

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL



**Desenvolvimento de um ensaio para genotipagem
do vírus da Hepatite C numa população de
infectados em ambiente prisional**

Ana Patrícia Custodinho da Avó

Mestrado em Biologia Humana e Ambiente

2011

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL



**Desenvolvimento de um ensaio para
genotipagem do vírus da Hepatite C numa
população de infectados em ambiente prisional**

Ana Patrícia Custodinho da Avó

Dissertação para obtenção do grau de Mestre orientada por
Doutora Elizabeth Pádua, Instituto Nacional de Saúde Dr. Ricardo Jorge
Prof. Doutora Ana Crespo, Faculdade de Ciências da Universidade de Lisboa

Mestrado em Biologia Humana e Ambiente

2011

AGRADECIMENTOS

Gostaria de expressar os meus sinceros agradecimentos a todas as pessoas que de alguma forma contribuíram e apoiaram a realização deste estudo, em particular:

À Doutora Elizabeth Pádua, pela oportunidade de realizar este estudo sob a sua orientação, pelos valiosos conhecimentos científicos, teóricos e práticos transmitidos e, principalmente, pela confiança demonstrada no meu trabalho, e pelo constante apoio, motivação e dedicação demonstrados ao longo deste ano.

À Professora Doutora Ana Maria Crespo por ter aceite a co-orientação deste projecto a partir da Faculdade de Ciências da Universidade de Lisboa, e pela disponibilidade e apoio demonstrados.

À Doutora Eleonora Paixão pelo auxílio na realização do tratamento estatístico e pela disponibilidade demonstrada em responder às minhas dúvidas.

À Filipa Sena e Patrícia Francisco pela preciosa ajuda sobre os programas bioinformáticos, pela partilha de conhecimentos e prontidão em ajudar.

Às colegas de laboratório, Ivone e Catarina, pela ajuda e disponibilidade demonstradas e pelos momentos de descontração.

À colega Cláudia Bispo, pela entreatuda e bons momentos passados sobretudo durante os períodos de almoço.

A todos os meus amigos pelo apoio e excelentes momentos passados juntos, sem os quais este trabalho teria sido muito mais difícil.

Ao Gonçalo, pela compreensão demonstrada em todas as minhas ausências, e pela companhia, apoio e paciência.

A toda a minha família que contribuiu para a minha formação pessoal e profissional, especialmente à minha avó Rogélia, pelo apoio e preocupação constantes.

Aos meus pais e irmão, um agradecimento especial, pelo apoio incondicional em todos os momentos, e pela compreensão, dedicação e paciência demonstradas ao longo da minha vida.

RESUMO

O vírus da hepatite C (VHC) é caracterizado por uma elevada variabilidade genética traduzida na sua classificação em 6 diferentes genótipos, e em mais de 80 subtipos com distintos padrões de distribuição epidemiológica no mundo. Os vários genótipos estão associados a uma diferente evolução da infecção e progressão da doença, sendo por isso a classificação do VHC a principal ferramenta do clínico para a determinação da dosagem e duração do tratamento dos doentes.

O presente estudo pretendeu desenvolver um método alternativo de genotipagem e subtipagem do VHC, no contexto de um laboratório de referência, que permitisse por um lado, reduzir os custos, e por outro lado, obter uma melhoria na classificação do vírus.

No presente trabalho foram analisadas amostras de referência e amostras clínicas de 72 reclusos infectados por VHC, com resultados previamente conhecidos dos níveis de RNA e de genotipagem obtidos pelo método comercial LiPA. As amostras foram sujeitas a extracção de RNA e síntese de cDNA, amplificação por *heminested* PCR e sequenciação, e ainda, a análise filogenética de sequências nucleotídicas das regiões C/E1 e NS5B do VHC. Os resultados foram comparados com os previamente conhecidos obtidos pelo método LiPA.

Independentemente das regiões genómicas, a subtipagem do VHC pelo método desenvolvido foi conseguida em 95,8% (69/72) dos casos, uma proporção bastante superior comparativamente à obtida no método LiPA (47,2%, 34/72), mostrando assim um melhor desempenho em identificar ou discriminar o subtipo VHC.

Informação molecular conjunta para as duas regiões genómicas foi obtida em 88,9% (64/72) dos casos com identificação de um recombinante intergenótipo 1b/2k e um recombinante intragenótipo 2q/2k para a região C/E1 e 2k/2a para a região NS5B. Adicionalmente, 2 casos apresentaram uma classificação de VHC discordante, 1b/4a e 3a/4a, respectivamente para as regiões C/E1 e NS5B analisadas. Contudo, nestes casos não foi possível descartar a hipótese de possíveis infecções mistas.

O elevado desempenho do método desenvolvido na detecção e classificação do VHC pode possibilitar de um modo mais preciso uma correlação clínica com os subtipos virais do VHC, e deste modo viabilizar um melhor entendimento do papel da variabilidade genómica na história natural e progressão da infecção do VHC.

PALAVRAS-CHAVE: Vírus da hepatite C (VHC); Genotipagem; Subtipagem; Regiões genómicas C/E1 e NS5B.

ABSTRACT

The hepatitis C virus (HCV) is characterized by a high genetic variability which is reflected in their classification into 6 different genotypes, and in more than 80 subtypes with distinct epidemiological patterns of distribution in the world. The several HCV genotypes are associated with a different evolution of infection and disease progression, and therefore the HCV classification is the clinician main tool to determine the patients' treatment dosage and duration.

This study aimed to develop an alternative method of HCV genotyping and subtyping, in the context of a reference laboratory, to enable lower costs, and the other hand, to obtain an improvement in virus classification.

In this work we analyzed reference samples and clinic specimens from 72 HCV-infected inmates, with previously known results of HCV viral load and genotyping obtained using the commercial method, LiPA. The samples were subjected to RNA extraction and cDNA synthesis, amplification by heminested PCR and sequencing, and also, to phylogenetic analysis of C/E1 and NS5B nucleotide sequences. The results were compared with those previously obtained by LiPA.

Regardless of the genomic regions, the HCV subtyping with the developed method was achieved for 95.8% (69/72) of cases, a much higher success rate when compared with 47.2% (69/72) in LiPA. It showed a better performance in identifying or discriminating the HCV subtype when using the method developed

The joint molecular information for the two genomic regions was obtained in 88.9% (64/72) of cases with identification of an intergenotypic recombinant form 1b/2k, and an intragenotypic recombinant 2q/2k for C/E1 region and 2k/2a for NS5B region. Additionally, 2 cases showed a discordant HCV classification, 1b/4a and 3a/4a, respectively for the analyzed regions, C/E1 and NS5B. However, in these cases it was not possible to discard the hypothesis of possible mixed infections.

The high performance of the developed method in HCV detection and classification may allow a more precise clinic correlation with the HCV subtypes, and thus to enable a better understanding of the genomic variability's role in natural history and progression of HCV infection.

KEY WORDS: Hepatitis C virus (HCV); Genotyping; Subtyping; C/E1 and NS5B genomic regions.

ÍNDICE GERAL

AGRADECIMENTOS.....	i
RESUMO	ii
<i>ABSTRACT</i>	iii
LISTA DE ABREVIATURAS E SIGLAS.....	vi
ÍNDICE DE FIGURAS.....	viii
ÍNDICE DE TABELAS	ix
1. INTRODUÇÃO	2
1.1 A infecção pelo Vírus da Hepatite C	2
1.1.1 Classificação	2
1.1.2 Morfologia do Virião	2
1.1.3 Organização Genómica	3
1.1.4 Ciclo Replicativo.....	4
1.2 Diversidade Molecular do VHC	6
1.3 Epidemiologia da Infecção pelo VHC no Mundo.....	7
1.3.1 Prevalência da Infecção.....	7
1.3.2 Distribuição dos Genótipos e Subtipos	9
1.4 Modos de Transmissão do VHC	12
1.5 Consequências da Infecção pelo VHC.....	12
1.6 Tratamento e Correlação dos Genótipos do VHC com Prognóstico da Infecção.....	13
1.7 Diagnóstico e Genotipagem do VHC	14
1.7.1 Ensaio para Pesquisa de Anticorpos	14
1.7.2 Ensaio para Pesquisa do Ácido Nucleico	15
1.7.3 Ensaio para Genotipagem	15
2. OBJECTIVOS	18
3. MATERIAL E MÉTODOS	20
3.1 Critérios de Selecção da Amostra.....	20
3.1.1 Amostras de Referência.....	20
3.1.2 Amostras Clínicas	21
3.2 Desenho do Estudo.....	23
3.2.1 Estratégia para Análise das Amostras	23

3.3	Extracção do RNA Viral.....	24
3.4	Síntese do cDNA.....	24
3.5	Amplificação por PCR das Regiões C/E1 e NS5B do VHC	25
3.5.1	Desenho e Escolha de <i>Primers</i>	25
3.5.2	Misturas de Reacção e Condições de Amplificação	26
3.5.3	Análise dos Produtos Amplificados.....	28
3.6	Purificação dos Produtos de PCR.....	29
3.7	Sequenciação dos Produtos Amplificados.....	29
3.8	Análise das Sequências Nucleotídicas	30
3.8.1	Obtenção das Sequências Consenso.....	30
3.8.2	Inferência Filogenética.....	30
3.8.3	Detecção de Recombinação Genética.....	31
3.9	Análise Estatística.....	32
4.	RESULTADOS	34
4.1	Classificação de Sequências de VHC em Amostras de Referência	34
4.2	Classificação de Sequências de VHC em Amostras Clínicas.....	34
4.2.1	Amplificação e Sequenciação das Regiões C/E1 e NS5B do VHC.....	34
4.2.2	Análise Molecular das Sequências de VHC.....	36
a)	Análise da Região C/E1	36
b)	Análise da Região NS5B.....	41
c)	Análise Conjunta das Regiões C/E1 e NS5B.....	45
4.3	Comparação de Resultados de Classificação do VHC pelo Método de Sequenciação e Análise Filogenética com o Método LiPA.....	47
a)	Comparação dos Resultados da Análise da Região C/E1 com LiPA	47
b)	Comparação dos Resultados da Análise da Região NS5B com LiPA	50
5.	DISCUSSÃO.....	54
6.	CONCLUSÕES.....	66
7.	BIBLIOGRAFIA	68
8.	ANEXOS.....	75

LISTA DE ABREVIATURAS E SIGLAS

BLAST	Do inglês, <i>Basic Local Alignment Search Tool</i>
C	Proteína do <i>core</i> , do inglês <i>core protein</i>
cDNA	Ácido desoxirribonucleico complementar, do inglês <i>complementary Deoxyribonucleic Acid</i>
CHC	Carcinoma Hepatocelular
CV	Carga Viral, RNA de VHC
DNA	Ácido desoxirribonucleico, do inglês <i>Deoxyribonucleic Acid</i>
dNTP	Desoxirribonucleico 5'-trifosfatados
E1	Proteína do invólucro 1 (“envelope”) do vírus da hepatite C
E2	Proteína do invólucro 2 (“envelope”) do vírus da hepatite C
EDTA	Etilenodiaminotetracetato
eIF2 e 3	Factor de iniciação eucariótico, do inglês <i>eukaryotic Initiation Factor</i>
EUA	Estados Unidos da América
f	Utilizado para designar o <i>primer</i> normal, do inglês <i>forward</i>
F/ARFP	Proteína do quadro de leitura alternativo, do inglês <i>Protein Alternate Reading Frame Protein</i>
HCV	Vírus da Hepatite C, do inglês <i>Hepatitis C Virus</i>
HIV	Vírus da Imunodeficiência Humana, do inglês <i>Human Immunodeficiency Virus</i>
i	Utilizado para designar o <i>primer interno</i> , do inglês <i>inner</i>
IFN-α-PEG	Interferão alfa peguilado
INSA	Instituto Nacional de Saúde
IRES	Local interno de entrada do ribossoma, do inglês <i>Internal Ribosomal Entry Site</i>
kb	Kilobase, equivalente a 1000 pb
KCl	Cloreto de Potássio
M	Molaridade
máx	Máximo
MEGA	Do inglês, <i>Molecular Evolutionary Genetics Analysis</i>
MgCl₂	Cloreto de Magnésio
min	Minuto
mín	Mínimo
mL	Mililitro
mM	Milimolar
nm	Nanómetro, equivalente a $1,0 \times 10^{-9}$ metros
NS	Não estrutural, do inglês <i>Non Structural</i>
nt	Nucleótido
o	Utilizado para designar o <i>primer externo</i> , do inglês <i>outer</i>
ORF	Grelha de leitura aberta, do inglês <i>Open Reading Frame</i>
pb	Pares de bases
PBMC	Célula mononuclear sanguínea periférica, do inglês <i>Peripheral Blood Mononuclear Cell</i>
PCR	Reacção em cadeia da polimerase, do inglês <i>Polymerase Chain Reaction</i>
pH	Potencial Hidrogeniônico de uma solução
r	Utilizado para designar o <i>primer complementar</i> , do inglês <i>reverse</i>
RBV	Ribavirina
RdRp	Polimerase de RNA dependente de RNA, do inglês <i>RNA-dependent RNA polymerase</i>

RE	Retículo Endoplasmático
RNA	Ácido ribonucleico, do inglês <i>Ribonucleic Acid</i>
RNase	Ribonuclease
RT	Transcrição reversa, do inglês <i>Reverse Transcription</i>
RT-PCR	Do inglês, <i>Reverse Transcriptase-Polymerase Chain Reaction</i>
RVS	Resposta Viroológica Sustentada
seg	Segundos
SIDA	Síndrome da Imunodeficiência Adquirida
SPSS	Do inglês, <i>Statistical Package for the Social Sciences</i>
SR-BI	Do inglês <i>Scavenger Receptor Class B type I</i>
TAE	Tampão Tris-acetato EDTA
TMA	Ensaio mediado por transcrição, do inglês <i>Transcription Mediated Assay</i>
Tris-HCl	Tris Hidrocloro
U	Unidade
UDI	Utilizadores de Drogas Injectáveis
UI/mL	Unidades Internacionais/mililitro
UKNEQAS	<i>United Kingdom National External Quality Assessment Service</i>
USA	Estados Unidos da América, do inglês <i>United States of America</i>
UTR	Região não traduzida, do inglês <i>Untranslated Region</i>
VHC	Vírus da Hepatite C
°C	Graus Celsius
µl	Microlitro, equivalente a 10 ⁻⁶ litros
µM	Micromolar, equivalente a 10 ⁻³ mol/m ³
σ	Desvio-padrão
\bar{x}	Média
κ	Coeficiente de <i>Kappa (Cohen)</i>

ÍNDICE DE FIGURAS

1. INTRODUÇÃO

Figura 1.1	Esquema representativo da estrutura do virião da hepatite C.....	2
Figura 1.2	Representação esquemática do genoma do VHC	4
Figura 1.3	Ciclo replicativo do VHC	5
Figura 1.4	Prevalência global da infecção por VHC em 2007	8
Figura 1.5	Distribuição dos principais genótipos do VHC no mundo e a sua associação a grupos com comportamento de risco para a infecção e transmissão	9
Figura 1.6	Distribuição dos genótipos de VHC entre indivíduos infectados no mundo	10

3. MATERIAL E MÉTODOS

Figura 3.1	Diagrama esquemático contendo informação das amostras de referência usadas no presente estudo	21
Figura 3.2	Diagrama esquemático contendo informação para a caracterização da amostra de conveniência usada no presente estudo	22
Figura 3.3	Diagrama correspondente à estratégia molecular planeada para o estudo de genotipagem de amostras VHC positivas	23
Figura 3.4	Esquema das regiões genômicas C/E1 e NS5B amplificadas, primers utilizados na amplificação e o seu local de hibridação na estirpe H77, e o tamanho dos fragmentos obtidos	26

4. RESULTADOS

Figura 4.1	Produtos amplificados das regiões C/E1 e NS5B visualizados em gel de agarose	35
Figura 4.2	Diagrama esquemático dos resultados de amplificação e sequenciação das amostras clínicas para as regiões genômicas C/E1 e NS5B do VHC	36
Figura 4.3	Árvore Filogenética com base na região C/E1 do VHC	38
Figura 4.4	Árvore Filogenética contendo majoritariamente sequências do genótipo 2 com base na região C/E1 do VHC	39
Figura 4.5	Análise por bootscannig da região C/E1 da sequência C4	40
Figura 4.6	Análise por bootscannig da região C/E1 da sequência C7	41
Figura 4.7	Árvore Filogenética com base na região NS5B do VHC	43
Figura 4.8	Árvore Filogenética contendo majoritariamente sequências do genótipo 2 com base na região NS5B do VHC	44
Figura 4.9	Análise por bootscannig da região NS5B da sequência C4	45
Figura 4.10	Distribuição da proporção (%) de subtipos do VHC identificados na população estudada por análise conjunta das regiões C/E1 e NS5B do VHC	44

ÍNDICE DE TABELAS

1. INTRODUÇÃO

Tabela 1.1	Genótipos e subtipos de VHC submetidos à base de dados de sequências de VHC de <i>Los Alamos</i> ; distribuição de sequências de cada genótipo (%) e número total de sequências no mundo, na Europa e em Portugal 7
Tabela 1.2	Proporção de genótipos e subtipos identificados no Laboratório Nacional de Referência IST - HIV e Hepatites B e C, nos anos de 2009 e 2010 11
Tabela 1.3	Resultados com subtipo não discriminado, incompletos e completos obtidos pelo método comercial VERSANT™ HCV Genotype 2.0 Assay-LiPA no Laboratório Nacional de Referência IST - HIV e Hepatites B e C, nos anos de 2009 e 2010 11

3. MATERIAL E MÉTODOS

Tabela 3.1	Sequência dos primers utilizados na amplificação das regiões C/E1 e NS5B do genoma do VHC, e respectiva posição nucleotídica em acordo com a sequência de referência H77 26
Tabela 3.2	Misturas de reacção correspondentes a cada enzima utilizada para amplificação das regiões C/E1 e NS5B do VHC 27
Tabela 3.3	Condições de amplificação da região C/E1 do VHC 27
Tabela 3.4	Condições de amplificação da região NS5B do VHC 28

4. RESULTADOS

Tabela 4.1	Comparação dos resultados de classificação de VHC por sequenciação e filogenia entre as regiões C/E1 e NS5B 46
Tabela 4.2	Comparação entre os resultados de genotipagem obtidos pelo método LiPA com os resultados de sequenciação e filogenia da região C/E1 do VHC 48
Tabela 4.3	Número de amostras em que a classificação do subtipo do VHC está ausente, é discordante e/ou indiscriminada pelo método LiPA em comparação com a classificação obtida por sequenciação e estimativa filogenética com base na região C/E1 48
Tabela 4.4	Comparação entre os resultados de genotipagem obtidos pelo método LiPA com os resultados de sequenciação e filogenia da região NS5B do VHC 50
Tabela 4.5	Número de amostras com indicação do subtipo do VHC ausente, discordante e/ou indiscriminado pelo método LiPA em comparação com a classificação obtida por sequenciação e estimativa filogenética com base na região NS5B 51

8. ANEXOS

Tabela 8.1	Dados demográficos e resultados laboratoriais das amostras clínicas analisadas no presente estudo 74
-------------------	--

1. INTRODUÇÃO

1. INTRODUÇÃO

1.1 A infecção pelo Vírus da Hepatite C

1.1.1 Classificação

O vírus da hepatite C (VHC) foi identificado em 1989, por Choo e seus colaboradores, no soro de um paciente com hepatite adquirida após transfusão e não causada pelos vírus da hepatite A ou da hepatite B [1].

O VHC pertence à família *Flaviviridae* do género *Hepacivirus* (do grego hepar, hepatos, *fígado*). O uso da nomenclatura internacional das variantes do VHC torna-se essencial quanto mais se descobre sobre a doença hepática relacionada com o VHC, e mais se conhece sobre as diferenças biológicas e antigénicas que existem entre as variantes virais identificadas [2].

1.1.2 Morfologia do Virião

O virião da hepatite C consiste numa partícula esférica com um diâmetro entre 40 a 70 nm, envolvida por um invólucro de constituição lipoproteica no qual estão ancoradas as glicoproteínas (E1 e E2) [3]. No interior do invólucro encontra-se a proteína do core (C) que protege o RNA genómico viral [4] (Figura 1.1).

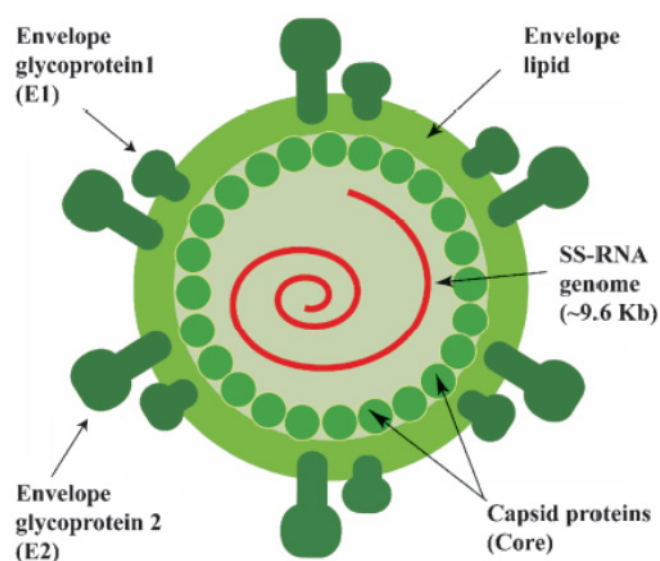


Figura 1.1 – Esquema representativo da estrutura do virião da hepatite C [5].

1.1.3 Organização Genômica

O genoma do VHC é constituído por RNA de cadeia simples com polaridade positiva, composto por aproximadamente 9,6 kb (Figura. 1.2) [5]. Contém apenas uma única grelha de leitura (*open reading frame*, ORF), precursor da poliproteína de cerca de 3000 aminoácidos, que é flanqueada por regiões de RNA não traduzidas (*untranslated region*, UTR) designadas respectivamente de 5'UTR e 3'UTR [4-6]. A região 5'UTR tem aproximadamente 341 nucleotídeos (nt) e é composta por quatro domínios altamente estruturados numerados de I a IV [4, 7]. Os domínios II, III e IV constituem um local interno de entrada do ribossoma (*Internal Ribosomal Entry Site*, IRES) que é responsável pela iniciação da tradução da poliproteína viral [8]. A região 3'UTR tem cerca de 269 nucleotídeos e, tal como a região 5'UTR, está envolvida na replicação do RNA e na tradução proteica [5, 7].

