla AS	CCA CAG GTA GCT GGA GGC AG
Puromicina S	CGC AAC CTC CCC TTC TAC G
Puromicina AS	GGT GAC GGT GAA GCC GAG
βglobina S	GTG GAT CCT GAG AAC TTC AGG CT
βglobina AS	CAG CAC ACA GAC CAG CAC GT
<b>Resultados - <i>splicing</i></b>	
5'UTR S	CTT CTC TGGTTC CCT GAC C
aug S	GCT AGC ATG GGG GAG CCA GGC CAG
ex3 AS	TGT ACC GCC GAA TGG AAG G
BamHI AS	CGA TTC AAA CTG CTA ACC CTG
ex4 AS	GAT CGA GAG AGT CGC TGA C