As proteínas estruturais são codificadas pelo terminal amina da ORF, enquanto a restante região da ORF codifica as proteínas não estruturais. Uma única poliproteína sofre proteólise mediado por proteases virais e da célula hospedeira, originando pelo menos 10 proteínas, incluindo três proteínas estruturais (C ou *core*, E1 e E2), uma pequena proteína, p7 que faz a ligação entre as proteínas estruturais e não estruturais e que forma canais iónicos (viroporina), e seis proteínas não estruturais (NS2, NS3, NS4A, NS4B, NS5A e NS5B) [4, 9, 10] (Figura 1.2). A proteína do quadro de leitura alternativo (*alternate reading frame protein*, *protein F/ARFP*) é codificada pelo quadro de leitura +1 do genoma viral, a qual se sobrepõe com a sequência codificante da proteína do *core* [11]. As funções, se for o caso, das proteínas F/ARFP no ciclo de vida e patogénese da infecção pelo VHC, permanecem por serem elucidadas [4].

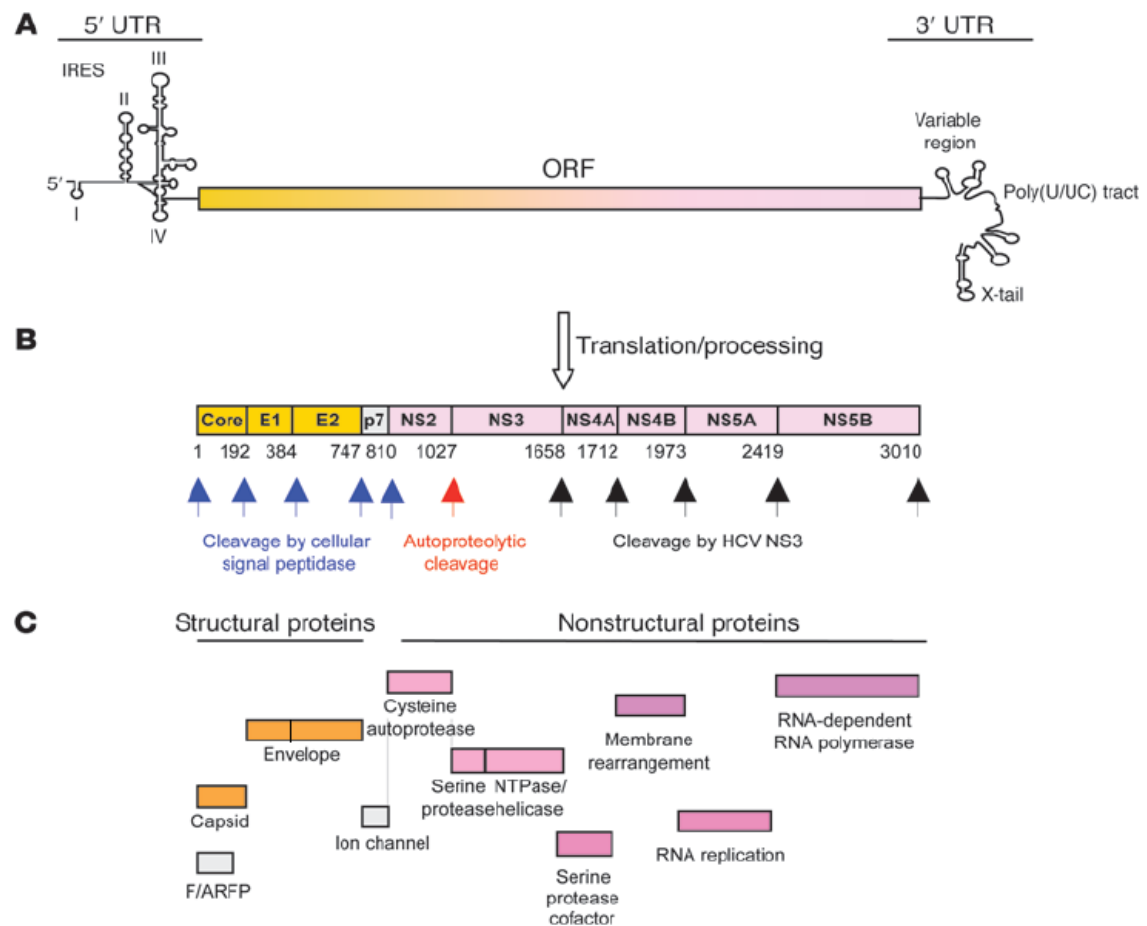


Figura 1.2 - Representação esquemática do genoma do VHC. **(A)** ORF flanqueada por 5'UTR e 3'UTR. **(B)** A poliproteína traduzida é processada por proteases celulares e virais. **(C)** Resultado da clivagem: 3 proteínas estruturais, a proteína p7 e 6 proteínas não estruturais, e as suas respectivas funções. A proteína F/ARFP é traduzida a partir de um pequeno quadro de leitura alternativo sem funções conhecidas. Adaptado de [10].

1.1.4 Ciclo Replicativo

O VHC infecta apenas humanos e chimpanzés e é considerado como sendo essencialmente hepatotrópico [4, 12]. Contudo, pode infectar células mononucleares do sangue periférico (PBMC), linfócitos B e células dendríticas. Pode ser encontrado em locais extra-hepáticos como no sistema nervoso central e na medula óssea de doentes com infecção VHC crónica, o que dificulta o sucesso do tratamento na eliminação do vírus do organismo [13-15].

O ciclo de replicação do VHC possui várias etapas. A primeira etapa corresponde à ligação das partículas virais a receptores celulares específicos. Entre outros, têm sido propostos como receptores do VHC a molécula CD81, o receptor SR-BI (*Scavenger Receptor Class B type I*) e a claudina-1 [16-18].

A internalização do VHC ocorre por endocitose mediada por clatrina que é iniciada pelas interações específicas entre as glicoproteínas virais E1 e E2 e os receptores na superfície da célula hospedeira, que provocam a libertação da nucleocápside viral no interior do citoplasma, após fusão entre as membranas que é induzida pelo pH baixo do endossoma [4, 19, 20] (Figura 1.3).

A desencapsidação da nucleocápside viral liberta cadeias genómicas positivas de RNA livres no citoplasma que juntamente com o RNA recém sintetizado, servem de RNA mensageiros para a síntese da poliproteína do VHC [8, 21]. O IRES liga-se à subunidade ribossomal 40S em associação com factores de iniciação eucarióticos 2 e 3 (eIF2 e eIF3) resultando no complexo de pré-iniciação 48S. Posteriormente, a subunidade ribossomal 60S associa-se ao complexo, levando à formação do complexo activo de tradução 80S para a síntese da poliproteína do VHC no retículo endoplasmático (RE) [22, 23]. As proteínas estruturais e a proteína p7 são libertadas da poliproteína após clivagem pelos sinais das peptidases do RE. As proteínas não estruturais são clivadas por proteases virais NS2-3 e NS3-4A [4]. A proteína F/ARFP provém de um *frameshift* ribossomal no primeiro codão da região do genoma que codifica o *core* [24].

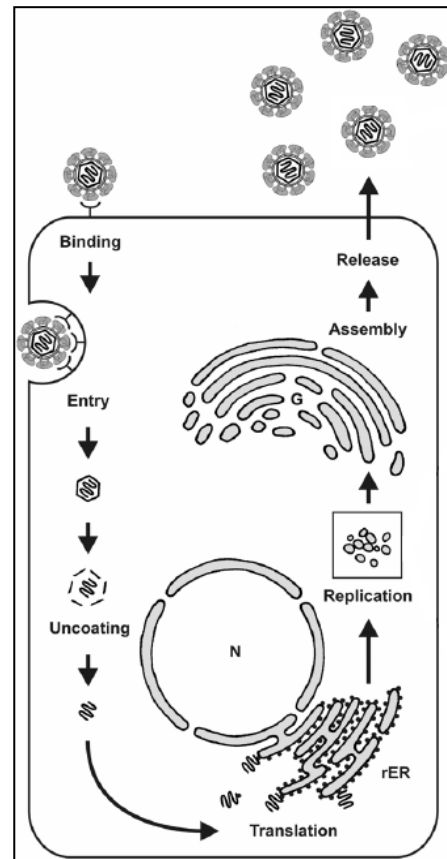


Figura 1.3 – Ciclo replicativo do VHC [3].

Por analogia com outros vírus de RNA de cadeia positiva, pensa-se que a replicação do VHC seja semi-conservativa e assimétrica. A replicação do RNA viral ocorre no citoplasma em duas fases, ambas catalisadas pela polimerase de RNA dependente de RNA (RdRp, NS5B) [9]. Além disso, vários factores celulares, tal como factores virais têm sido descritos como parte do complexo de replicação do RNA. Um factor viral importante para a formação desse complexo parece ser NS4B, que é capaz de induzir uma rede membranosa derivada do RE e/ou aparelho de Golgi. A rede é rica em colesterol e ácidos gordos que regula a replicação do VHC [25].

Na primeira fase da replicação, a cadeia positiva do RNA serve de modelo para a síntese de uma cadeia negativa intermediária de replicação. Na segunda fase, a cadeia negativa do RNA serve como um modelo para produzir inúmeras cadeias de polaridade positiva, que posteriormente serão utilizadas para a tradução da

poliproteína, síntese de novos intermediários de replicação ou empacotamento em novas partículas virais [9]. Após as proteínas virais, as glicoproteínas, e o RNA genômico do VHC terem sido sintetizados, os componentes individuais têm de ser organizados de modo a produzir viriões infecciosos [26]. A nucleocápside forma-se por oligomerização da proteína C e sua associação com o RNA genômico, processo essencial para a paragem da replicação do genoma e início da sua encapsidação. A formação do invólucro ocorre por interação da nucleocápside com as glicoproteínas E1 e E2 associadas à membrana do RE, através de gemulação. Por fim, as partículas do VHC saem da célula por exocitose [4, 27].

1.2 Diversidade Molecular do VHC

A infecção pelo VHC é um processo adaptativo e altamente dinâmico, com a produção e libertação de cerca de 10^{12} viriões por dia [28]. Esta elevada taxa de replicação em conjunto com os erros cometidos pela polimerase de RNA do VHC (ausência de actividade de *proofreading*), são responsáveis pela grande diversidade genética das estirpes de VHC [29, 30]. A evolução é rápida e o grau de diversificação das estirpes de VHC no indivíduo aumenta significativamente ao longo do tempo, resultando no desenvolvimento de *quasispecies* que são importantes na história natural da infecção do VHC, por possibilitarem o escape do vírus ao sistema imunitário do hospedeiro [29].

Comparações de sequências do VHC derivadas de indivíduos de diferentes regiões geográficas revelaram a presença de seis genótipos de VHC, que em acordo com a alteração na classificação ocorrida em 1998, foram designados por algarismos de 1 a 6 [31]. Estes genótipos diferem em 30-35% na sua sequência genómica [32]. As regiões correspondentes às glicoproteínas E1 e E2 são as que apresentam maior variabilidade, enquanto as regiões do *core* e de alguns genes das proteínas não estruturais, tais como NS3, são mais conservadas. A região 5'UTR é a que possui menor variabilidade entre genótipos [33-35]. Mais recentemente foram identificadas sequências de VHC divergentes e consideradas como potencialmente pertencerem a um novo genótipo, que foi designado de genótipo 7 [36, 37]. Com excepção dos genótipos 5 e 7, todos os restantes genótipos do VHC possuem diversos subtipos relacionados mas geneticamente distintos entre si, com as sequências a diferirem em 20-25% [38]. Mundialmente, estão relatados mais de 80 subtipos diferentes de VHC, tendo sido designados por letras minúsculas associadas ao algarismo do respectivo genótipo [35, 36, 39-41] (Tabela 1.1).

A recombinação também contribui para a diversidade genética do VHC. Nos vírus de RNA de polaridade positiva como o VHC, a recombinação ocorre durante a replicação do RNA genómico quando existe co-infecção com duas variantes virais. Embora a recombinação seja pouco frequente, já foram identificados alguns recombinantes intergenótipos naturais, tais como 2k/1b [42, 43], 2b/1b [44], 2/5 [45] e 2i/6p [46], e também intragenótipos designados como 1a/1b [47] e 1a/1c [48]; ainda de referir, que foi também descrita a recombinação de intrasubtipos de VHC nos subtipos 1a, 1b e 3a [49, 50].

A limitação na detecção de recombinação entre os genótipos do VHC também tem implicações para a classificação do vírus. Isto porque na maioria dos estudos de variabilidade do VHC em áreas com elevada diversidade são baseadas na análise de uma única região sub-genómica, o que pode potencialmente limitar a detecção de formas geneticamente recombinadas, subestimando a verdadeira frequência de recombinação [35].

Tabela 1.1 – Genótipos e subtipos de VHC submetidos à base de dados de sequências de VHC de *Los Alamos*; distribuição de sequências de cada genótipo (%) e número total de sequências no mundo, na Europa e em Portugal [36, 39-41]. (Acedido a 15/08/2011).

Genótipo	Mundo	Europa	Portugal	Subtipos
1	76,7%	82,2%	50,2%	a, b, c, d, e, f, g, h, i, j, k, l, m
2	4%	1,9%	-	a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r
3	10,2%	9,4%	21%	a, b, c, d, e, f, g, h, i, k, l
4	2,5%	2,1%	28,1%	a, b, c, d, e, f, g, h, k, l, m, n, o, p, q, r, s, t, u
5	1,4%	2,1%	-	a
6	1,2%	-	-	a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u
7	-	-	-	a
Outro	3,8%	2,3%	0,8%	
Total de sequências	80215	44468	257	

1.3 Epidemiologia da Infecção pelo VHC no Mundo

1.3.1 Prevalência da Infecção

A infecção por VHC tem um elevado impacto na saúde pública da maioria dos países do mundo, estimando-se que 130 a 170 milhões de pessoas, ou seja, cerca de 3% da população mundial, se encontre infectada pelo vírus [51]. Contudo, conhecer a incidência da infecção de VHC é mais difícil, pois a maioria das infecções são assintomáticas e os ensaios de diagnóstico mais correntemente usados não

distinguem entre a infecção aguda, crónica ou resolvida [52]. Para além disso, a maioria dos estudos de prevalência usam populações seleccionadas, nomeadamente dadores de sangue, o que pode subestimar a prevalência real do vírus na população [53].

Embora o VHC seja endémico em todo o mundo, há um elevado grau de variabilidade na sua distribuição geográfica, quer entre países do mesmo continente, quer entre zonas geográficas do mesmo país [54]. As regiões com menores taxas de prevalência incluem os países industrializados da América do Norte, norte e oeste da Europa, e Austrália (1-2,5%). Os países com maior prevalência localizam-se em África, Ásia e América do Sul (>2,5%), sendo o Egipto, o país com maior prevalência de infecção por VHC (>15%) [51, 54, 55] (Figura 1.4).

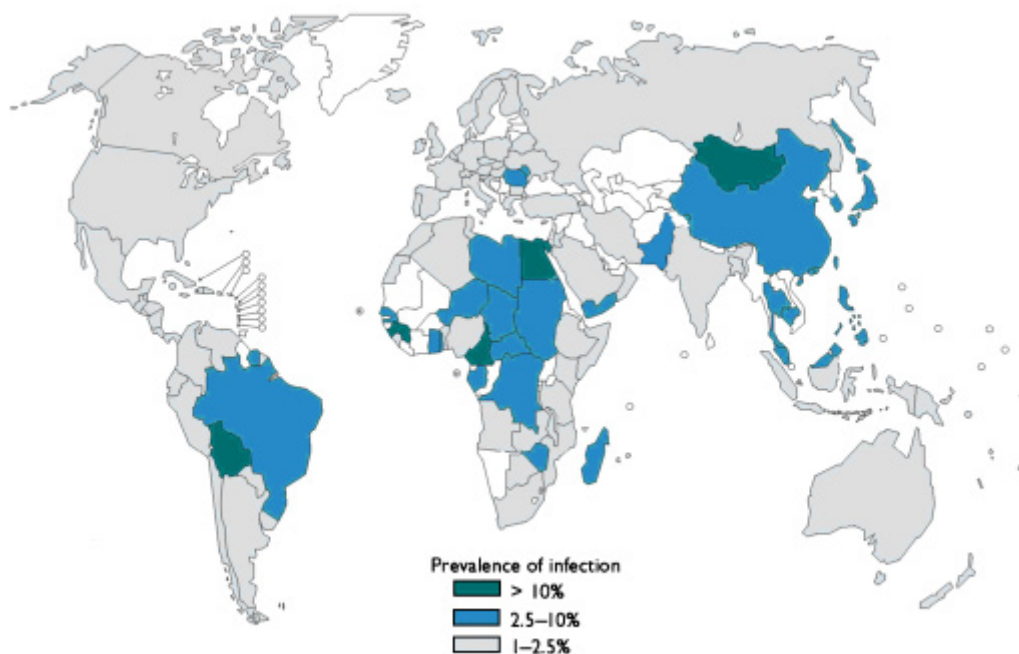


Figura 1.4 – Prevalência global da infecção por VHC em 2007 [55].

Em Portugal, existem poucos estudos sobre a infecção pelo VHC estimando-se uma prevalência de 1,5% na população portuguesa (cerca de 150 mil indivíduos), sendo o VHC responsável por 45% dos casos notificados de hepatite, embora nem todos os casos estejam notificados e muitos dos indivíduos infectados não estejam diagnosticados [56]. De acordo com dados da Direcção-Geral de Saúde, existiam 85 novos casos notificados de hepatite C em 2009, mais 39 casos do que no ano anterior [57]. Também, segundo um estudo do *Observatório Europeu da Droga e da Toxicoddependência*, Portugal é um dos países europeus a apresentar as maiores taxas de infecção pelo VHC, podendo atingir valores entre 60-80% em populações de toxicoddependentes [58].

1.3.2 Distribuição dos Genótipos e Subtipos

A distribuição dos genótipos e subtipos do VHC reflecte-se na epidemiologia do vírus e está fortemente associada com as principais vias de transmissão e disseminação da infecção [59]. Os subtipos 1a, 1b e 3a, encontram-se amplamente distribuídos na população. Este facto, resulta, quer da transmissão do vírus por transfusão de sangue (subtipo 1b), quer por troca de agulhas contaminadas entre UDI (utilizadores de drogas injectáveis) (subtipos 1a e 3a) nos últimos 30 a 70 anos, e que são transmissões que presentemente representam a grande maioria das infecções identificadas na Europa Ocidental [33] (Figura 1.5).

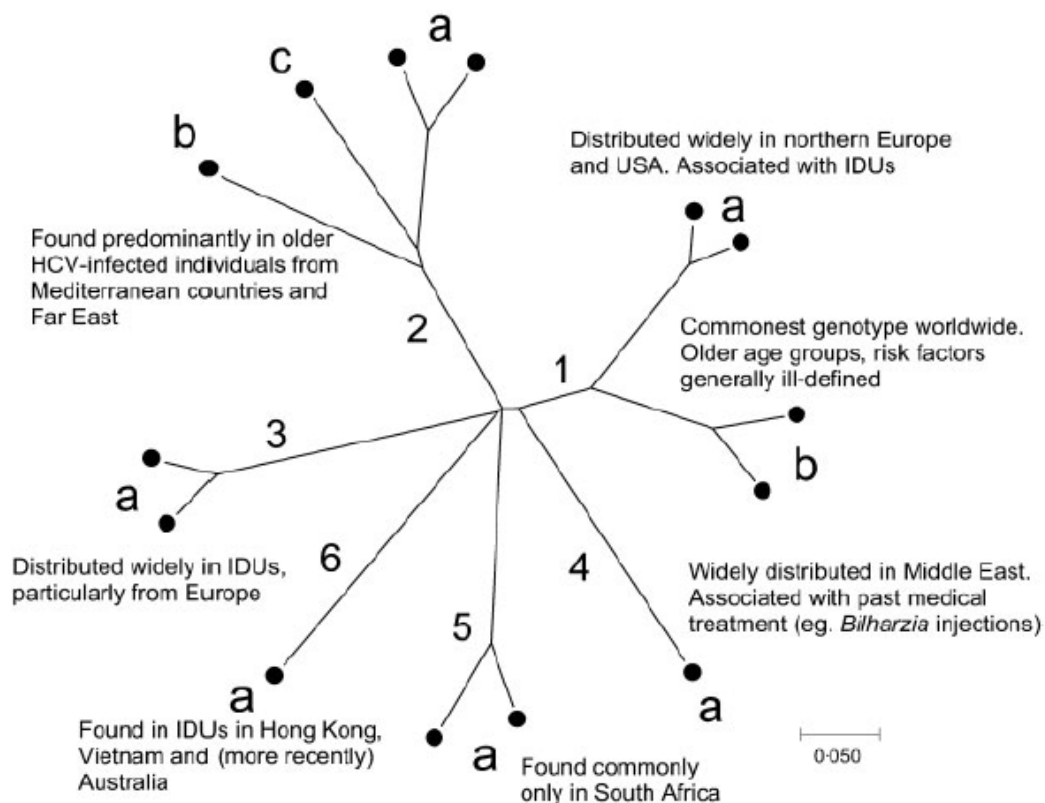


Figura 1.5 – Distribuição dos principais genótipos do VHC no mundo e a sua associação a grupos com comportamento de risco para a infecção e transmissão do vírus [33].

O genótipo 1, seguido pelos genótipos 2 e 3, frequentemente detectados na América e também na Europa, tornaram-se o foco de interesse da maioria dos ensaios clínicos para desenvolvimento de fármacos [60] (Figura 1.6). Enquanto o subtipo 1a está relacionado com UDI e encontra-se globalmente distribuído pelo norte da Europa e EUA, o subtipo 1b está associado a indivíduos de idade mais avançada e a transfusões sanguíneas realizadas no passado. Os subtipos 2a, 2b e 2c são

encontrados predominantemente nos países mediterrânicos e do Extremo Oriente. O subtipo 3a afecta predominantemente os UDI na Europa [33] (Figura 1.5).

O genótipo 4 é a causa mais frequente de hepatite C crónica no Médio Oriente, no norte de África e na África Subsaariana. Recentemente, observou-se a disseminação para o sul da Europa, dos subtipos 4a e 4d, entre UDI e populações de imigrantes [61, 62]. O genótipo 5 tem sido quase exclusivamente descrito na África do Sul [63]. O genótipo 6 é raro e está confinado ao sudeste da Ásia [64, 65] (Figura 1.6).

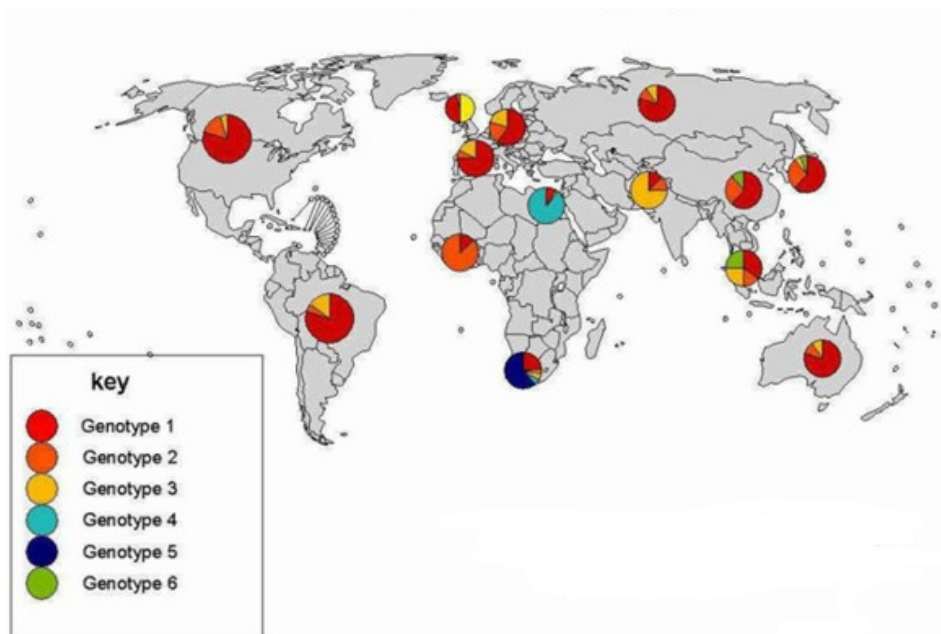


Figura 1.6 – Distribuição dos genótipos de VHC entre indivíduos infectados no mundo [65].

Em Portugal existem poucos estudos sobre a prevalência dos genótipos do VHC. Um estudo realizado entre 1995 e 2004, com base na genotipagem por um método comercial, indicou o genótipo 1 (subtipos 1a e 1b) como sendo o mais prevalente (60,8%), seguido dos genótipos 3 (subtipo 3a) (30,6%), 4 (7,6%) e 2 (1%) [66]. Mais recentemente, outro estudo (Dezembro de 2009 a Abril de 2011) apresentou uma distribuição de genótipos de 57,3%, 25,5% e 15,4%, respectivamente para os genótipos 1, 3 e 4, através de genotipagem por um método comercial e sequenciação [67]. Os UDI são o grupo particularmente mais afectado pela infecção por VHC no nosso país (58,7%) [66].

Segundo dados do Laboratório Nacional de Referência de IST - HIV e Hepatites B e C, do Instituto Nacional de Saúde Dr. Ricardo Jorge, foram testados em 2009 e em 2010, para genotipagem do VHC, respectivamente 226 e 208 soros de

indivíduos em ambiente prisional. Nesta população infectada foram identificados os 4 primeiros genótipos de VHC. E em ambos os anos, o genótipo mais frequente foi o genótipo 1 e o menos frequente foi o genótipo 2. Em termos de classificação do genótipo e do subtipo viral, os mais comuns foram 1a e 3a (Tabela 1.2). Embora a maioria dos resultados obtidos tenham sido designados de completos, ou seja, resultados em que a classificação do vírus identifica simultaneamente o genótipo e o subtipo do VHC em causa, registou-se para estes dois anos 59 (13,6%; 59/434) casos com resultados não discriminados relativamente ao subtipo do VHC. Observou-se ainda 25 (5,8%; 25/434) casos de resultados incompletos por não ter sido possível determinar o subtipo viral em causa (Tabela 1.3).

Tabela 1.2 – Proporção de genótipos e subtipos identificados no Laboratório Nacional de Referência IST - HIV e Hepatites B e C, nos anos de 2009 e 2010, utilizando o método comercial *VERSANT™ HCV Genotype 2.0 Assay-LiPA* (Siemens Healthcare Diagnostics Inc., Tarrytown, NY, USA).

Resultados	2009		2010	
	Nº de casos	%	Nº de casos	%
1	10	4,4	1	0,5
1a	91	40,3	102	49,0
1a/1b	1	0,4	2	1,0
1b	32	14,2	18	8,7
2	2	0,9	1	0,5
2a/2c	2	0,9	3	1,4
2b	-	-	1	0,5
3	1	0,4	1	0,5
3a	61	27,0	45	21,6
4	6	2,7	3	1,4
4a/4c/4d	20	8,8	31	14,9
Total	226	-	208	-

Tabela 1.3 – Resultados com subtipo não discriminado (vários subtipos), incompletos (sem indicação do subtipo) e completos (com classificação do genótipo e subtipo do VHC) obtidos pelo método comercial *VERSANT™ HCV Genotype 2.0 Assay-LiPA* (Siemens Healthcare Diagnostics Inc., Tarrytown, NY, USA) no Laboratório Nacional de Referência IST - HIV e Hepatites B e C, nos anos de 2009 e 2010.

Resultados não discriminados		Resultados incompletos				Resultados completos					
2009		2010		2009		2010		2009		2010	
Nº de casos	%	Nº de casos	%	Nº de casos	%	Nº de casos	%	Nº de casos	%	Nº de casos	%
23	10,2	36	17,3	19	8,4	6	2,9	184	81,4	166	79,8

1.4 Modos de Transmissão do VHC

Na maioria dos países desenvolvidos, a fonte predominante de novas infecções pelo VHC tem sido a partilha de agulhas e material contaminado no grupo de UDI [68]. Nesta população, a seroprevalência de VHC pode atingir 70% [69]. O rastreio de anticorpos anti-VHC nos bancos de sangue, desde o início de 1990, praticamente eliminou a transmissão do vírus pela via sanguínea, mas o comportamento de risco associado ao consumo de drogas injectáveis tornou-se actualmente o principal factor de risco para a aquisição e transmissão do VHC [70, 71]. Contrariamente, nos países em desenvolvimento, com escassos recursos económicos, as transfusões e as injeções terapêuticas inseguras ainda podem ser as principais vias de transmissão do vírus [68].

Como consequência das medidas de segurança implementadas em bancos de sangue nos países desenvolvidos, observou-se uma alteração da predominância dos subtipos de VHC de 1b para 1a e 3a, sendo a infecção pelo VHC em crianças exclusivamente associada à transmissão vertical do vírus [62, 72]

Outras vias de transmissão do VHC incluem, o contacto sexual, com uma eficiência de transmissão do VHC de 0,01% em casais heterossexuais monogâmicos, e também a lesão não intencional com agulhas ou exposição a outros objectos pontiagudos contaminados em profissionais de saúde, com uma incidência descrita que pode atingir os 10% [73, 74].

1.5 Consequências da Infecção pelo VHC

A infecção pelo vírus da hepatite C num indivíduo pode ser classificada como aguda ou crónica. A forma aguda pode apresentar sinais clínicos de uma hepatite até 6 meses após a exposição ao VHC e, ocorre em cerca de 25% dos infectados. Nos restantes indivíduos é assintomática, manifestando apenas sinais ligeiros no curso da infecção [51, 75]. Com a resolução clínica dos sintomas, os níveis de aminotransferase normalizam em cerca de 40% dos pacientes. A perda de RNA do VHC, indica cura da hepatite C, mas ocorre em menos de 20% dos infectados, sendo independente da normalização das aminotransferases [75].

O risco de infecção crónica de VHC é elevado, e cerca de 80% dos pacientes mantêm níveis de RNA de VHC quantificáveis seis meses após a infecção aguda [53, 75]. O mecanismo que leva ao estabelecimento da infecção crónica ainda não é conhecido. O processo adaptativo e a evolução rápida do vírus podem permitir que o VHC escape ao sistema imunológico do hospedeiro infectado [75]. A infecção pelo

VHC é geralmente lenta e progressiva e os sintomas de infecção crónica variam. Aproximadamente 10-20% dos indivíduos têm elevado risco de desenvolver cirrose num período de 10 a 30 anos, e 1-5% dos doentes podem desenvolver carcinoma hepatocelular (CHC) [75, 76].

É estimado que a infecção crónica de VHC seja responsável por aproximadamente 250 mil a 350 mil mortes por ano, principalmente relacionadas com cirrose, doença hepática e CHC de fase terminal [77].

Dados da Direcção Geral de Saúde, referentes ao período de 2001 a 2005, fazem prever que o número de mortes por doença hepática crónica associada ao VHC duplicará no país dentro de 10 a 20 anos [referido em 66, 78] Em Portugal, a doença hepática é a oitava causa de morte, sendo a hepatite C responsável por 500 mortes por ano. Cerca de 60% das mortes por CHC e 20% por SIDA (Síndrome da Imunodeficiência Adquirida) também estão relacionadas com o VHC [79].

1.6 Tratamento e Correlação dos Genótipos do VHC com Prognóstico da Infecção

Desde a introdução da terapia combinada de interferão-alfa peguilado (INF- α -PEG) com ribavirina (RBV), padrão actual de tratamento, a eficácia de alcançar a resposta virológica sustentada (RVS), ou seja, um resultado de RNA de VHC indetectável 24 semanas após o fim do tratamento, tem melhorado significativamente [80, 81]. Estes fármacos são conhecidos por activar diversos mecanismos antivirais directos e indirectos, tais como por exemplo a degradação do RNA viral e a paragem da tradução viral [82, 83]. Infelizmente, estes dois compostos são tóxicos e a sua administração causa efeitos colaterais, entre os quais, cefaleia, febre, depressão grave, mialgia, artralgia e anemia hemolítica. A interrupção precoce pode ocorrer em 20% dos pacientes [84]. O custo do tratamento é muito elevado e pode prolongar-se por vários meses (até 72 semanas). Em alguns casos a RVS não é alcançada no final do tratamento [85].

A combinação de INF- α -PEG e RBV induz uma RVS que varia entre 42-82% dos pacientes, dependente do genótipo em causa [81, 84]. Em infecções pelo genótipo 1 observam-se as taxas mais baixas de RVS (40-50%), mesmo após 48 semanas de tratamento. Em contraste, para os genótipos 2 ou 3 foram obtidas elevadas taxas de RVS que atingiam os 80% [81, 84, 86]. No entanto, taxas de RVS de valores intermédios (43-70%) após 36-48 semanas de tratamento foram associadas ao

genótipo 4. Em relação aos genótipos 5 e 6, as taxas de RVS alcançadas podem ser as mais elevadas, variando entre 60-85% após 48 semanas de tratamento [87].

Recentemente, dois estudos revelaram a existência de uma eficácia de tratamento superior, através da utilização de um novo fármaco, o telaprevir, que interfere na replicação do VHC pela inibição da enzima protease serínica NS3/4A. Dois ensaios clínicos de fase 3 no tratamento da hepatite C crónica associada ao genótipo 1 demonstraram que o tratamento com telaprevir, em combinação com INF- α -PEG e RBV, conduz a uma RVS significativamente superior à do tratamento usual com apenas INF- α -PEG e RBV [88].

1.7 Diagnóstico e Genotipagem do VHC

Por apresentar sintomas inespecíficos, leves ou mesmo inexistentes, a infecção por VHC não é frequentemente diagnosticada. Estima-se que apenas um em cada quatro indivíduos infectados reconheça a sua situação e, portanto, a maioria dos infectados não segue as opções de tratamento e não toma medidas individuais de prevenção para diminuir o risco da transmissão do vírus [86]. No diagnóstico da infecção existem comercialmente disponíveis ensaios serológicos para pesquisa de anticorpos anti-VHC, ensaios para a pesquisa de RNA e para a genotipagem do VHC [89].

1.7.1 Ensaios para Pesquisa de Anticorpos

Na prática laboratorial corrente, os ensaios serológicos baseiam-se em testes imunoenzimáticos que utilizam antigénios recombinantes da proteína C e NS3-5 na captura de anticorpos específicos. Estes ensaios detectam com uma sensibilidade superior a 99% a presença de anticorpos anti-VHC, em amostras (soro ou plasma) colhidas 4 a 10 semanas após a infecção [3, 89, 90].

Contudo, resultados serológicos positivos exigem a quantificação de RNA do VHC para discriminação entre pacientes com hepatite C crónica ou infecção pelo VHC resolvida. Numa infecção recente por VHC (<10 semanas) os testes serológicos também podem conduzir a resultados falsos negativos. Em contraste, a pesquisa de RNA do VHC é detectável alguns dias após a infecção, sendo preferencialmente o ensaio escolhido para o diagnóstico de hepatite C aguda [91].

1.7.2 Ensaio para Pesquisa do Ácido Nucleico

A determinação de RNA do VHC pode ser essencial para diagnosticar uma infecção aguda e monitorizar a eficácia do tratamento de indivíduos infectados [91]. Existem ensaios qualitativos e quantitativos para a detecção de RNA do VHC. Nos ensaios qualitativos comercializados incluem-se o RT-PCR (*Reverse Transcriptase-Polymerase Chain Reaction*) e TMA (*Transcription Mediated Assay*) do RNA do VHC. Os ensaios quantitativos possibilitam a determinação do número de cópias virais num volume específico de amostra (soro ou plasma) através da amplificação da sequência alvo (PCR competitivo e PCR em tempo real) ou da amplificação de sinal (*branched DNA*) [92].

Os ensaios qualitativos têm normalmente limites mais baixos de detecção de cópias virais em comparação com os ensaios quantitativos. Os resultados da aplicação destes testes devem ser conhecidos antes da realização de um teste de genotipagem do VHC [92].

1.7.3 Ensaio para Genotipagem

Uma vez que os vários genótipos de VHC estão associados a uma diferente evolução da infecção e doença, a genotipagem do vírus torna-se a principal ferramenta que o clínico dispõe para a determinação da dosagem e duração do tratamento, e prognóstico da infecção [37].

Tanto a análise directa de sequências como a tecnologia de hibridação reversa permite a genotipagem do VHC. Os ensaios iniciais foram desenvolvidos para analisar exclusivamente a região 5'UTR. A escolha desta região deve-se ao seu elevado grau de conservação, sendo frequentemente utilizada nos ensaios validados para a prestação de serviços efectuada nos laboratórios em pedidos de genotipagem do VHC [37, 93]. No entanto, pode suportar erros ou incapacidade de classificação, pois apesar de ser bastante conservada, esta região possui vários polimorfismos que não permitem, por vezes, a classificação correcta do subtipo em causa. Por exemplo, o polimorfismo da posição 243, frequentemente utilizada para diferenciar os subtipos 1a e 1b, é incerta. Um estudo demonstrou que 6 (7,5%) das 80 sequências com o subtipo 1a e 1 (1,3%) de 79 sequências com o subtipo 1b foram incorrectamente classificadas tendo como base esse polimorfismo [35, 94]. Também a diferenciação entre os subtipos 1a e 1b pode ser prejudicada pelo facto do fragmento analisado diferir apenas num nucleótido na posição 99 [95]. A análise da região 5'UTR também se tem demonstrado limitada na discriminação entre os genótipos 6 e 1, devido à semelhança da região 5'UTR destes dois genótipos [96].

O método laboratorial correntemente utilizado para genotipagem do VHC é uma segunda versão do ensaio designado *VERSANT™ HCV Genotype 2.0 Assay-LiPA* (Siemens Healthcare Diagnostics Inc., Tarrytown, NY, USA), o qual usa informação da região 5'UTR e do *core* do VHC, permitindo uma melhor e mais precisa diferenciação entre o genótipo 1 e os subtipos 6c a 6l, e entre os subtipos 1a e 1b. Este método é um ensaio de sondas em linha, e destina-se a amostras com resultado positivo para RNA do VHC. O ensaio tem como princípio a hibridação reversa, em que o produto de DNA biotilado, resultante da amplificação das regiões 5'UTR e do *core* do VHC por RT-PCR, é hibridado com sondas de oligonucleótidos específicas dessas regiões para os seis genótipos do VHC. No final, a formação de um precipitado de cor púrpura/castanha produz um padrão de bandas visível numa tira de nitrocelulose que será posteriormente comparada com as linhas da tira padrão. Este ensaio de genotipagem apresenta como desvantagem o seu elevado custo [97].

2. OBJETIVOS

2. OBJECTIVOS

A caracterização molecular da infecção pelo vírus da Hepatite C com identificação dos genótipos e subtipos virais na população infectada é essencial não só para a área clínica, implementação de procedimentos terapêuticos e a produção de vacinas eficazes, como ainda contribui para aperfeiçoar e desenvolver testes de diagnóstico. Os ensaios de genotipagem do VHC mais vulgarmente disponíveis e utilizados no laboratório, usam exclusivamente a análise da regiões 5'UTR e do *core* do genoma viral, e possuem um elevado custo, o que impede o seu uso em estudos populacionais de larga escala.

Atendendo a estes factores, os principais objectivos deste trabalho foram:

1. Desenvolver e avaliar uma nova metodologia *in house* de baixo custo, tendo como alvo uma população de indivíduos infectados por VHC que permitisse identificar e classificar os vários genótipos e subtipos virais;
2. Amplificar fragmentos virais de diferentes regiões do genoma a partir do soro de indivíduos infectados por VHC, através de extracção do RNA seguido de um RT-PCR e sequenciação, com posterior análise e classificação das sequências recorrendo a uma base de dados de VHC e estimativa filogenética;
3. Analisar diferentes regiões genómicas de forma a não limitar a possibilidade de detectar a presença de eventuais formas recombinantes de VHC;
4. Comparar os resultados obtidos pelo método a implementar com os resultados já conhecidos pelo método correntemente utilizado no laboratório (*VERSANTTM HCV Genotype 2.0 Assay-LiPA (Siemens Healthcare Diagnostics Inc., Tarrytown, NY, USA)*) para a genotipagem do VHC;
5. Estabelecer possíveis associações entre algumas características demográficas (idade) e laboratoriais (valores de RNA de VHC) com os resultados obtidos de genotipagem e subtipagem;
6. Reduzir custos para viabilizar futuros estudos que conduzam a um melhor conhecimento molecular da infecção do VHC no país.

3. MATERIAL E MÉTODOS

3. MATERIAL E MÉTODOS

3.1 Critérios de Selecção da Amostra

Os soros foram seleccionados de uma população infectada por VHC (presença de anticorpos anti-VHC), com níveis de RNA de VHC detectáveis (limite de detecção de 50 UI/mL) e resultados de genotipagem pelo método *VERSANT™ HCV Genotype 2.0 Assay-LiPA* (*Siemens Healthcare Diagnostics Inc., Tarrytown, NY, USA*) (designado genericamente por método LiPA) conhecidos. Os critérios privilegiaram a escolha de soros com resultados de genotipagem incompletos (sem indicação do subtipo de VHC) ou indiscriminados (com indicação de vários subtipos virais).

3.1.1 Amostras de Referência

Neste estudo foram analisados 4 soros de referência, obtidos a partir do programa de controlo de qualidade externa de genotipagem do VHC designado de *UKNEQAS* (*United Kingdom National External Quality Assessment Service*), integrados em dois painéis distribuídos em 2009 e 2010 por diferentes laboratórios Europeus. Estes soros de referência com um número de cópias de RNA de VHC conhecido, estavam classificados quanto ao genótipo e subtipo do VHC. No âmbito do controlo de qualidade foram submetidos na altura ao método LiPA (Figura 3.1). A utilização das amostras de referência no presente estudo teve como objectivo uma primeira avaliação do desempenho da metodologia a implementar.

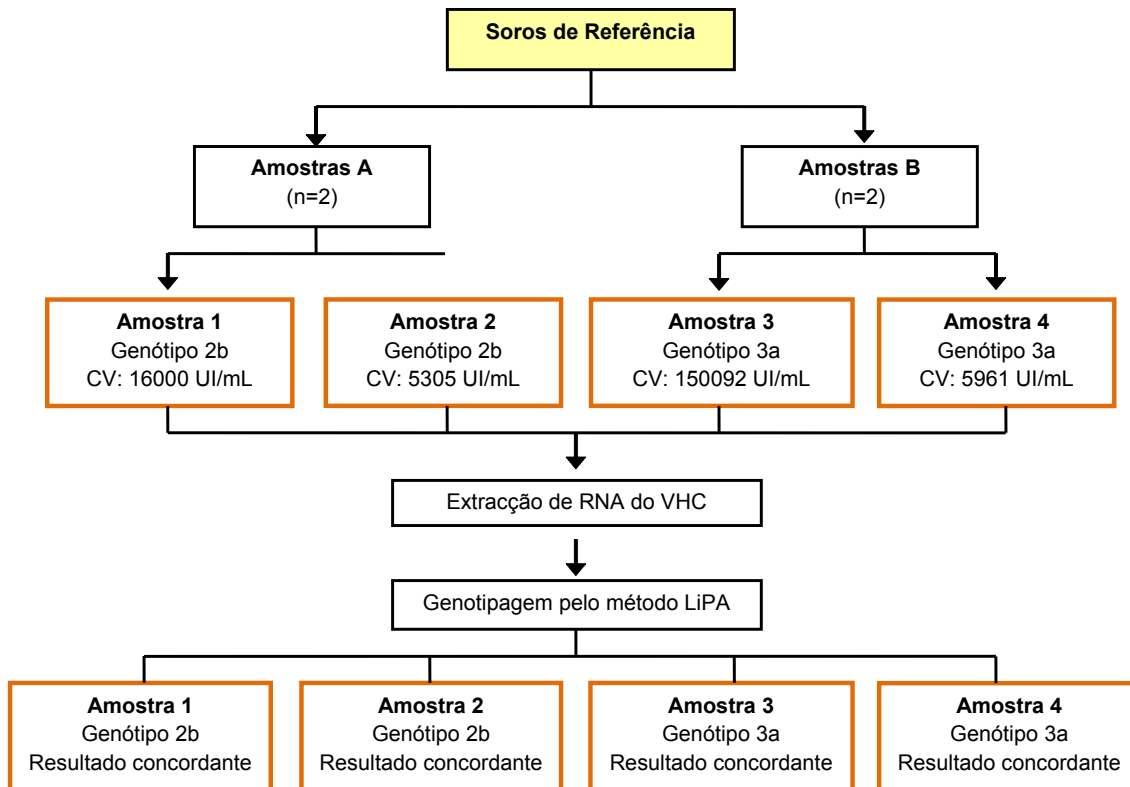


Figura 3.1 – Diagrama esquemático contendo informação das amostras de referência usadas no presente estudo com a apresentação de resultados obtidos após a sua submissão ao ensaio de genotipagem do VHC pelo método LiPA. (CV = carga viral, RNA de VHC).

3.1.2 Amostras Clínicas

No presente estudo foi analisada uma amostra de conveniência constituída por 72 soros de reclusos infectados por VHC de vários Estabelecimentos Prisionais do Distrito de Lisboa. A maioria dos soros foi colhida em 2008 para determinação da carga vírica e genotipagem do VHC.

Atendendo a que a maioria dos estabelecimentos prisionais albergava homens, apenas 6 casos estudados eram do sexo feminino (8,3%) sendo os restantes 66 do sexo masculino (91,7%). As idades dos indivíduos estavam compreendidas entre os 22 e os 64 anos (idade média de 39,5 anos; $\sigma=8,8$) (Anexos, Tabela 8.1).

A população de amostras estudada possuía níveis detectáveis de RNA de VHC que variava entre 470 a 50 200 000 UI/mL quantificados pelo sistema *COBAS AmpliCor® HCV Monitor v2.0* (Roche, Roche Diagnostics, Branchburg, NJ, USA). A posterior genotipagem do vírus com a utilização do kit comercial LiPA identificou a presença dos genótipos 1 (n=31, 43%), 2 (n=12; 16,7%), 3 (n=10; 13,9%) e 4 (n=19; 26,4%), sendo as formas virais mais prevalentes no grupo populacional estudado

classificadas de 1a (n=12; 16,7%), de 4a/4c/4d (n=11; 15,3%) e de 1b (n=10; 13,9%) (Figura 3.2).

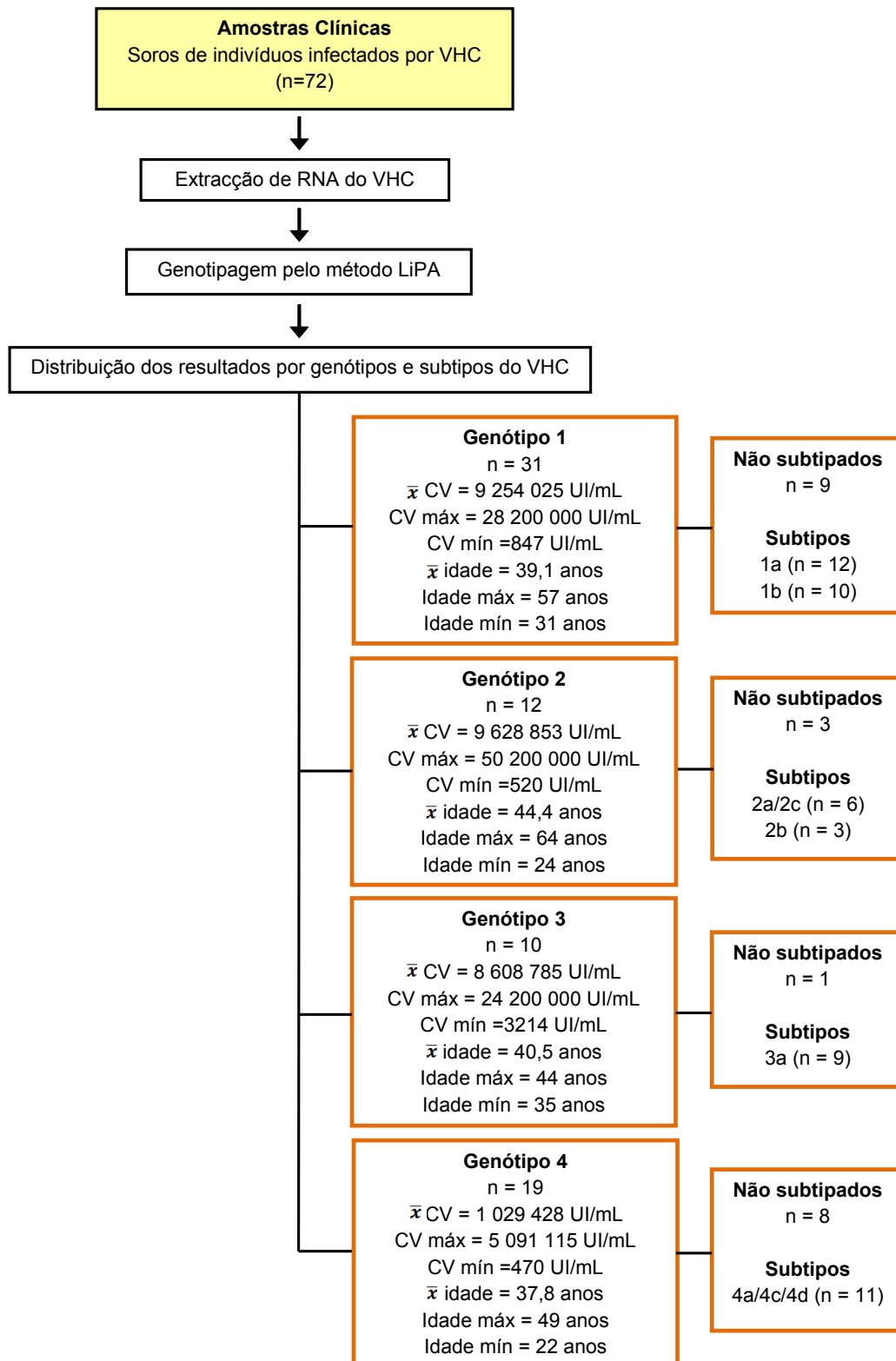


Figura 3.2 - Diagrama esquemático contendo informação para a caracterização da amostra de conveniência usada no presente estudo com apresentação dos resultados obtidos pelo método LiPA, valores de RNA de VHC (CV, carga viral) e valores de idade para o total das amostras clínicas a submeter ao método a desenvolver. \bar{x} CV, média da carga viral; CV máx, valor máximo de carga viral; CV mín, valor mínimo de carga viral. \bar{x} idade, média da idade; Idade máx, valor máximo da idade; Idade mín, valor mínimo da idade.

3.2 Desenho do Estudo

3.2.1 Estratégia para Análise das Amostras

A abordagem metodológica adoptada para a análise das amostras de referência e clínicas baseou-se numa estratégia molecular para classificação das formas virais. Os resultados da análise molecular das regiões C/E1 e NS5B foram comparados entre si e posteriormente comparados com os resultados previamente conhecidos obtidos pelo método comercial LiPA (Figura 3.3).

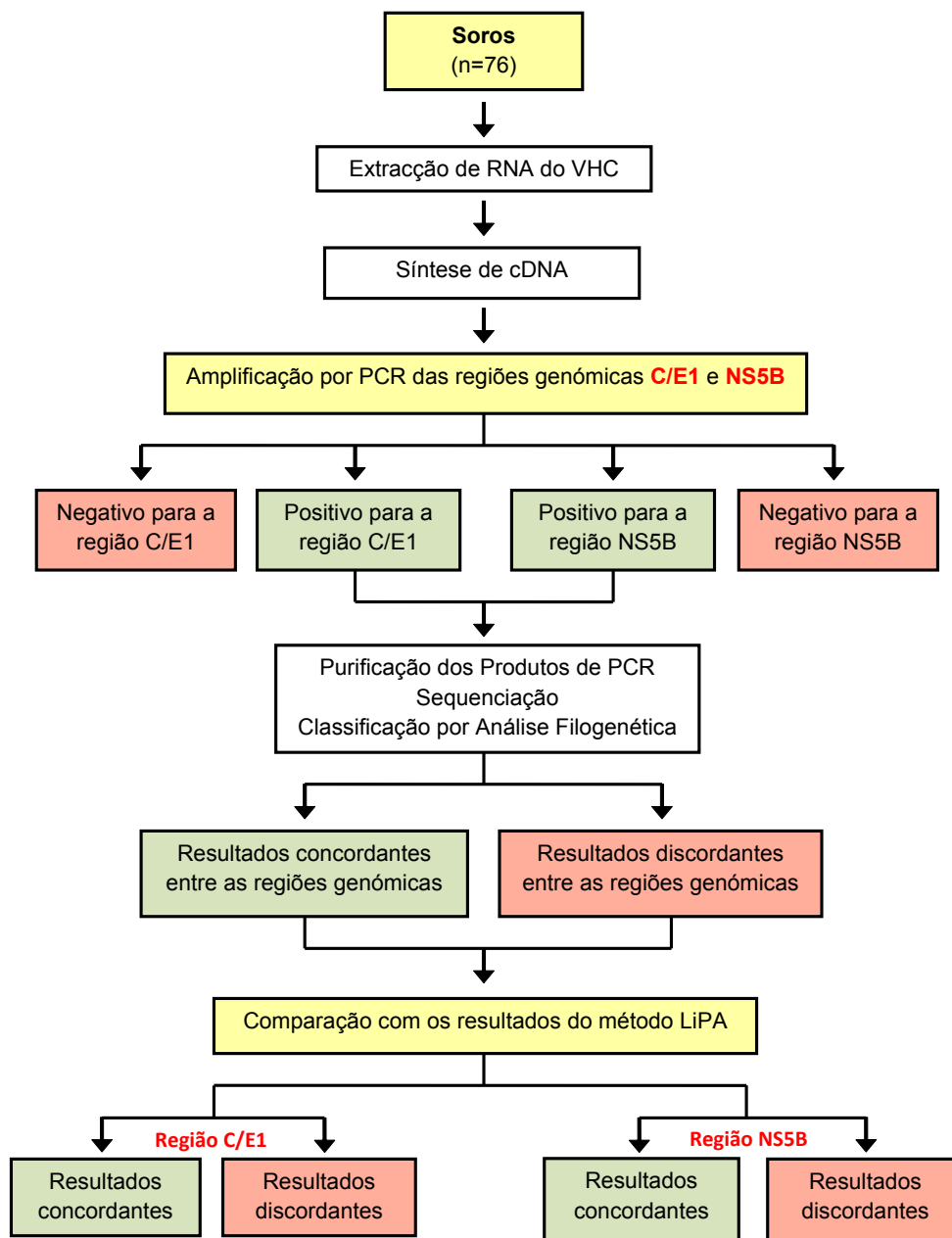


Figura 3.3 – Diagrama correspondente à estratégia molecular planeada para o estudo de genotipagem de amostras VHC positivas. Resultados concordantes indicam resultados iguais a nível da classificação do genótipo e subtipo, e resultados discordantes indicam resultados diferentes a nível do genótipo e/ou do subtipo do vírus.

3.3 Extracção do RNA Viral

A extracção do RNA viral a partir de soro foi realizada com o kit comercial “*QIAmp® Viral RNA Mini Kit*” (QIAGEN, Hilden, Germany) seguindo as instruções do fabricante. Na primeira etapa da extracção do RNA efectuou-se a lise das partículas virais através da adição de 560 µl de tampão de lise (Buffer AVL) contendo 5,6 µl de *carrier* RNA (facilita a ligação do RNA à membrana da coluna) a um volume de 140 µl de amostra. Posteriormente, misturou-se no vórtex e após uma incubação à temperatura ambiente durante 10 min, efectuou-se uma breve centrifugação. De seguida adicionou-se 560 µl de etanol (96-100%), agitou-se a amostra no vórtex e centrifugou-se por breves segundos. Depois transferiu-se cerca de 630 µl da mistura para a coluna e centrifugou-se a 6000 x g durante 1 min, posteriormente transferiu-se a coluna para um novo tubo colector e repetiu-se o procedimento anterior até toda a mistura ser transferida para a coluna. O RNA aderente na coluna foi purificado com 500 µl de tampão de lavagem AW1 e a coluna centrifugada a 6000 x g durante 1 min, seguida de nova lavagem colocando 500 µl de tampão AW2 com centrifugação à velocidade máxima durante 3 min, mudando os tubos colectores após cada centrifugação. A eluição do RNA foi realizada com a adição de 50 µl de Buffer AVE seguida de uma incubação de 1 min à temperatura ambiente, e centrifugação a 6000 x g durante 1 min. Este processo de eluição foi realizado duas vezes com a obtenção de cerca 80 µl de RNA viral puro que foi guardado a -80°C.

3.4 Síntese do cDNA

O cDNA foi sintetizado a partir do RNA extraído de forma indicada no ponto anterior, através de uma reacção de transcrição reversa (RT) usando o kit comercial “*RT – kit plus*” (Nanogen Advanced Diagnostics S.r.L., Italy), seguindo as indicações do fabricante. Na sala denominada de “limpa” preparou-se uma mistura de reacção com 1 µl de transcriptase reversa (MMLV-RT), 0,5 µl de inibidor de RNase (RNasi STOP) e 3,5 µl de água sem nucleases (Acqua Ultrapura) por cada amostra, incluindo também um controlo negativo. Essa mistura de reacção preparada foi pipetada para os “RT-MIX monotest tubes” que já possuem 10 µl de uma mistura de reacção de RT. Na sala de adição de RNA adicionou-se 10 µl do extraído enquanto para o controlo negativo da reacção, a adição de RNA foi substituída por água. Para alguns casos, aumentou-se o volume de RNA extraído para 12 µl e ajustou-se o volume de água sem nucleases (Acqua Ultrapura) para 1,5 µl, de modo a perfazer-se um volume final

de 25 µl. A síntese do cDNA decorreu num termociclador programado para efectuar as seguintes etapas: 25°C – 10 min; 37°C – 45 min; 95°C – 5 min.

3.5 Amplificação por PCR das Regiões C/E1 e NS5B do VHC

3.5.1 Desenho e Escolha de *Primers*

No desenho dos *primers* para a amplificação de fragmentos de VHC foram seleccionadas duas regiões genómicas, C/E1 e NS5B. A escolha da região C/E1 recaiu no facto de ser uma das regiões mais conservadas do genoma viral do VHC e também por ser a região já utilizada em ensaios comerciais, nomeadamente no método LiPA. Por outro lado, a preferência pela região NS5B é justificada por ser uma região relativamente conservada, mas que permite a diferenciação entre genótipos e subtipos [33]. Segundo a literatura, as regiões C/E1 e NS5B são os principais alvos para análise do VHC [98].

Os *primers* utilizados na amplificação da região NS5B do VHC foram desenhados com base num alinhamento de sequências nucleotídicas de diferentes genótipos e subtipos do VHC, disponíveis na base de dados *Los Alamos sequence database* (<http://hcv.lanl.gov/content/sequence/NEWALIGN/align.html>), incluindo a estirpe H77 (subtipo 1a) que foi usada como sequência de referência para a determinação da posição dos *primers*. A amplificação da região genómica NS5B foi realizada com os *primers* NS5B8245_f, NS5B8256_f e NS5B8616_r.

Contudo, para a amplificação da região genómica C/E1 do VHC foram utilizados *primers* designados de CEo_f, CEi_f e CE_r já publicados na literatura [99].

A Tabela 3.1 mostra as sequências dos *primers* e as suas respectivas posições na estirpe de referência H77, e a Figura 3.4 mostra esquematicamente o local de hibridação dos *primers* na estirpe H77 e o tamanho dos fragmentos obtidos.

Tabela 3.1 - Sequência dos *primers* utilizados na amplificação das regiões C/E1 e NS5B do genoma do VHC, e respectiva posição nucleotídica em acordo com a sequência de referência H77 (nº de acesso NC_004102).

Designação do <i>Primer</i> *	Direcção da sequência (5' → 3')	Localização no genoma
CEo_f [99]	GCA ACA GGG AA(CT) (CT)T(AGT) CC(CT) GGT TGC TC	834-859
CEi_f [99]	AA(CT) (CT)T(AGT) CCC GGT TGC TC(CT) TT(CT) TCT AT	843-868
CE_r [99]	TTC ATC ATC AT(AG) TCC CA(AGCT) GCC AT	1293-1315
NS5B8245_f	TGG GGT T(CT)T C(AGCT)T ATG A(CT)A CCC G(CA)T G	8245-8269
NS5B8256_f	TAT GA(CT) ACC CGC TG(CT) TTT GAC TC	8256-8278
NS5B8616_r	GG(AGCT) GC(AGCT) GA(AG) TAC CT(AGC) GTC ATA GC(CT) TCC GTG AA	8616-8647

* f = forward; r = reverse; i = inner; o = outer

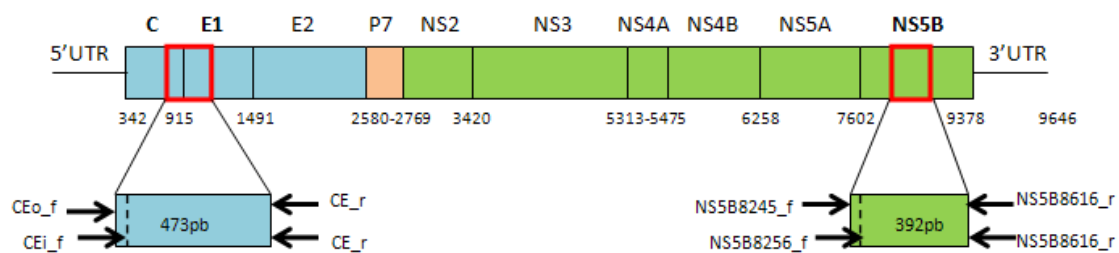


Figura 3.4 – Esquema das regiões genómicas C/E1 e NS5B amplificadas, *primers* utilizados na amplificação e o seu local de hibridação na estirpe H77 (nº de acesso NC_004102), e o tamanho dos fragmentos obtidos [7]. Caixas vermelhas indicam as regiões amplificadas.

3.5.2 Misturas de Reacção e Condições de Amplificação

Numa primeira fase, na amplificação das regiões C/E1 e NS5B foi utilizada a enzima *AmpliTaq® Gold DNA Polymerase* (Applied Biosystems, Foster City, USA), sendo que as misturas de reacção encontram-se descritas na Tabela 3.2.

Em alternativa, também foi utilizado o Sistema *illustra™ puReaq Ready-To-Go PCR Beads* (GE Healthcare, England), em que cada tubo possui uma esfera liofilizada, que quando hidratada para um volume final de 25 µl origina uma mistura de reacção constituída por 200 µM dNTPs, 10 mM Tris.HCl (pH 9,0 a temperatura ambiente), 50 mM KCl e 1,5 mM MgCl₂ (Tabela 3.2).

Na primeira reacção de amplificação do *heminested* PCR, foram usados os *primers* CEo_f e CE_r e os *primers* CEi_f e CE_r na segunda reacção, para

amplificação da região C/E1 do VHC. A amplificação da região NS5B do VHC foi realizada por *heminested* PCR com os *primers* NS5B8245_f e NS5B8616_r na primeira reacção e os *primers* NS5B8256_f e NS5B8616_r na segunda reacção.

Em ambas as misturas de reacção, a quantidade de cDNA adicionada na primeira reacção de amplificação variou entre os 5 µl e os 7 µl, enquanto na segunda reacção transferiu-se 2-3 µl de produto amplificado na reacção anterior. A água foi adicionada até perfazer um volume final de 25 µl.

Tabela 3.2 – Misturas de reacção correspondentes a cada enzima utilizada para amplificação das regiões C/E1 e NS5B do VHC.

Misturas de Reacção	Volume	Misturas de Reacção	Volume
<i>GeneAmp 10x PCR Gold Buffer</i> (150 mM Tris-HCl, pH 8.0, 500 mM KCl)	2,5 µl	<i>puReaq Ready-To-Go</i>	-
25 mM MgCl ₂	1,5 µl	<i>PCR Beads</i>	-
1,25 mM dNTPs	4 µl	10 µM <i>Primer forward</i>	1 µl
10 µM <i>Primer forward</i>	1 µl	10 µM <i>Primer reverse</i>	1 µl
10 µM <i>Primer reverse</i>	1 µl		
5U/µl <i>AmpliTaQ® Gold</i>	0,2 µl		

As condições de amplificação das regiões C/E1 e NS5B do VHC encontram-se descritas, respectivamente, nas Tabelas 3.3 e 3.4.

Tabela 3.3 - Condições de amplificação da região C/E1 do VHC por *heminested* PCR (*primers* CEo_f e CE_r seguidos de CEi_f e CE_r).

Amplificação da região C/E1 do VHC	Temperatura	Tempo de duração	Número de ciclos
Desnaturação inicial	95°C	5 min	1
Desnaturação	95°C	30 seg	
Hibridação	60°C	45 seg	5
Síntese	72°C	1 min	
Desnaturação	95°C	30 seg	
Hibridação	65°C	45 seg	35
Síntese	72°C	1 min	
Extensão final	72°C	7 min	1

Tabela 3.4 - Condições de amplificação da região NS5B do VHC por *heminested* PCR (*primers* NS5B8245_f e NS5B8616_r seguidos de NS5B8256_f e NS5B8616_r).

Amplificação da região NS5B do VHC	Temperatura	Tempo de duração	Número de ciclos
Desnaturação inicial	95 °C	5 min	1
Desnaturação	95°C	30 seg	
Hibridação	50°C	30 seg	5
Síntese	72°C	45 seg	
Desnaturação	95°C	30 seg	
Hibridação	55°C	30 seg	15
Síntese	72°C	45 seg	
Desnaturação	95°C	30 seg	
Hibridação	60°C	30 seg	25
Síntese	72°C	45 seg	
Extensão final	72°C	7 min	1

De modo a evitar possíveis contaminações durante as reacções de amplificação, a preparação da mistura de reacção e a adição de cDNA e de produtos de PCR foram realizadas em salas diferentes. Também se utilizou controlos negativos em que o cDNA ou o produto de PCR foi substituído por água.

3.5.3 Análise dos Produtos Amplificados

Os produtos de PCR amplificados foram sujeitos a electroforese em gel de agarose a 1,8% em tampão TAE 1x (4,84 g de Tris.base; 1,14 ml de ácido acético; 2 ml de EDTA a 0,5M pH 8,0; e 1000 ml de água bidestilada) com 5 µl de SybrSafe (*Invitrogen, California, USA*) para a sua visualização sob luz ultravioleta. Em cada poço do gel de agarose foram pipetados 2 µl de DNA amplificado de cada amostra com 2 µl de corante azul de bromofenol. Para identificação do fragmento genómico amplificado foi também aplicado no gel o marcador *DNA Molecular Weight Marker VIII* (*Roche, Roche Diagnostics, Germany*). A electroforese foi realizada a 120 volts durante aproximadamente 1 hora. A identificação das bandas de DNA realizou-se num transiluminador ultravioleta (*Gel Doc 2000, BioRad, USA*), em que se comparou a massa molecular dos fragmentos amplificados com o perfil de bandas do marcador molecular utilizado.

3.6 Purificação dos Produtos de PCR

Os produtos de PCR foram purificados através do kit comercial “*QIAquick® PCR Purification Kit*” (QIAGEN, Valencia, USA) seguindo as instruções do fabricante. Resumidamente, a cada volume de produto de PCR amplificado adicionou-se 5 volumes de Buffer PBI. Colocou-se a mistura numa coluna para ligação do DNA à respectiva membrana e centrifugou-se a 17 900 x g durante 1 min. Transferiu-se a coluna para um novo tubo colector. Na etapa seguinte das lavagens, adicionou-se 750 µl de Buffer PE e centrifugou-se a 17 900 x g durante 1 min, repetindo o processo de centrifugação em iguais condições e colocando a coluna em novos tubos colectores. O DNA foi eluído com 30 µl de Buffer EB (Tris-Cl 10 mM, pH 8,5) após centrifugação da coluna a 17 900 x g durante 1 min. O DNA purificado foi analisado por electroforese em gel de agarose a 1,8% tal como anteriormente referido.

3.7 Sequenciação dos Produtos Amplificados

Após a purificação realizou-se a sequenciação utilizando o kit comercial “*BigDye™ Terminator Cycle Sequencing Ready Reaction Kit*” (PE Applied Biosystems, Foster City, CA, USA), sempre seguindo as indicações do fabricante. Utilizou-se dois protocolos experimentais de sequenciação para cada amostra, respectivamente com e sem tampão de sequenciação na mistura de reacção efectuada para um volume total de 10 µl. Nas reacções com tampão, usou-se 1 µl de Big Dye terminator v.3.1, 3 µl de tampão de sequenciação, 1 µl de *primer* a 0,1 µM e 3 µl de água por amostra. Nas reacções em que não se utilizou tampão, este foi substituído por água. A mistura de reacção para sequenciação foi realizada na sala denominada de “limpa”. Posteriormente na sala de adição de produtos amplificados foram adicionados 2 µl de produto de PCR purificado. Nos casos em que o gel de agarose pós purificação dos produtos de PCR revelou uma concentração baixa de DNA, a quantidade foi aumentada na mistura de reacção tendo depois o volume de água sido ajustado. As reacções de sequenciação decorreram num termociclador programado com as seguintes condições: 1 ciclo a 96°C – 30 seg; 25 ciclos a 96°C – 30 seg, 50°C – 15 seg e 60°C – 4 min.

Os *primers* de sequenciação utilizados foram CEi_f e NS5B8256_f respectivamente para as regiões C/E1 e NS5B do VHC.

Os posteriores procedimentos de sequenciação para a obtenção dos cromatogramas foram realizados pela Unidade de Tecnologia e Inovação do INSA Dr. Ricardo Jorge.

3.8 Análise das Sequências Nucleotídicas

3.8.1 Obtenção das Sequências Consenso

Após finalizados os procedimentos práticos da sequenciação foram obtidos os respectivos cromatogramas. As sequências consenso foram construídas para cada amostra e região (C/E1 e NS5B) analisadas. Para tal, utilizou-se o programa *ClustalX 2.0.10*, onde se efectuou o alinhamento das sequências [100], sendo o tratamento manual das mesmas e a posterior obtenção da sequência consenso efectuados com auxílio do programa *Bioedit Sequence Alignment Editor 7.0.9.0* [101]. A cada sequência consenso foi atribuído um sistema de numeração iniciado pela letra “R” para as amostras de referência, e pela letra “C” para os casos clínicos, seguido pelo número da amostra.

3.8.2 Inferência Filogenética

Numa primeira abordagem, as sequências consenso foram submetidas a análise preliminar por BLAST, com recurso à ferramenta disponível na base de dados de VHC de Los Alamos, *HCV sequence database Blast search* (http://hcv.lanl.gov/content/sequence/BASIC_BLAST/basic_blast.html).

Posteriormente, as sequências consenso relativas às duas regiões (total dos casos estudados) foram alinhadas entre si e com sequências de referência de diversos subtipos e formas genéticas do VHC disponíveis na base de dados, com recurso ao programa *ClustalX 2.0.10* [100]. Através do *BioEdit* procedeu-se à edição e correcção manual dos alinhamentos [101].

A análise filogenética foi realizada no programa MEGA 4, versão 4.0.2 [102]. Neste programa, a construção das árvores, foi baseada no método de *Neighbor-Joining* a partir de múltiplos alinhamentos, e a matriz de distâncias foi determinada com base no modelo de *Kimura* a dois parâmetros [103]. A robustez da inferência filogenética foi avaliada através de *bootstrapping* de 1000 réplicas, estabelecendo relações entre as sequências analisadas e as sequências de referência de vários subtipos do VHC [104]. Os agrupamentos com valores de *bootstrap* iguais ou superiores a 70%, foram considerados estatisticamente significativos. Na construção das árvores usou-se o subtipo 7a do VHC como *outgroup*.

As sequências de referência dos vários genótipos e subtipos do VHC usadas na análise filogenética foram obtidas através da base de dados de VHC de Los Alamos, já referida anteriormente, disponível em

<http://hcv.lanl.gov/components/sequence/HCV/search/searchi.html>. Simultaneamente para ambas as regiões genômicas, C/E1 e NS5B do VHC, foram usadas as sequências de referência com os seguintes números de acesso: **Genótipo 1** – NC_004102, D10749, M62321, D10934, U01214, AY587016, AY051292, AM910652; **Genótipo 2** – AY746460, AB047645, AY232733, FN666428, AB031663; D00944; AF238482; AB047639; AB030907; AY232738; AB661373; D50409; FN666429; DQ155561; HM777358; HM777359; **Genótipo 3** – NC_009824, AF046866, D28917, D49374, F63821; **Genótipo 4** – DQ418783, DQ516084, DQ418784, EF589160, DQ418786, DQ516083; **Genótipo 5** – NC_009826, Y13184; **Genótipo 6** – AY859526, NC_009827, EF424629, D84263; **Genótipo 7** – EF108306; **Recombinante 1b/2k** – AY587845; **Recombinante 2k/1b** – HQ537005; **Recombinante 2b/1b** – DQ364460; **Recombinante 2/5** – AM408911; EF026073; **Recombinante 2i/6p** – DQ155560.

Na construção das árvores filogenéticas, também foram incluídas sequências de referência diferentes devido ao facto de algumas não possuírem a informação molecular para ambas as regiões genômicas em análise. Neste contexto, para a **região C/E1 do VHC** foram usadas as sequências de referência seguintes: **Genótipo 2** – D31972; AB327022; L38333; D49745; D49746; EF115979; D49754; D49757; HM777422; **Recombinante 1b/2k** – AY070214, AY070215; para a **região NS5B do VHC** foram usadas as sequências de referência seguintes: **Genótipo 2** – FJ872250; D10076; FJ872252; FJ872263; FJ872254; FJ872264; DQ220876; **Recombinante 1b/2k** – AB327052, DQ417439; **Recombinante 2k/1b** - HQ537006

3.8.3 Detecção de Recombinação Genética

Os casos cuja localização na árvore filogenética originou dúvidas foram analisados com maior detalhe por *bootscannig*, com o objectivo de investigar a ocorrência de eventos de recombinação ao longo da região estudada, e identificar quais as formas genéticas de VHC envolvidas.

Numa primeira fase utilizou-se a ferramenta bioinformática *Viral Genotyping* (<http://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi>), com o intuito de determinar quais as potenciais formas virais envolvidas nos eventos de recombinação. Posteriormente, a sequência de interesse e as sequências de referência representativas das várias formas virais, em particular, as previamente indicadas pela ferramenta bioinformática anteriormente referida, foram analisadas com auxílio do programa *SimPlot* 3.5.1 [105]. Este programa permitiu determinar o grau de

semelhança entre as sequências, que numa fase seguinte foram analisadas por *bootscannig*, no sentido de se obter informação sobre o número dos possíveis eventos de recombinação e o tamanho dos fragmentos que constituem o respectivo mosaico genético.

3.9 Análise Estatística

Para a análise de concordância dos dados foi aplicado o coeficiente de *Kappa* (*Cohen*) através do programa *Statistical Package for Social Studies* (SPSS) versão 17.0 (SPSS Inc., USA). Foi assumida uma boa concordância para valores de *Kappa* a partir de 0,61 [106].

4. RESULTADOS

4. RESULTADOS

4.1 Classificação de Sequências de VHC em Amostras de Referência

Foram analisados 4 soros de referência com informação previamente conhecida quanto à classificação do genótipo e subtipo do VHC. Duas amostras estavam classificadas como 2b e as outras duas como 3a. Por amplificação das regiões genómicas C/E1 e NS5B do VHC e, posterior sequenciação, análise das sequências por BLAST e estimativa filogenética, obtiveram-se resultados de classificação concordantes, tanto a nível do genótipo como do subtipo de VHC (Figuras 4.3 e 4.4). Atendendo a que um valor baixo de RNA de VHC (ou indetectável) pode ser um factor limitativo para a amplificação de fragmentos virais, observou-se que o método implementado demonstrava possuir uma boa sensibilidade para amplificação das regiões em análise das 2 amostras que apresentavam valores baixos de RNA de VHC (<6000 UI/mL).

4.2 Classificação de Sequências de VHC em Amostras Clínicas

4.2.1 Amplificação e Sequenciação das Regiões C/E1 e NS5B do VHC

Analisou-se uma população de conveniência constituída por 72 amostras clínicas. As amostras foram distribuídas em acordo com a classificação previamente conhecida dos genótipos de VHC obtida pelo método LiPA (Figura 3.2), e os grupos foram caracterizados pelos seguintes valores médios de RNA de VHC: 9 254 025 UI/mL (genótipo 1), 9 628 853 UI/mL (genótipo 2), 8 608 785 UI/mL (genótipo 3) e 1 029 428 UI/mL (genótipo 4). Com a excepção do grupo de amostras classificadas em genótipo 4, todos os restantes apresentavam valores de quantificação semelhantes. De igual forma, a mediana também foi calculada para os grupos dos diferentes genótipos, apresentando os valores de RNA de VHC seguintes: 8 500 000 UI/mL (genótipo 1), 2 510 000 UI/mL (genótipo 2), 4 152 447 UI/mL (genótipo 3) e 264 536 UI/mL (genótipo 4). Alguns casos incluídos nos 4 diferentes grupos apresentavam valores baixos de RNA de VHC (> 470 UI/mL e <9 720 UI/mL).

A partir do soro das amostras clínicas procedeu-se à extracção de RNA, síntese de cDNA, amplificação das regiões C/E1 e NS5B do VHC por *heminested*

PCR, e posterior visualização dos fragmentos obtidos em gel de agarose. A título de exemplo, alguns resultados obtidos são mostrados na Figura 4.1.

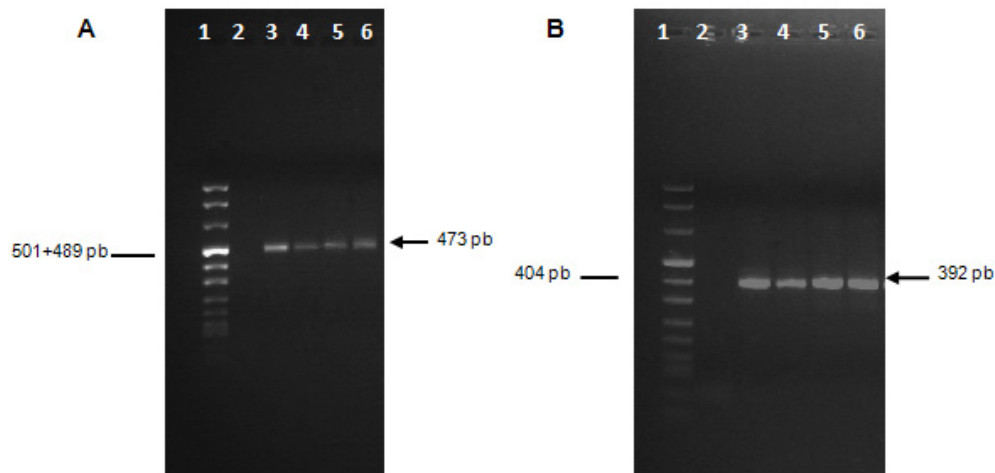


Figura 4.1 – Produtos amplificados das regiões C/E1 (A) e NS5B (B), visualizados em gel de agarose. No poço 1 foi aplicado o marcador *DNA Molecular Weight Marker VIII* (Roche, Roche Diagnostics, Germany) com um perfil de bandas variando entre 19 e 1114 pb. O poço 2 corresponde à aplicação de controles negativos e, respectivamente nos poços 3 a 6 foram aplicados 4 produtos amplificados com um fragmento C/E1 de 473 pb e um fragmento NS5B de 392 pb.

Dos 72 soros analisados, obteve-se sucesso na amplificação de 69 (95,8%) para a região C/E1 e de 71 (98,6%) para a região NS5B (Figura 4.2). Nas 3 amostras em que não se obteve amplificação da região C/E1, verificou-se que 2 apresentavam valores de RNA de VHC baixos (<3700 UI/mL), e a restante, um valor mais elevado (57500 UI/mL), mas bastante inferior à média de RNA de VHC quantificado na população de amostras estudada. Para a região NS5B, a única amostra que não amplificou apresentava um valor muito baixo de quantificação de RNA de VHC (847 UI/mL).

Para a região C/E1, as 69 amostras amplificadas foram purificadas e sequenciadas, obtendo-se, com exceção de 1 caso, cromatogramas de leitura molecular fácil e definida (n=68, 98,6%). De modo geral, os mesmos resultados foram obtidos para 68 (95,8%) das 71 amostras sequenciadas na região NS5B (Figura 4.2).

Contabilizando os resultados, obteve-se um total de 64 amostras (88,9%) com informação molecular disponível para ambas as regiões genômicas analisadas (C/E1 e NS5B).

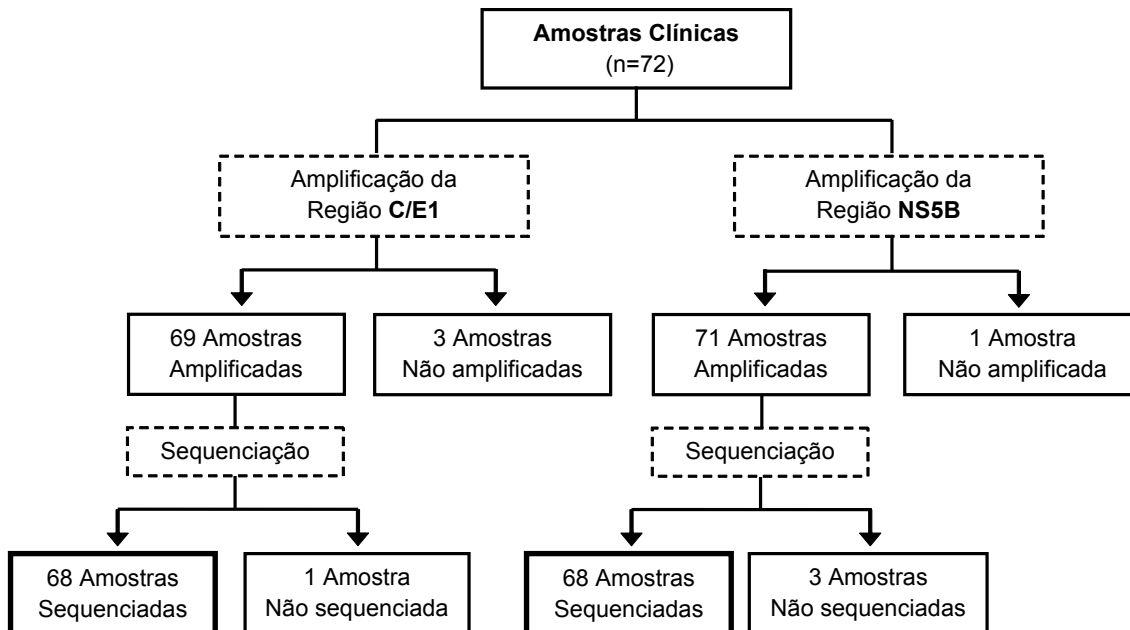


Figura 4.2 – Diagrama esquemático dos resultados de amplificação e sequenciação das amostras clínicas (n=72) para as regiões genômicas C/E1 e NS5B do VHC.

4.2.2 Análise Molecular das Sequências de VHC

Após a construção das sequências consenso (n=68 para cada região), estas foram submetidas a uma análise prévia por BLAST com recurso à base de dados de VHC de *Los Alamos*. Esta primeira abordagem permitiu a obtenção de uma classificação preliminar das sequências em estudo (dados não apresentados), orientando a escolha de sequências de referência nas posteriores análises a efectuar. Assim, foi seleccionado um total de 69 sequências de referência dos subtipos mais prevalentes na Europa, e também de subtipos e formas recombinadas de interesse para o presente estudo.

a) Análise da Região C/E1

A região C/E1 incluiu a análise de sequências nucleotídicas com comprimentos entre 389 e 470 pb, cujas diferenças no comprimento resultaram do grau de definição dos picos obtidos nas partes inicial e final dos cromatogramas analisados.

Na construção da árvore filogenética da região C/E1 foram comparados fragmentos com 448 pb de 68 sequências em estudo e 35 sequências de referência seleccionadas. As sequências de referência representando os diferentes genótipos de VHC formaram agrupamentos robustos e definidos na árvore, suportados por valores

de *bootstrap* iguais ou superiores a 78%, como se pode observar na Figura 4.3. A maioria das sequências C/E1 de amostras clínicas integraram o agrupamento com as sequências de referência do genótipo 1 (33/68). Contudo, 13 agruparam com sequências do genótipo 4, 11 com sequências do genótipo 2 e as restantes 11 com sequências do genótipo 3.

Os agrupamentos de sequências correspondentes aos subtipos 1a, 1b, 2a, 2b, 3a, 4a e 4d apresentaram-se pela topologia da árvore de forma clara e bastante robusta, apoiadas por elevados valores de *bootstrap* (1a, 1b, 2a, 2b, 3a e 4a, 100%, e 4d, 99%). A sequência C1 dentro do agrupamento de sequências classificadas como sequências do genótipo 2, agrupou com sequências de referência 1b/2k com um valor de *bootstrap* de 73%.

Atendendo à topologia da árvore, não foi possível classificar as sequências C4 e C7 quanto ao subtipo de VHC. No entanto, integradas no agrupamento filogenético representativo do genótipo 2, a análise prévia por BLAST indicou serem respectivamente semelhantes a sequências 2q e 2c. As sequências C4 e C7 apresentavam elevados valores de RNA de VHC, respectivamente de 5 790 000 UI/mL e 1 670 000 UI/mL.

A análise molecular das sequências correspondente à região C/E1 permitiu, assim, classificar por inferência filogenética 66 sequências (97,1%) das 68 em estudo. A maioria das sequências (25/66) foram classificadas em subtipo 1a, enquanto 11 foram classificadas em subtipo 3a, 9 em subtipo 4a, 8 em subtipo 1b, 4 em subtipo 2a, 4 em subtipo 2b, 4 em subtipo 4d e 1 na potencial forma recombinada 1b/2k (Tabela 8.1).

Entre as 66 sequências analisadas, não houve discrepâncias entre a classificação prévia obtida por BLAST e a classificação por análise filogenética.

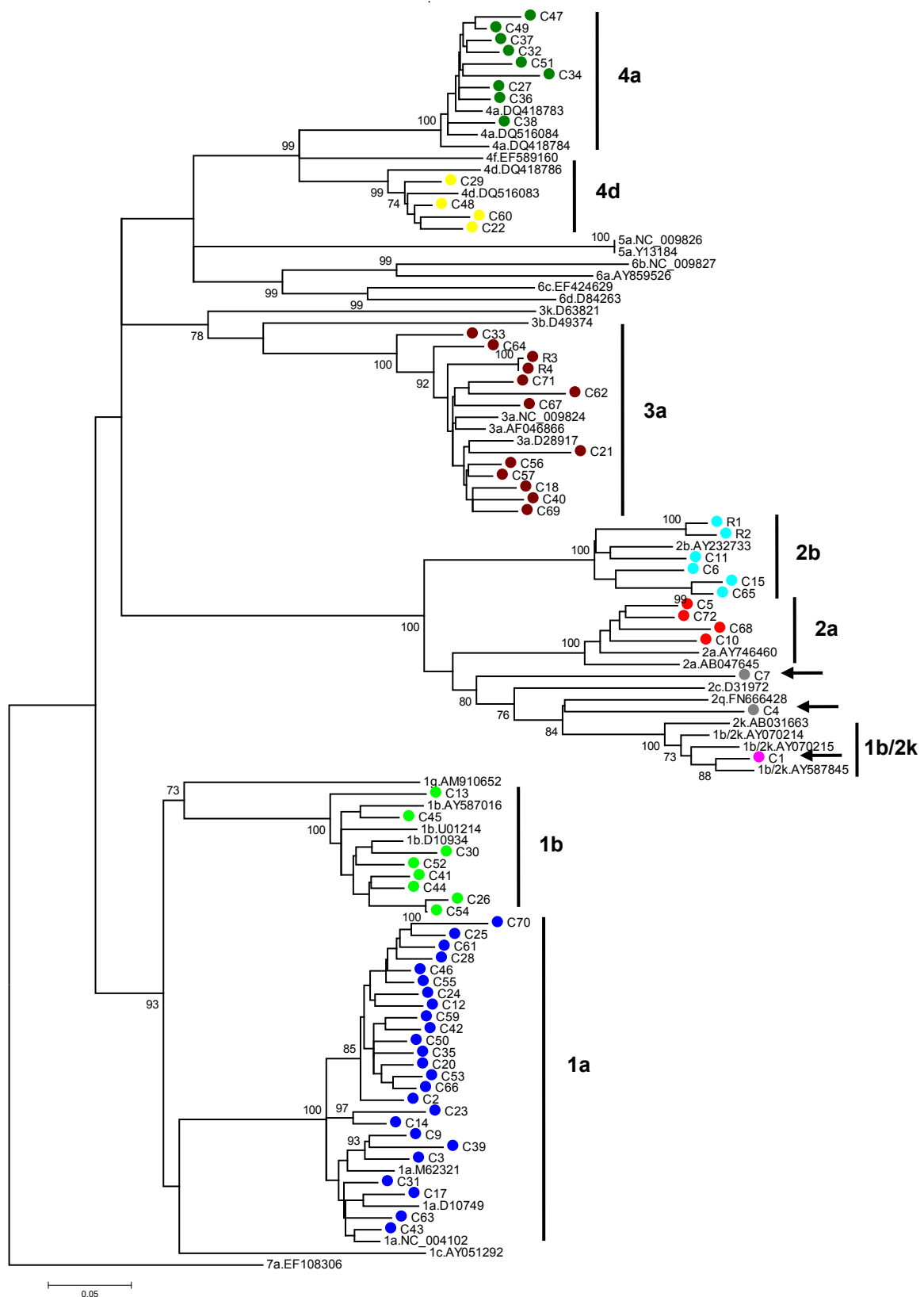


Figura 4.3 – Árvore Filogenética com base na região C/E1 do VHC. As cores representam as várias sequências estudadas agrupadas com sequências de referência dos subtipos indicados à direita. A cor cinzenta corresponde a amostras não subtipadas. As 4 amostras de referência estão representadas por “R” e as amostras clínicas por “C”. As setas assinalam as sequências analisadas por *bootscanning*.

Para região C/E1, apenas 2 dos 68 casos iniciais não foram subtipados por filogenia (sequências C4 e C7). Na tentativa de classificar estes 2 casos, construiu-se uma segunda árvore com um incremento de sequências representativas de todos os subtipos do genótipo 2 que se encontravam disponíveis na base de dados de VHC de *Los Alamos*. As 11 sequências em estudo classificadas na análise anterior como genótipo 2 foram comparadas com 31 sequências de referência. No entanto, a classificação do subtipo de VHC para ambos os casos continuou a não ser possível recorrendo apenas à inferência filogenética (Figura 4.4).

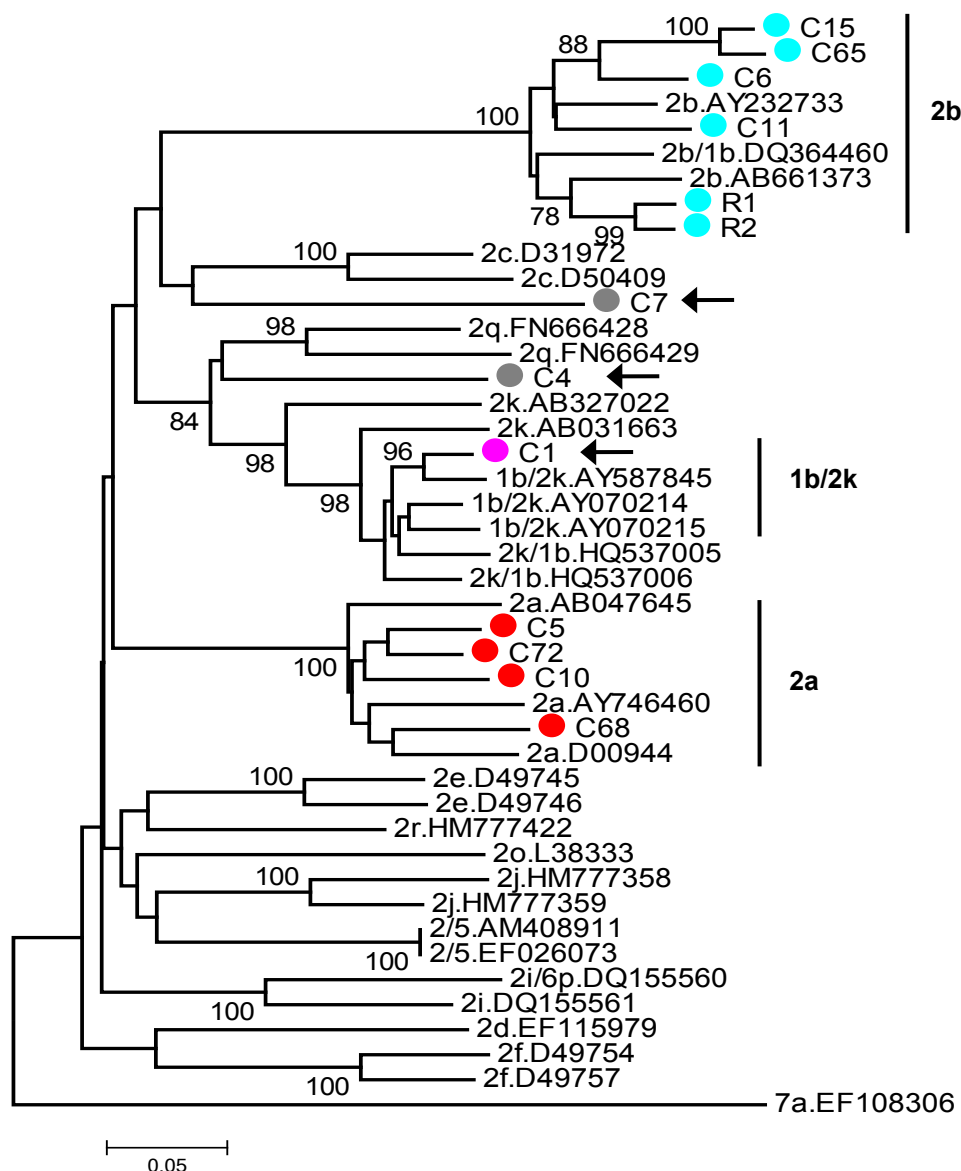


Figura 4.4 - Árvore Filogenética contendo maioritariamente sequências representativas do genótipo 2 com base na região C/E1 do VHC. As cores representam as várias sequências estudadas agrupadas com sequências de referência dos subtipos indicados à direita. A cor cinzenta corresponde a amostras não subtipadas. As 4 amostras de referência estão representadas por "R" e as amostras clínicas por "C". As setas assinalam as sequências analisadas por *bootscanning*.

Tanto a sequência C4 como a C7 encontram-se isoladas de qualquer agrupamento que possa definir um subtipo ou forma recombinada do VHC conhecida, sugerindo que eventualmente possa existir eventos de recombinação intragénica ao nível da região C/E1 analisada. Atendendo a esse facto, procedeu-se à análise por *bootscannig* para investigar essa hipótese. Primeiramente, utilizou-se a ferramenta *Viral Genotyping* que indicou os subtipos 2q e 2k, no caso da sequência C4, e os subtipos 2q e 2c, no caso da sequência C7, como sendo as formas genéticas envolvidas nos potenciais eventos de recombinação. Consequentemente, na construção do ficheiro a ser analisado no programa *SimPlot* foram incluídas sequências adicionais representativas de vários subtipos do VHC. A análise por *bootscannig* confirmou a presença das duas formas previamente indicadas. A sequência C4 era constituída por um primeiro segmento, com cerca de 250 pb, classificado como subtipo 2q, e por um segundo segmento, com aproximadamente 200 pb, classificado como subtipo 2k (Figura 4.5). No caso da sequência C7, o primeiro segmento, com cerca de 245 pb, foi classificado como subtipo 2q, e o segundo segmento, com aproximadamente 205 pb, foi classificado como subtipo 2c (Figura 4.6).

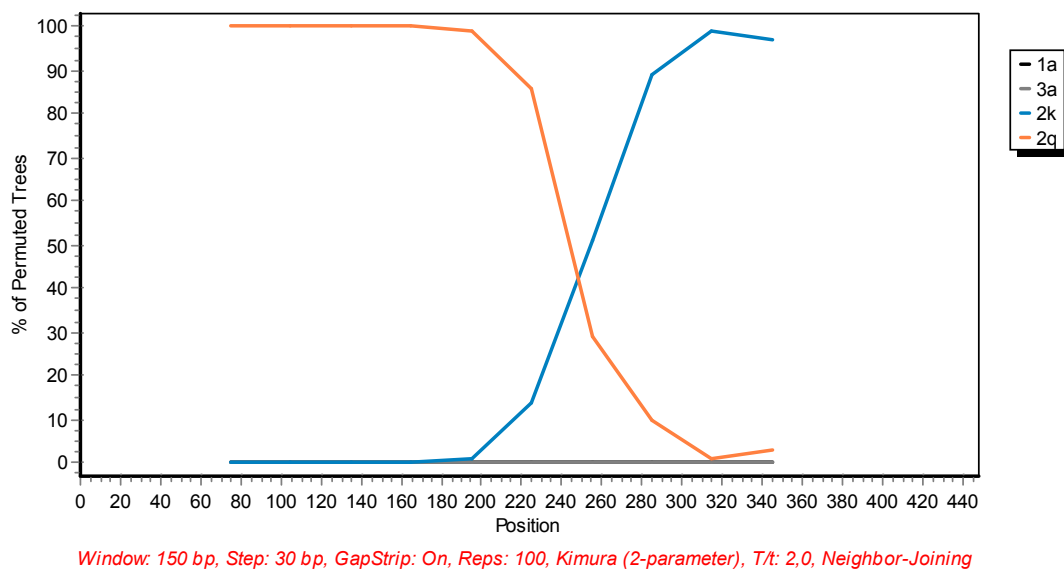
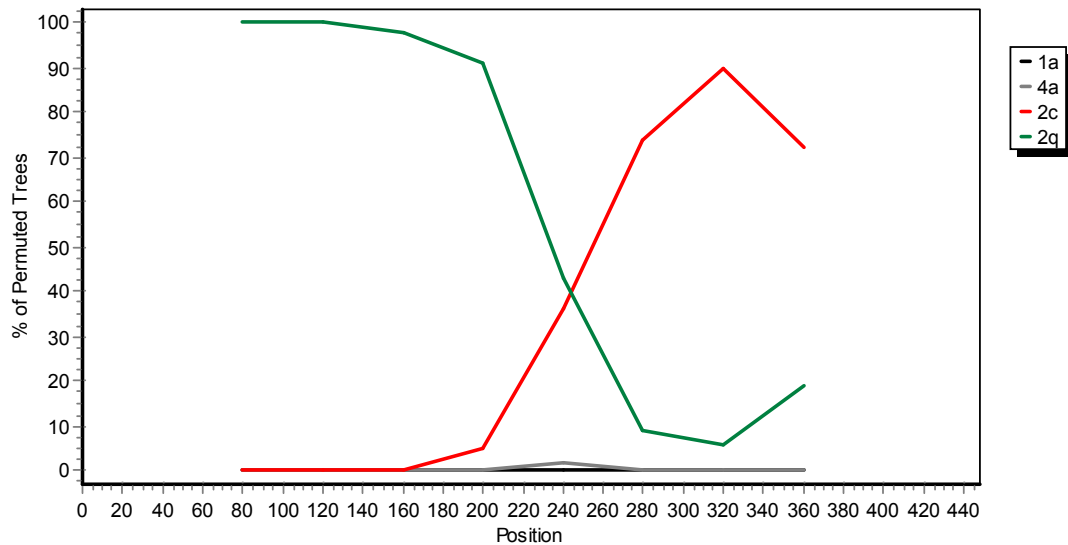


Figura 4.5 – Análise por *bootscannig* da região C/E1 da sequência C4. A vermelho estão indicados os parâmetros utilizados, e na legenda estão assinalados os subtipos do VHC incluídos na análise.



Window: 160 bp, Step: 40 bp, GapStrip: On, Reps: 100, Kimura (2-parameter), T/t: 2,0, Neighbor-Joining

Figura 4.6 - Análise por *bootscannig* da região C/E1 da sequência C7. A vermelho estão indicados os parâmetros utilizados, e na legenda estão assinalados os subtipos do VHC incluídos na análise.

A sequência C1, por estar agrupada com sequências de referência da forma recombinada 1b/2k, também foi analisada por *bootscannig*. A análise prévia com a ferramenta *Viral Genotyping* indicou apenas o subtipo 2k como estando envolvido na sua composição molecular. Seguidamente, a análise por *bootscannig* não indicou a ocorrência eventos de recombinação (dados não apresentados).

b) Análise da Região NS5B

As sequências nucleotídicas obtidas para a região NS5B (n=68) variaram entre 339 e 391 pb. A árvore filogenética para a região NS5B foi construída com base na comparação de fragmentos com 362 pb das 68 sequências em estudo e 35 sequências de referência seleccionadas. O conjunto de sequências referência representativas de cada um dos genótipos do VHC formaram agrupamentos bem definidos na árvore, suportados por valores de *bootstrap* iguais ou superiores a 85%, como se pode observar na Figura 4.7. A maioria das sequências NS5B obtidas a partir de amostras clínicas agrupou com as sequências de referência do genótipo 1 (31/68). Adicionalmente, 17 agruparam com sequências do genótipo 4, 10 com sequências do genótipo 2 e as restantes 10 com sequências do genótipo 3.

Os agrupamentos de sequências correspondentes aos subtipos 1b, 3a e 4d apresentaram-se na árvore de forma bastante robusta, com um valor de *bootstrap* de 100%. O mesmo foi observado para as sequências correspondentes aos subtipos 1a, 2a, 2b e 4a, apoiadas por elevados valores de *bootstrap* (respectivamente, 99%, 93%, 88% e 99%).

A topologia da árvore mostrou que a sequência C1 estava agrupada com sequências de referência 1b/2k, ainda que com um valor de *bootstrap* inferior a 70%. No entanto, este conjunto de sequências (C1 e referências 1b/2k) na análise da região IS5B integra o agrupamento definido como genótipo 1 com um *bootstrap* de 99%, e recorda-se, que na análise anterior efectuada para a região C/E1, incluíam o agrupamento de sequências definido como genótipo 2. Desta forma, pode ser colocada a hipótese de que se esteja em presença de uma forma de vírus recombinada.

Para as sequências C4 e C8, incluídas no agrupamento correspondente a sequências do genótipo 2, não foi possível definir uma classificação quanto ao subtipo e vírus em causa. No entanto, através de análise prévia por BLAST, as sequências C4 e C8 foram consideradas ser semelhantes a sequências respectivamente dos subtipos 2k e 2a. As sequências C4 e C8 apresentavam elevados valores de RNA de HC, respectivamente 5 790 000 UI/mL e 853 780 UI/mL.

Assim sendo, a análise molecular das sequências correspondente à região IS5B permitiu classificar filogeneticamente 66 (97,1%) das 68 sequências iniciais. A maioria das sequências (23/66) foram classificadas em subtipo 1a. No entanto, 10 sequências foram classificadas em subtipo 3a, 10 em subtipo 4a, 7 em subtipo 1b, 7 em subtipo 4d, 4 em subtipo 2a, 4 em subtipo 2b e 1 numa forma potencialmente recombinada 1b/2k (Tabela 8.1).

Apenas se verificou uma discrepância identificada para a sequência C1, entre a classificação prévia por BLAST e a classificação por análise filogenética. O resultado obtido em BLAST foi de subtipo 1b, e por filogenia, apesar de C1 se encontrar integrada no agrupamento definindo de sequências do genótipo 1, sub-agrupou com um valor de *bootstrap* (19%) com sequências de referência do subtipo 1b/2k.

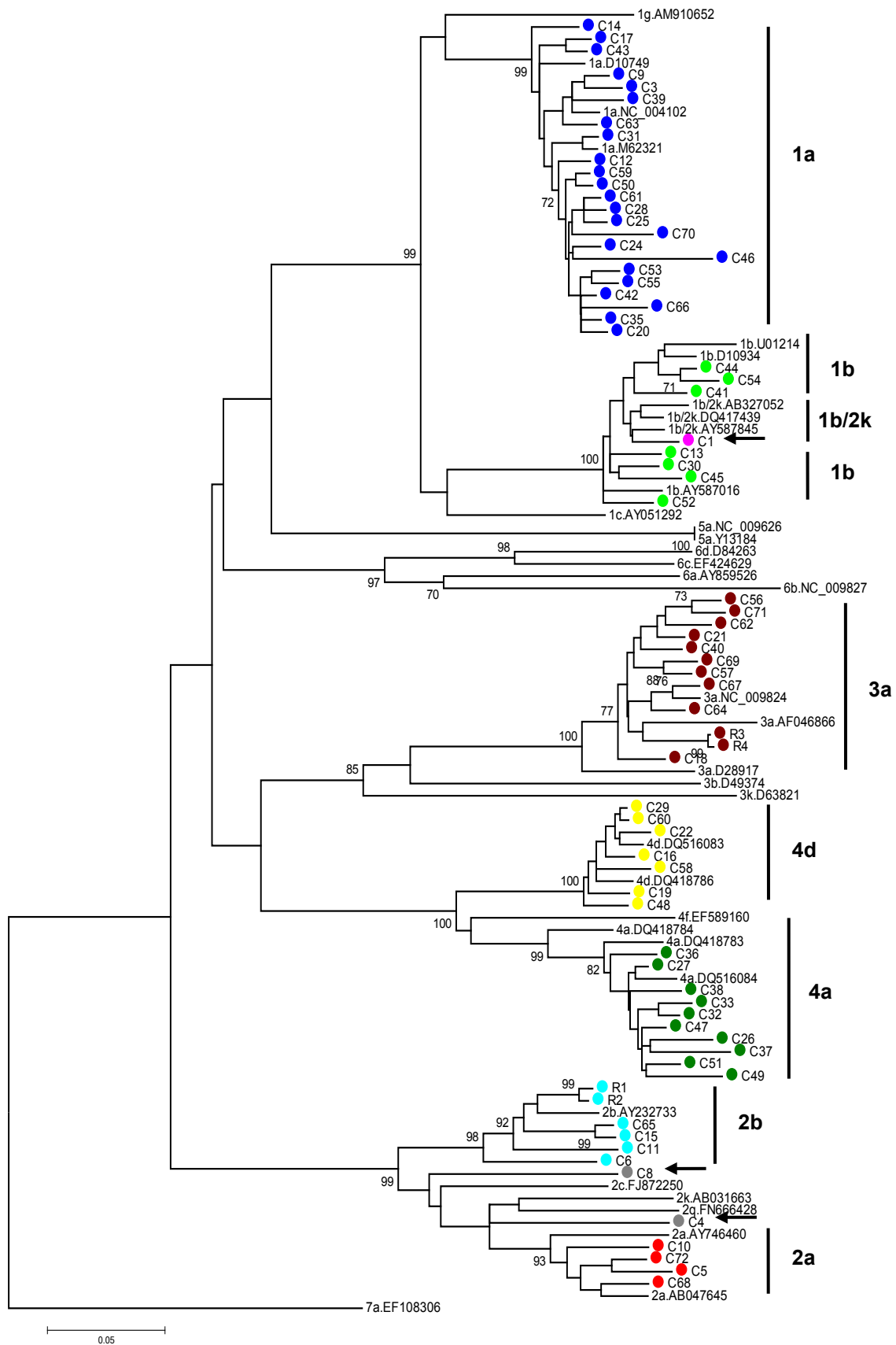


Figura 4.7 - Árvore Filogenética com base na região NS5B do VHC. As cores representam as várias sequências estudadas agrupadas com sequências de referência dos subtipos indicados à direita. A cor cinzenta corresponde a amostras não subtipadas. As 4 amostras de referência estão representadas por “R” e as amostras clínicas por “C”. As setas assinalam as sequências analisadas por *bootscanning*.

Na análise filogenética correspondente à região NS5B, observou-se insucesso na subtipagem de 2 das 68 sequências estudadas, embora ambas tenham sido incluídas no agrupamento definindo sequências do genótipo 2 (C4 e C8). No esforço de classificar estes 2 casos, construiu-se uma segunda árvore para a região NS5B com incremento de sequências representativas de todos os subtipos do genótipo 2 disponíveis na base de dados de VHC de *Los Alamos*. As 10 sequências em estudo classificadas na análise anterior como genótipo 2 foram comparadas com 19 sequências de referência. No entanto, não foi possível para ambos os casos, a classificação do subtipo do VHC (Figura 4.8).

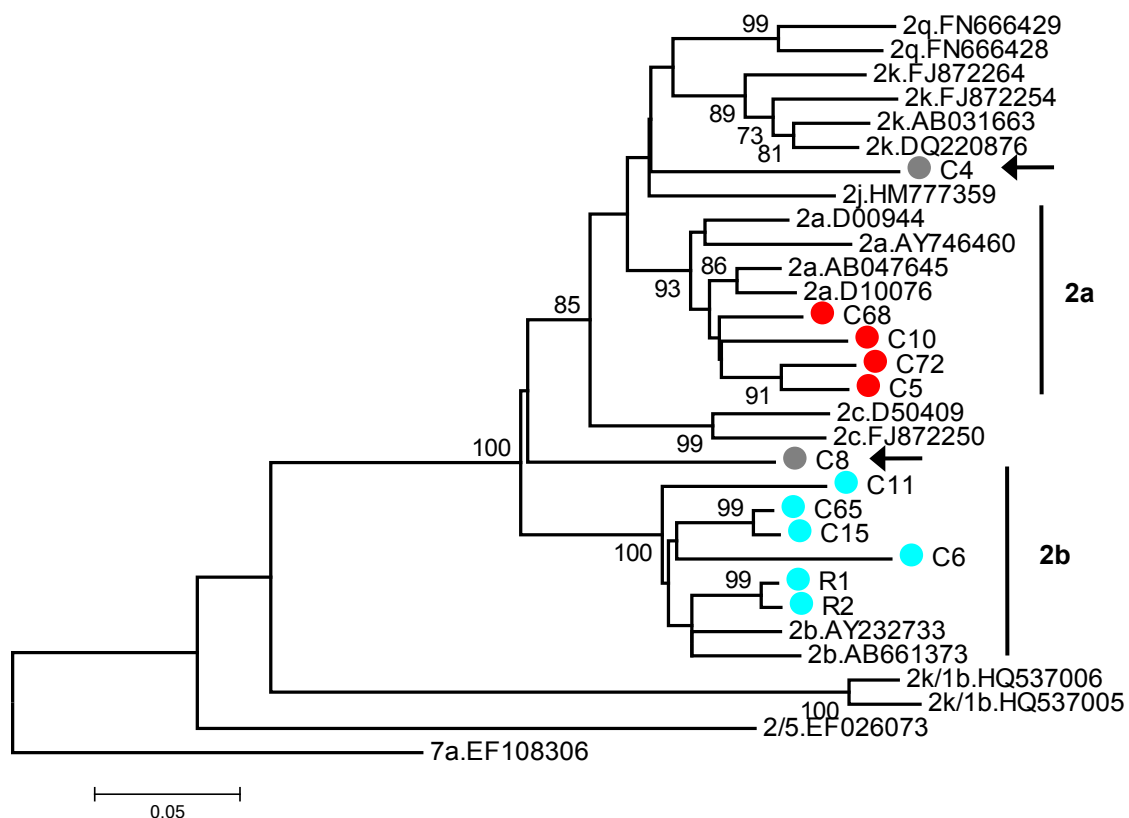
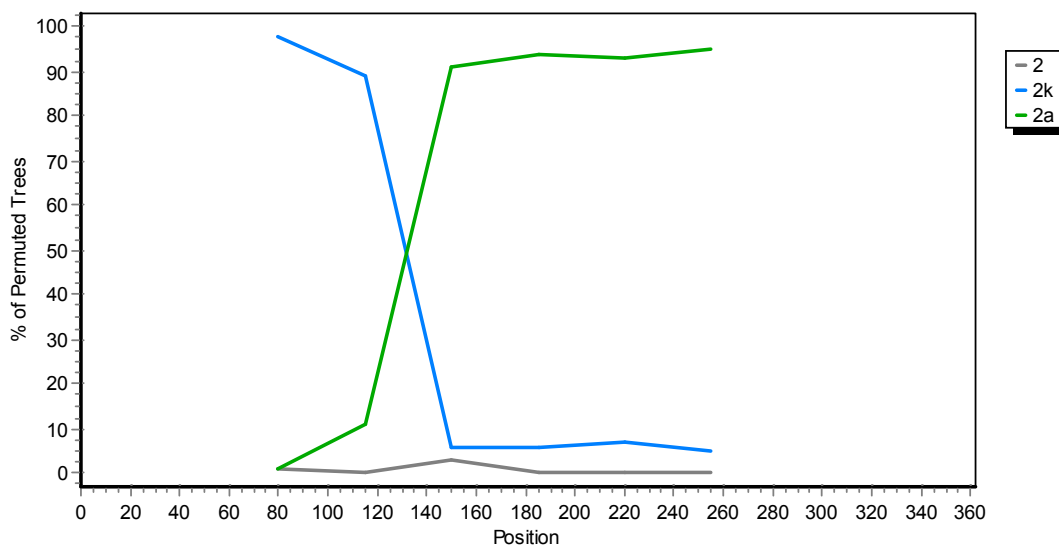


Figura 4.8 - Árvore Filogenética contendo maioritariamente sequências representativas do genótipo 2 com base na região NS5B do VHC. As cores representam as várias sequências estudadas agrupadas com sequências de referência dos subtipos indicados à direita. A cor cinzenta corresponde a amostras não subtipadas. As 4 amostras de referência estão representadas por "R" e as amostras clínicas por "C". As setas assinalam as sequências analisadas por *bootscanning*.

À semelhança da região C/E1, também para a região NS5B a sequência C4 manteve uma posição dúbia na nova árvore filogenética construída. Assim sendo, a posição das sequências C4 e C8, levantaram a hipótese da presença de possíveis formas recombinadas que foi posteriormente investigada.

A ferramenta *Viral Genotyping* indicou apenas o subtipo 2a como constituinte molecular da sequência C4. No entanto, a análise por *bootscanning* revelou um evento

de recombinação, sendo a sequência C4 constituída por um primeiro segmento, com cerca de 130 pb, semelhante a sequências classificadas como subtipo 2k, e por um segundo segmento, com aproximadamente 230 pb, classificado como semelhante a sequências do subtipo 2a (Figura 4.9). Para a sequência C8, a ferramenta *Viral Genotyping* indicou apenas o subtipo 2a, e a posterior análise por *bootscanning* não revelou qualquer evento de recombinação (dados não apresentados).



Window: 160 bp, Step: 35 bp, GapStrip: On, Reps: 100, Kimura (2-parameter), T/t: 2,0, Neighbor-Joining

Figura 4.9 - Análise por *bootscanning* da região NS5B da sequência C4. A vermelho estão indicados os parâmetros utilizados, e na legenda estão assinalados os subtipos do VHC incluídos na análise.

A sequência C1, por estar agrupada com sequências de referência da forma recombinada 1b/2k, embora sem valor de *bootstrap* significativo, também foi analisada por *bootscanning*. A análise prévia com a ferramenta *Viral Genotyping* indicou apenas o subtipo 1b, e seguidamente, a análise por *bootscanning* não confirmou a ocorrência de eventos de recombinação (dados não apresentados).

c) Análise Conjunta das Regiões C/E1 e NS5B

Nas 64 amostras em que se dispunha de informação molecular simultânea para ambas as regiões analisadas não foi possível por filogenia obter a classificação do subtipo de VHC para um único caso (C4). Foi obtida uma boa concordância pelo coeficiente de *Kappa* ($\kappa=0,931$) na comparação de resultados de classificação de genótipos de ambas as regiões genômicas. Na amostra C4 foi sugerido por análise de *bootscanning* potenciais eventos de recombinação em cada uma das regiões

genômicas analisadas, respectivamente 2q/2k para a região C/E1 e 2k/2a para a região NS5B do VHC.

Resultados de classificação discordantes entre as regiões genômicas analisadas foram repetidamente obtidos para 2 casos, C26 e C33, os quais foram respectivamente classificadas de subtipos 1b/4a (C/E1, NS5B), e 3a/4a (C/E1, NS5B) (Tabela 4.1). A análise por *bootscanning* destes 2 casos não sugeriu a existência de qualquer potencial evento de recombinação (dados não apresentados). De referir que a repetição dos procedimentos técnicos foi efectuada a partir do soro, incluindo a utilização de diluições que pudessem testar a presença de um vírus prevalente na amostra.

Tabela 4.1 - Comparação dos resultados de classificação de VHC por sequenciação e filogenia entre as regiões C/E1 e NS5B. Nas caixas a verde é indicado o número de casos com resultados discordantes. Na caixa a vermelho é indicado o caso filogeneticamente não subtipado (potencial evento de recombinação sugerido por análise de *bootscannig*: 2q/2k para a região C/E1 e 2k/2a para a região NS5B do VHC). Na caixa a amarelo é indicado o número de casos que agrupou com sequências de referência classificadas como 1b/2k em ambas as regiões analisadas.

Classificação de Casos de VHC Sequenciação e Estimativa Filogenética											
Região C/E1	Região NS5B									Total	
	1a	1b	1b/2k	2	2a	2b	3a	4a	4d		
1a	23									23	
1b	7							1		8	
1b/2k			1								1
2				1						1	
2a					4					4	
2b						4				4	
3a							10	1		11	
4a								8		8	
4d									4	4	
Total	23	7	1	1	4	4	10	10	4	64	

Em termos gerais, a classificação filogenética das sequências e a posterior análise por *bootscannig* dos casos considerados de interesse, revelou para a população de 64 amostras estudadas para ambas as regiões genômicas, a seguinte distribuição de subtipos: **1a**, 35,9% (n=23); **3a**, 15,6% (n=10); **4a**, 12,5% (n=8); **1b**, 10,9% (n=7); **4d**, 6,3% (n=4); **2a**, 6,3% (n=4) e **2b**, 6,3% (n=4). Adicionalmente, foram identificados 4 casos como formas potencialmente recombinadas (C1, C4, C26 e C33),

mas que em 2 deles também não pode ser excluída a hipótese de eventuais infecções VHC mistas (C26 e C33).

Prevalência de subtipos do VHC na população estudada

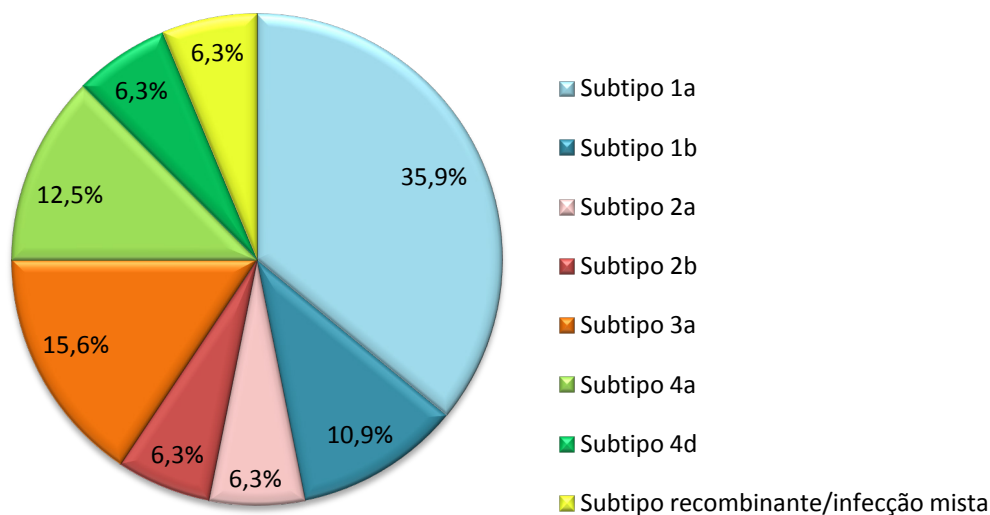


Figura 4.10 – Distribuição da proporção (%) de subtipos do VHC identificados na população estudada por análise conjunta das regiões C/E1 e NS5B do VHC.

4.3 Comparação de Resultados de Classificação do VHC pelo Método de Sequenciação e Análise Filogenética com o Método LiPA

a) Comparação dos Resultados da Análise da Região C/E1 com LiPA

Os resultados de genotipagem da região C/E1 do VHC obtidos através do novo método implementado foram comparados com os resultados previamente conhecidos obtidos pelo método LiPA. A comparação da classificação do genótipo do VHC mostrou a existência de diferenças em apenas 3 casos (C23, C26 e C33), como se pode observar em caixas assinaladas a verde na Tabela 4.2. Foi obtida uma boa concordância entre os dois métodos avaliada através do coeficiente de *Kappa* ($\kappa=0,935$).

Tabela 4.2 - Comparação entre os resultados de genotipagem obtidos pelo método LiPA com os resultados de sequenciação e filogenia da região C/E1 do VHC. Números em caixa a amarelo indicam o número de amostras com resultados concordantes entre os dois métodos. Números em caixa a verde indicam o número de amostras com resultados discordantes ao nível do genótipo entre os dois métodos. Números em caixa vermelha indicam o número de amostras não subtipadas por filogenia mas com sendo um potencial recombinante por análise de *bootscannig*, 2q/2k e 2q/2c. *Caso agrupado com sequências de referência classificadas como 1b/2k no agrupamento definido como genótipo 2.

LiPA (5'UTR + core)	Classificação do VHC										
	Resultados de Sequenciação e Filogenia (Região C/E1)										
	1a	1b	1b/2k*	2	2a	2b	3a	4a	4d	Sem resultado	Total
1	9										9
1a	12										12
1b	3	7									10
2				1	1	1					3
2a/2c			1	1	3					1	6
2b						3					3
3							1				1
3a							9				9
4							1	3	2	2	8
4a/4c/4d	1	1						6	2	1	11
Total	25	8	1	2	4	4	11	9	4	4	72

Tabela 4.3 - Número de amostras em que a classificação do subtipo do VHC está ausente, é discordante e/ou indiscriminada pelo método LiPA em comparação com a classificação obtida por sequenciação e estimativa filogenética com base na região C/E1. Resultado incompleto indica resultado não subtipado. Resultado discordante indica resultado diferente entre os métodos LiPA e análise filogenética, a nível do genótipo e/ou subtipo viral. Resultado indiscriminado indica resultado com indicação de vários subtipos. *Caso em que agrupou com sequências de referência classificadas como 1b/2k no agrupamento definido como genótipo 2. ⁸Casos não subtipados por análise filogenética mas com informação de ocorrência de eventos de recombinação por análise de *bootscannig*, 2q/2k e 2q/2c.

Número de amostras (n=68)	Resultados de Sequenciação e Filogenia (Região C/E1)	LiPA (Regiões 5'UTR + core)	
		Casos com resultado incompleto (n=18)	Casos com resultado discordante ou indiscriminado (n=19)
25	1a	9	3 (1b), 1 (4a/4c/4d)
8	1b	-	1 (4a/4c/4d)
1	1b/2k*	-	1 (2a/2c)
2	2 ⁸	1	1 (2a/2c)
4	2a	1	3 (2a/2c)
4	2b	1	-
11	3a	1	1 (4)
9	4a	3	6 (4a/4c/4d)
4	4d	2	2 (4a/4c/4d)

Nas 68 amostras em análise de comparação de resultados, em 19 amostras (27,9%) o método LiPA não conseguiu determinar o subtipo do VHC e em 15 amostras (22,1%) não conseguiu discriminar o subtipo do VHC em causa (Tabelas 4.2 e 4.3). O método LiPA conseguiu identificar o subtipo do VHC em apenas 34 amostras (50% dos casos). No entanto, nessas observou-se concordância entre a classificação pelos dois métodos em 31 amostras. Foi obtida uma baixa concordância dos resultados de subtipagem entre os dois métodos através do coeficiente de *Kappa* ($\kappa=0,283$).

Em 25 amostras classificadas pelo método de sequenciação como subtipo 1a, o método LiPA classificou 3 amostras em 1b e não conseguiu identificar o subtipo do VHC em 10 amostras: 9 delas foram classificadas em genótipo 1 e a restante foi incorrectamente classificada em 4a/4c/4d. Pela análise de sequenciação da região C/E1, observou-se também que 8 amostras foram classificadas em 1b. Contudo, uma delas foi classificada incorrectamente pelo método LiPA em 4a/4c/4d. (Tabela 4.3).

Relativamente a amostras classificadas na região C/E1 em genótipo 2, e excluindo 2 delas em que se verificou dificuldades na identificação do subtipo de VHC por ambos os métodos (C4 e C7), observou-se que o método LiPA não conseguiu identificar ou discriminar o subtipo do VHC em 6 amostras. Contudo, a análise da região C/E1 classificou 4 amostras em subtipo 2a, 1 em subtipo 2b e 1 num potencial recombinante 1b/2k (Tabela 4.2).

De um total de 11 amostras classificadas na região C/E1 em subtipo 3a, apenas 2 apresentaram resultados de classificação diferente dos obtidos pelo método LiPA: uma delas foi classificada em genótipo 3 e a outra em genótipo 4, ambas sem identificação do subtipo de VHC em causa (Tabela 4.3).

Relativamente aos 13 casos classificados na região C/E1 em genótipo 4 (4a ou 4d), não foi identificado para nenhum deles o subtipo de VHC pelo método LiPA. Adicionalmente, verificou-se também que para 3 casos classificados pelo método de sequenciação em 1a, 1b e 3a (caixas assinaladas a verde na Tabela 4.2), o método LiPA classificou-os incorrectamente em genótipo 4 (2 casos em genótipo 4 e 1 caso em 4a/4c/4d).

b) Comparação dos Resultados da Análise da Região NS5B com LiPA

A classificação dos casos de VHC para a região NS5B obtida através do novo método implementado foi comparada com os resultados previamente conhecidos obtidos pelo método LiPA. A comparação dos resultados de classificação quanto ao genótipo do VHC mostrou a existência de diferenças em apenas uma amostra, como se pode observar na Tabela 4.4 (caixa assinalada a verde. Assim, entre os dois métodos foi obtida uma boa concordância pelo coeficiente de *Kappa* ($\kappa=0,979$) para a classificação dos genótipos de VHC.

Tabela 4.4 - Comparação entre os resultados de genotipagem obtidos pelo método LiPA com os resultados de sequenciação e filogenia da região NS5B do VHC. Números em caixa a amarelo indicam o número de amostras com resultados concordantes entre os dois métodos. Números em caixa a verde indicam o número de amostras com resultados discordantes ao nível do genótipo entre os dois métodos. Número em caixa vermelha indica o número de amostras não subtipadas por filogenia mas com sendo um potencial recombinante por análise de *bootscannig*, 2k/2a. *Caso agrupado com sequências de referência classificadas como 1b/2k no agrupamento definido como genótipo 1. &Casos não subtipados por análise filogenética.

LiPA (5'UTR + core)	Classificação do VHC										Sem resultado	Total
	Resultados de Sequenciação e Filogenia (Região NS5B)											
	1a	1b	1b/2k*	2&	2a	2b	3a	4a	4d			
1	9											9
1a	12											12
1b	2	7								1		10
2				1	1	1						3
2a/2c			1	1	3					1		6
2b						3						3
3							1					1
3a							9					9
4								4	4			8
4a/4c/4d								6	3	2		11
Total	23	7	1	2	4	4	10	10	7	4		72

Tabela 4.5 - Número de amostras em que a indicação do subtipo do VHC está ausente, é discordante e/ou indiscriminada pelo método LiPA em comparação com a classificação obtida por sequenciação e estimativa filogenética com base na região NS5B. Resultado incompleto indica resultado não subtipado. Resultado discordante indica resultado diferente entre os métodos LiPA e análise filogenética, a nível do genótipo e/ou subtipo viral. Resultado indiscriminado indica resultado com indicação de vários subtipos. *Caso em que agrupou com sequências de referência classificadas como 1b/2k no agrupamento definido como genótipo 1. [‡]Casos não subtipados por análise filogenética mas com informação de ocorrência de eventos de recombinação na amostra não subtipada por análise de *bootscannig*, 2k/2a.

Número de amostras (n=68)	Resultados de Sequenciação e Filogenia (Região NS5B)	LiPA (Regiões 5'UTR + core)	
		Resultado não subtipado (n=21)	Resultado discordante ou indiscriminado (n=16)
23	1a	9	2 (1b)
7	1b	-	-
1	1b/2k*	-	1 (2a/2c)
2	2 [‡]	1	1 (2a/2c)
4	2a	1	3 (2a/2c)
4	2b	1	-
10	3a	1	-
10	4a	4	6 (4a/4c/4d)
7	4d	4	3 (4a/4c/4d)

Nas 68 amostras em análise de comparação de resultados, em 21 amostras (30,9%), o método LiPA não conseguiu identificar o subtipo de VHC e em 14 amostras (20,6%) também não conseguiu discriminar o subtipo do VHC em causa (Tabelas 4.4 e 4.5). O método LiPA conseguiu identificar o subtipo do VHC em apenas 33 amostras (48,5% dos casos). No entanto, observou-se concordância entre a classificação pelos dois métodos em 31 amostras. Quando agrupados os subtipos de VHC, foi obtida uma fraca concordância pelo coeficiente de *Kappa* ($\kappa=0,284$) na classificação de subtipos de VHC entre os dois métodos.

Em 23 amostras classificadas pelo método de sequenciação como subtipo 1a, o método LiPA classificou incorrectamente 2 amostras em 1b e não conseguiu identificar o subtipo do VHC em 9 amostras (classificadas em genótipo 1) (Tabela 4.5).

Pela análise de sequenciação da região NS5B, observou-se que 7 amostras foram classificadas em 1b sendo a classificação concordante com a obtida pelo método LiPA (Tabela 4.4).

Quanto as amostras classificadas na região NS5B em genótipo 2, e excluindo 2 delas em que se verificou dificuldades na identificação do subtipo de VHC por ambos os métodos (C4 e C8), observou-se que o método LiPA não conseguiu identificar ou

discriminar o subtipo de VHC em 6 amostras. Contudo, a análise da região NS5B classificou 4 amostras em subtipo 2a, 1 em subtipo 2b e 1 numa forma potencialmente recombinada 1b/2k (Tabela 4.4).

De um total de 10 amostras classificadas na região NS5B em subtipo 3a, apenas 1 amostra apresentou um resultado de classificação incompleta pelo método LiPA (classificada em genótipo 3) (Tabela 4.5).

Para o genótipo 4, em 8 amostras não subtipadas pelo método LiPA, 4 foram classificadas na região NS5B de subtipo 4a e outras 4 de subtipo 4d. O método LiPA também não conseguiu discriminar o subtipo de VHC em 9 amostras classificando-as em 4a/4c/4d. Porém, a análise da região NS5B revelou que 6 amostras eram do subtipo 4a e 3 eram do subtipo 4d. O método LiPA não identificou para nenhuma amostra do genótipo 4 o subtipo de VHC em causa (Tabela 4.4).

5. DISCUSSÃO

5. DISCUSSÃO

O VHC apresenta uma elevada variabilidade genética traduzida na sua classificação em 6 diferentes genótipos que diferem entre si em 30-35% por análise das suas sequências genómicas [32]. Recentemente, foram identificadas sequências de VHC divergentes e que integram uma proposta de um novo potencial genótipo que foi designado de genótipo 7 [36, 37]. Além da distribuição em genótipos, o VHC possui mais de 80 subtipos que diferem entre si em 20-25% nas suas sequências genómicas [38]. A grande diversidade genética das estirpes de VHC é fundamentada, em parte, pelo erros cometidos pela polimerase de RNA do VHC, e também, pelo consequente desenvolvimento de *quasispecies* cujas sequências virais podem diferir até 10%. Estes factos contribuem para a virulência e persistência da infecção do VHC e para a dinâmica da transmissão e evolução do vírus [29, 30, 35].

A genotipagem do VHC é considerada uma importante ferramenta no prognóstico e evolução da infecção/doença, possuindo impacto na dosagem e duração do tratamento. A epidemiologia molecular do vírus, a sua distribuição geográfica e a evolução no tempo, podem estar correlacionados com factores de risco para a aquisição da infecção, nomeadamente o uso de drogas injectáveis ou transfusões de sangue não seguras, sobretudo em países com fracos recursos económicos [33].

Existem diferentes métodos moleculares para a classificação do VHC, sendo a sequenciação considerada um método de referência. Contudo, a forma mais correcta de classificar o VHC implica a sequenciação do genoma completo, facto que por ser moroso se torna impraticável em laboratórios que pretendem dar uma resposta analítica em tempo útil [95]. Em alternativa, vários kits comerciais foram desenvolvidos e validados para os procedimentos de rotina na determinação dos genótipos e subtipos do VHC, tendo como alvo a detecção por hibridação de sondas com produtos amplificados de uma porção do genoma do VHC [97].

A confiança nos métodos de genotipagem depende fortemente da quantidade de informação, ou seja, do número de *sites* informativos que são utilizados para a discriminação das variantes genéticas. A região 5'UTR é suficientemente variável para a discriminação dos genótipos 1 a 5 do VHC e da maioria dos subtipos do genótipo 6 [94, 107-109]. No entanto, esta região pode não permitir a discriminação entre os subtipos 6c a 6l e o genótipo 1, e também entre os subtipos 1a e 1b [96, 108]. Erros de classificação em genótipo 6 quando em presença de genótipo 1, podem conduzir a

uma duração do tratamento do doente inadequada e a uma interpretação incorrecta na resposta do paciente à terapia. De facto, indivíduos infectados pelo genótipo 6 apresentam uma melhor resposta à terapia que indivíduos infectados pelo genótipo 1 [110]. Em métodos comerciais, a correcta identificação dos subtipos 1a e 1b pela análise da região 5'UTR ocorre em 80 a 95% dos casos, nomeadamente por falta de especificidade das sondas utilizadas na detecção destes subtipos virais [95]. Para ultrapassar estas limitações foi desenvolvida uma nova versão do ensaio comercial, *Versant HCV genotype assay (LiPA) 2.0* (genericamente designado por método LiPA), que para além de ter como alvo uma região 5'UTR de 240 pb inclui também uma região do *core* do VHC de 270 pb. Desta forma, potencialmente existirá uma melhoria no desempenho do método para a determinação do genótipo e subtipo do VHC [46, 111]. Este ensaio de genotipagem do VHC é o mais vulgarmente utilizado nos laboratórios do país.

Em estudos moleculares tem-se observado a implementação da sequenciação de outras regiões genómicas do VHC em particular os genes codificantes da proteína não estrutural NS5B e das proteínas estruturais do *core* e E1, as quais, por serem regiões menos conservadas, proporcionam a identificação de diferenças nas sequências que permitem distinguir e classificar genótipos e subtipos virais do VHC [93]. A sequenciação de uma porção específica do genoma é assim o método de referência alternativo, com amplificação e sequenciação de um fragmento genómico, e uma posterior análise preliminar das sequências geradas por BLAST recorrendo a uma importante base de dados de VHC, que actualmente contém mais de 80 mil sequências identificadas no mundo [39]. Adicionalmente, a estimativa filogenética para classificação das sequências do VHC pode também ser uma abordagem a considerar.

Embora a sequenciação necessite de equipamento sofisticado de elevado custo e de técnicos qualificados para executar as respectivas tarefas, actualmente estes serviços encontram-se comercialmente disponíveis por empresas especializadas na área, oferecendo preços muito competitivos. Por outro lado, a análise das sequências pode ser actualmente mais rápida e acessível através de programas bioinformáticos disponíveis *online* e com uma relação benefício/custo muito favorável.

Deste modo, o objectivo do presente trabalho foi desenvolver um método alternativo de genotipagem do VHC, no contexto de um laboratório de referência, utilizando a amplificação, a sequenciação e a análise filogenética de duas regiões genómicas do VHC, distintas e relativamente conservadas, que por um lado permitissem diferenciar os subtipos virais de forma mais precisa, e por outro lado, não

limitassem a detecção de formas virais geneticamente recombinadas. Com esta abordagem pretendia-se obter uma melhoria na classificação do vírus e uma simultânea redução de custos.

Numa primeira fase, e também recorrendo à literatura, seleccionou-se as regiões C/E1 e NS5B como alvos principais da análise molecular do vírus. Foram seleccionados *primers* referenciados em publicações, ou desenhados *de novo*, a partir de alinhamentos de sequências de referência contendo sequências de diferentes génotipos e subtipos de VHC, preferencialmente descritos na Europa e provenientes de estudos efectuados na população de infectados por VHC residente em Portugal (base de dados de sequências de VHC de *Los Alamos*) [40, 41].

Alguns estudos mostraram o potencial de discriminação de subtipos para a região NS5B, mas outros descrevem proporções de 18,3% de resultados falso-negativos [112, 113]. Contudo outros autores descrevem uma elevada sensibilidade de detecção do VHC (97,3%), mas salvaguardam a necessidade de se utilizar diferentes pares de *primers* na amplificação do vírus [37].

Os protocolos de amplificação utilizados foram definidos após um período de optimização das condições de reacção em função das temperaturas de *melting* dos *primers* e das enzimas utilizadas. De modo a diminuir o risco de contaminação das amostras em estudo, os procedimentos de extracção de RNA, a preparação das misturas de reacção (de RT, de PCR e de sequenciação), a adição de amostra (de RNA e DNA), a amplificação e detecção dos produtos, e posteriormente, a sua purificação (pré-sequenciação) foram realizados em espaços fisicamente separados. Nos diferentes espaços foram também incluídos controlos negativos nas reacções.

A população estudada, constituída por 72 soros de reclusos infectados, foi submetida à amplificação por *heminested* PCR das regiões C/E1 e NS5B do VHC. Houve um sucesso de amplificação de 95,8% (69/72) e 98,6% (71/72) respectivamente para as regiões C/E1 e NS5B. Várias razões podem ser apontadas para os casos em que não se observou a amplificação de fragmentos virais. Entre elas, podemos salientar uma possível degradação do RNA viral resultante de vários ciclos de congelação e descongelação a quando da manipulação dos soros. Em alguns casos, também o baixo valor de carga viral e a elevada diversidade genética do VHC podem ter dificultado a hibridação dos *primers* nas reacções de amplificação.

Por outro lado, também para alguns casos amplificados que posteriormente foram sequenciados, não foi possível chegar a classificação do VHC pela baixa qualidade das sequências geradas. Assim, foram sequenciados com sucesso para as regiões C/E1 e NS5B, respectivamente, 98,6% (68/69) e 95,8% (68/71) dos casos.

Analisando apenas os 64 casos (88,9%; 64/72) em que foi possível obter informação molecular conjunta para as duas regiões genômicas, observou-se em 2 casos (C26 e C33) uma classificação discordante para as regiões C/E1 e NS5B do VHC, respectivamente 1b/4a e 3a/4a. Atendendo a que estes resultados foram mantidos, mesmo após a repetição de todo o procedimento de amplificação e sequenciação com utilização de amostras de RNA de VHC diluídas, esta classificação levantou a hipótese da presença de vírus recombinantes (e/ou eventualmente infecções mistas que discutiremos mais adiante). A classificação do VHC nos restantes 62 casos revelou a distribuição seguinte: 1a (37,1%, n=23), 3a (16,1%; n=10), 4a (12,9%; n=8), 1b (11,3%; n=7), 2a (6,5%; n=4), 2b (6,5%; n=4), 4d (6,5%; n=4), 2 (1,6%; n=1) e 1b/2k (1,6%; n=1).

Na comparação dos resultados obtidos pelo método de sequenciação desenvolvido com o método LiPA, observou-se uma concordância na classificação do genótipo do VHC de 95,6% (65/68) para a região C/E1, e de 98,5% (67/68) para a região NS5B. Estes resultados são semelhantes aos descritos na literatura que referem uma concordância de 99,5% entre métodos de genotipagem por LiPA e a sequenciação da região NS5B [114].

Na população de amostras em análise, resultados conhecidos de genotipagem pelo método LiPA revelaram que 29,2% (21/72) dos casos não tinham sido subtipados (a que designámos de resultados incompletos), e em 23,6% (17/72) dos casos não tinha sido possível discriminar o subtipo viral (a que designámos de resultados indiscriminados ou indefinidos). Os resultados incompletos foram principalmente obtidos para os casos de genótipo 1 (12,5%; 9/72) e de genótipo 4 (11,1%; 8/72), e os resultados indefinidos foram identificados nos casos 4a/4c/4d (15,3%; 11/72) e nos casos 2a/2c (8,3%; 6/72). Deste modo, a contabilização dos resultados incompletos e indefinidos pelo método LiPA na população estudada correspondia a uma proporção de 52,8% (38/72). O método implementado permitiu diminuir acentuadamente a proporção de indefinição na classificação do VHC, conseguindo a subtipagem em 95,8% (69/72) dos casos. Assim, pelo novo método implementado, apenas 3 (4,2%) casos (C4, C7 e C8) não foram subtipados e 2 deles (C4 e C7) correspondem a casos

com informação molecular disponível apenas para uma das regiões genómicas analisadas.

Foram descritos erros de classificação entre os subtipos 1a e 1b, associados à existência de um polimorfismo identificado na posição 99 do genoma do VHC, e que é utilizada frequentemente para diferenciar estes subtipos virais [95]. Por outro lado, o nucleótido G na posição 243 da região 5'UTR foi considerado representativo do subtipo 1b, embora já tenha sido encontrado em populações classificadas como subtipo 1a [112, 115]. Estudos descrevem que um desempenho de 97,7% pode ser atingido na identificação do subtipo 1a, e esta proporção pode descer para os 58,5% quando a informação se basearia apenas na análise da região 5'UTR. Assim, vários estudos confirmam o benefício da análise conjunta da região do *core* e da região 5'UTR para uma classificação correcta do subtipo 1a [97, 111].

No presente estudo, apesar do método LiPA utilizar as regiões do *core* e 5'UTR do VHC foram encontradas algumas discrepâncias relativamente à classificação de subtipos 1a e 1b. Assim, 3 casos classificados de subtipo 1a (C2, C59 e C63) foram identificados de subtipo 1b pelo método LiPA. Alguns estudos sobre este ensaio referem uma detecção do subtipo 1b exacerbada em comparação ao subtipo 1a, estimando uma taxa de detecção do subtipo 1b cerca de 13% maior do que a taxa real [95]. Assim, se o método LiPA ocasionalmente pode classificar casos de subtipo 1a como subtipo 1b, e se este é o método mais utilizado no país para classificação do VHC, seria de esperar a existência de muitos casos incorrectamente classificados como subtipo 1b.

A título de curiosidade, a consulta à base de dados de sequências de VHC de Los Alamos, revelou que o subtipo 1b possui a maior proporção de sequências de VHC tanto ao nível mundial (39,1%), como europeu (51,2%), sendo seguida pela proporção de sequências do subtipo 1a [39, 40]. No entanto, em Portugal, observou-se que a proporção de sequências 1b constitui apenas 4,7% da totalidade de sequências disponíveis, bastante inferior à proporção de 45,5% de sequências classificadas como subtipo 1a [41]. Aparentemente, a baixa frequência observada de subtipo 1b pode conduzir a uma baixa frequência de erro na classificação destes subtipos em 1a, quando utilizado o método LiPA. Contudo, é de salientar que a base de dados possui apenas 257 sequências de VHC identificadas em indivíduos residentes em Portugal, e pode não representar a distribuição dos genótipos e subtipos de VHC na população, mas sim as frequências das sequências dos vários genótipos e subtipos que foram submetidas nessa mesma base de dados.

Segundo um estudo com determinação do genótipo e subtipo de VHC pelos métodos de sequenciação e LiPA mostrou um padrão na população portuguesa em

geral, em que o subtipo 1a surge com elevada prevalência (36,8%), seguido do subtipo 1b (20,2%) e do subtipo 3a (24,2%) [67]. A análise das regiões C/E1 e NS5B do VHC permitiu classificar em subtipo 1a os 9 casos que com o método LiPA não tinha sido possível a subtipagem do VHC.

A correcta classificação dos subtipos 1a e 1b também adquiriu uma nova importância, após ter sido relatado que a infecção com o subtipo 1b pode estar associado a uma progressão mais agressiva da doença, e por isso ter impacto na duração e na resposta do doente ao tratamento [59, 116, 117].

Vários estudos demonstraram que a classificação do subtipo 2a e do subtipo 2c pode também não ser conseguida quando a análise é exclusivamente baseada na região 5'UTR. Deste modo, esta região parece não ser suficientemente heterogénea para possibilitar a discriminação dos diferentes subtipos do genótipo 2 [112, 118, 119]. No entanto, mesmo com a informação adicional da região do *core*, o ensaio LiPA parece não apresentar um bom desempenho na discriminação dos subtipos 2a e 2c. Dois estudos revelaram proporções de 50% e 42,1% de resultados indefinidos 2a/2c, e uma elevada proporção, de 30% e 21,1%, de casos de genótipo 2 que não foram subtipados [97, 111]. No presente estudo o método LiPA não identificou e/ou não discriminou o subtipo em respectivamente 50% (6/12) e 25% (3/12) dos casos. Contudo, no método de sequenciação implementado apenas não foi possível determinar o subtipo de VHC em 3 casos (C4, C7 e C8), salientando que em 2 deles (C7 e C8) apenas estava disponível a informação molecular para uma das 2 regiões genómicas estudadas.

Diversos estudos mostraram que o método LiPA não apresenta problemas para a identificação do genótipo 3, bem como para a classificação da maioria dos subtipos dentro deste grupo de vírus. Porém, contrariamente, existe um estudo publicado em que é descrita uma proporção de 13,3% de casos em que não foi possível a determinação do subtipo de VHC [97, 111]. No presente estudo, apenas 1 caso (10%, 1/10) não foi subtipado pelo método LiPA, embora tenha sido classificado em subtipo 3a pelo novo método implementado.

Para o genótipo 4, a região 5'UTR parece não ser suficientemente heterogénea para a classificação dos subtipos virais. Adicionalmente, a informação complementar da região do *core* parece também não proporcionar um melhor desempenho do método LiPA na determinação do subtipo viral em causa. No presente estudo, nenhum dos 19 casos foi subtipado pelo método LiPA. Analisando apenas o grupo de 14 casos

com informação molecular disponível para ambas as regiões C/E1 e NS5B, observou-se a determinação do subtipo de VHC em todos eles. No entanto, é de referir que se identificou 2 casos (C26 e C33) como podendo tratar-se de infecções mistas (e/ou vírus recombinantes), atendendo a que foi obtida uma classificação diferente para cada uma das regiões genómicas analisadas. No entanto, a proporção muito elevada de resultados incompletos ou de subtipos não discriminados obtidos pelo método LiPA, em casos de infecções com o genótipo 4 está em acordo com o descrito na literatura [97].

A recombinação desempenha um importante papel na evolução dos vírus pela criação de variação genética e produção de vírus com novas características [120]. Embora a recombinação não ocorra frequentemente, vários eventos de recombinação inter e intragenómica do VHC têm sido identificados em diferentes partes do mundo [42-50]. No entanto, acredita-se que a maioria das sequências geneticamente recombinadas pode resultar de eventos isolados cujas variantes posteriormente acabam por não circular na população infectada. De facto, ainda que os recombinantes não sejam necessariamente mais virulentos, eles podem contribuir para a evolução do VHC e ter implicações na compreensão da biologia do vírus, no diagnóstico e tratamento no contexto do controlo da infecção pelo VHC no indivíduo. A análise de mais do que uma região subgenómica torna-se necessária para não limitar a identificação de sequências virais recombinadas [42].

A primeira forma recombinada de VHC foi identificada por Kalinina em 2002 num indivíduo infectado de S. Petersburgo. A forma viral foi classificada de 2k/1b e parecia estar em circulação na população de UDI na Rússia [42]. Mais tarde, foi observado em novos isolados, formas denominadas de RF1_2k/1b que se estavam a propagar pela Europa, tendo já sido encontradas na Irlanda, Estónia, Uzbequistão, Sibéria e França [43, 121-124]. A sequenciação de um fragmento amplificado da região *core*-NS2 permitiu localizar o ponto de recombinação na região NS2 (posição 3175 ou 3176 do VHC, de acordo com o sistema de numeração para a sequência pj6CF). Verificou-se que a sequência era constituída por um primeiro segmento genómico classificado como subtipo 2k, e um segundo segmento mais curto classificado como subtipo 1b [42].

No presente estudo, as árvores filogenéticas construídas para as regiões C/E1 e NS5B revelaram que o caso C1 agrupava com sequências de referência classificadas de 1b/2k retiradas da base de dados de Los Alamos de VHC. No entanto, para a região C/E1, o conjunto das sequências estava incluído no agrupamento

correspondente ao genótipo 2, e para a região NS5B estava incluído no agrupamento correspondente ao genótipo 1. Desta forma, colocou-se a hipótese de estar em presença de uma forma de vírus potencialmente recombinada. Assim sendo, procedeu-se à análise por *bootscannig* que não indicou a ocorrência de qualquer evento de recombinação em ambas as regiões do VHC analisadas (dados não apresentados).

Embora as sequências de referência 1b/2k utilizadas neste estudo surjam na base de dados de *Los Alamos* classificadas como genótipo 1b/2k, na verdade essas sequências foram classificadas de 2k/1b pelos autores que as identificaram, uma vez que são constituídas por um primeiro fragmento pertencente ao subtipo 2k e um segundo fragmento classificado como subtipo 1b [43, 123, 125]. Existe assim, a possibilidade de o caso C1 analisado ser um putativo recombinante, pois o único local de recombinação identificado na forma recombinada 2k/1b localiza-se na proteína NS2, e neste estudo apenas foram analisadas as regiões C/E1 e NS5B do VHC. Sendo assim, para o caso C1 seria fundamental analisar a região NS2 ou até mesmo a totalidade do genoma para apurar se de facto se trata ou não da forma viral recombinada 2k/1b já descrita.

Por outro lado, as sequências C4, C7 e C8, embora observadas no agrupamento correspondente a sequências de referência classificadas em genótipo 2, a sua localização na árvore filogenética (C/E1 e/ou NS5B) não permitia determinar qual o subtipo viral a que pertenciam. Assim, no sentido de aprofundar um pouco mais o conhecimento molecular das sequências, procedeu-se à análise por *bootscannig*.

Para o caso da sequência C4 verificou-se a ocorrência de dois potenciais eventos de recombinação, um em cada uma das regiões genómicas analisadas, respectivamente, 2q/2k para a região C/E1 e 2k/2a para a região NS5B. No caso das sequências C7 e C8, infelizmente apenas foi possível analisar a informação molecular de uma das regiões genómicas estudadas. Para o caso da sequência C7, a análise de *bootscannig* identificou um possível evento de recombinação entre os subtipos virais 2q e 2c na região C/E1. Para o caso da sequência C8 não foi sugerido qualquer potencial evento de recombinação na região NS5B. Contudo, é de salientar que na literatura apenas estão descritos 2 recombinantes intragenótipo naturais designados de 1a/1b e de 1a/1c [47, 48]. De facto, para uma melhor compreensão dos possíveis acontecimentos genómicos ocorridos nos 3 casos incluídos no presente estudo, seria fundamental a análise do genoma completo com identificação dos eventuais pontos de recombinação genética.

Vários estudos têm relatado infecções de VHC mistas em hemofílicos, UDI e doentes em hemodiálise [126-128]. A detecção laboratorial de uma infecção mista é complicada, pois geralmente os ensaios acabam por identificar o genótipo de VHC dominante na infecção. Consequentemente, os genótipos presentes em frequências mais baixas podem não ser detectados [127, 129]. Recentemente, num estudo em que foi analisada uma população de hemodializados, constituída por 1292 indivíduos, foi detectada infecção mista em 22 (1,7%) casos com a presença de mais do que um genótipo de VHC [130]. Num estudo de avaliação multicêntrico, apenas 2 entre 10 laboratórios participantes foram capazes de detectar eficazmente casos de infecção mista com base na sequenciação da região NS5B do VHC [113]. O facto do método de genotipagem poder não ser suficientemente sensível para a detecção de infecções mistas pode conduzir à noção errada de uma baixa frequência destas infecções [113, 114]. De facto, em populações de risco para a infecção VHC, nomeadamente em UDI, a frequência de infecções mistas é sugerida ser maior do que as que se encontram descritas. A evidência científica constatada de uma maior prevalência de infecções mistas em doentes com cirrose e CHC torna fundamental a utilização de métodos suficientemente sensíveis para a sua detecção, e o respectivo ajuste na administração do tratamento em acordo com a identificação dos diferentes genótipos de VHC presentes na infecção [37].

No presente estudo foram identificados 2 casos (C26 e C33) em que não foi possível eliminar a hipótese de poderem constituir infecções mistas. No entanto, não existem dados epidemiológicos disponíveis que corroborem múltiplos contactos com vírus da hepatite C ao longo do tempo. Estes acontecimentos são frequentemente descritos em UDI com hábitos de partilha de seringas, e na verdade, não pode deixar de se ter em consideração que este tipo de comportamentos de risco pode estar associado a uma grande parte dos indivíduos em ambiente prisional. Por outro lado, e tal como já foi dito anteriormente, para estes mesmos 2 casos (C26 e C33), também não se pode descartar, a hipótese de serem infecções com variantes de VHC. De facto, o caso C26 foi classificado de subtipos 1b e 4a e o caso C33 de subtipos 3a e 4a, respectivamente para cada um deles atendendo às regiões C/E1 e NS5B. Embora a análise por *bootscanning* não tenha revelado eventos de recombinação nas regiões C/E1 e NS5B (dados não apresentados), a recombinação genética pode eventualmente ter ocorrido em outras regiões do genoma que não foram alvo da presente análise. Assim, para um melhor conhecimento das formas virais presentes nestes 2 casos, e a confirmação de estarmos ou não na presença de recombinantes de VHC, seria necessário analisar a totalidade do genoma recorrendo a técnicas de

clonagem e sequenciação dos vírus, e assim obter a informação molecular necessária para comprovar as hipóteses anteriormente formuladas.

A população de amostras clínicas estudada foi caracterizada por corresponder a indivíduos infectados por VHC com idades compreendidas entre os 22 e os 64 anos. Embora na literatura se descreva uma tendência para os casos de infecção pelo subtipo 1b estar associada a indivíduos de idade mais avançada, não foram encontradas diferenças significativas quando se relacionou a idade do indivíduo com o resultado de genotipagem do VHC no presente estudo [33].

No presente estudo foi obtida para o grupo de indivíduos infectados com o genótipo 4, a mediana na quantificação do RNA de VHC mais baixa (264 536 UI/mL), e contrariamente, para o grupo de indivíduos infectados com o genótipo 1 a mediana na quantificação de RNA VHC mais elevada (8 500 000 UI/mL). Na verdade, segundo a literatura casos de infecções com genótipo 1 apresentam pior prognóstico e uma menor eficácia em alcançar uma RVS. No entanto, a ausência de informação clínica relativamente ao tratamento dos casos em estudo e o desconhecimento se este teria sido já iniciado ou terminado, dificulta a correcta interpretação dos valores de quantificação de RNA VHC obtidos e a associação entre o prognóstico da infecção com a classificação dos genótipos de VHC.

O ensaio comercial LiPA é vulgarmente utilizado na esmagadora maioria dos laboratórios do país, e também em muitos estudos de prevalência de genótipos de VHC que infectam grupos populacionais específicos. É um teste validado para a prestação de serviços à comunidade, com uma abordagem técnica simples e um tempo de execução relativamente rápido. Contudo, os custos imputados pelo seu uso no Laboratório Nacional de Referência IST – HIV e Hepatites B e C, parecem ser demasiado elevados (15 euros/amostra para amplificação e 90 euros/amostra para detecção dos genótipos do VHC).

Para o método desenvolvido de amplificação e sequenciação (para uma das regiões genómicas do VHC, C/E1 ou NS5B) foi estimado um custo de 6 euros para a síntese de cDNA (RT-PCR), 2,5 euros para as 2 amplificações (*heminested* PCR) utilizando o Sistema *illustra™ puReaq Ready-To-Go PCR Beads* (GE Healthcare,

England) e 19 euros para de síntese de *primers* e para a sequenciação dos fragmentos amplificados (cadeia normal e complementar). De facto, excluindo à estimativa de custos, os consumíveis, o desgaste de equipamento, os recursos humanos e a extracção do RNA do VHC, que são factores em comum entre os métodos, a nova metodologia desenvolvida apresenta um custo 3,7 vezes inferior comparativamente ao método comercial LiPA. Pelo seu elevado desempenho, o método implementado pode possibilitar a correlação clínica de um modo mais preciso a diferentes subtipos de VHC e deste modo viabilizar um melhor entendimento do papel da variabilidade genética, na história natural e progressão da infecção do VHC.

Através do método implementado de amplificação e sequenciação foi possível obter um elevado sucesso (69 em 72 casos) na classificação do genótipo e do subtipo de VHC tendo por base, ou a região C/E1, ou a região NS5B, ou ambas as regiões genómicas do VHC. Por outro lado, os resultados obtidos possibilitaram a identificação de alguns casos de interesse epidemiológico e clínico. Assim, de futuro, seria importante aprofundar o conhecimento molecular dos casos C1, C4, C7 e C8 (potencialmente recombinantes), e também dos casos C26 e C33 (recombinantes ou infecções mistas) através da clonagem e análise do genoma completo do VHC.

6. CONCLUSÕES

6. CONCLUSÕES

No presente estudo, o método desenvolvido que incluiu amplificação e sequenciação das regiões C/E1 e NS5B com posterior análise filogenética das sequências de VHC possibilitou:

1. A classificação do genótipo e a discriminação do subtipo do VHC com maior clareza e rigor em comparação com o método comercial LiPA, permitindo reduzir acentuadamente a indefinição na classificação de subtipos de 58,5% (38/72) para 4,2% (3/72).
2. Mostrar, relativamente quanto à classificação dos genótipos do VHC, uma concordância de 95,6% (65/68) entre a análise de resultados da região C/E1 e LiPA, e de 98,5% (67/68) da região NS5B e LiPA.
3. Identificar casos de eventuais vírus recombinantes (ou infecções mistas) que não foram detectados pelo método LiPA.
4. Reduzir acentuadamente os custos inerentes à genotipagem do VHC em pelo menos 3x, e simultaneamente obter uma melhoria na qualidade de resultados.
5. Desenvolver as condições para que o laboratório possa disponibilizar serviço de apoio concreto, para a classificação de casos de infecção de VHC com resultados incompletos (sem subtipo) ou indefinidos (sem discriminação do subtipo) obtidos pelo método LiPA, e para o estudo de casos de interesse clínico.
6. A aplicação prática do método desenvolvido a futuros estudos de epidemiologia molecular do VHC alargados no tempo para um melhor conhecimento da variabilidade genómica, história natural e progressão da infecção do VHC.

7. BIBLIOGRAFIA

7. BIBLIOGRAFIA

1. Choo Q, Kuo G, Weiner A, Overby L, Bradley D, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 1989,**244**:359-362.
2. Kuiken C, Simmonds P. Nomenclature and numbering of the hepatitis C virus. *Methods Mol Biol* 2009,**510**:33-53.
3. Chevaliez S, Pawlotsky J. HCV Genome and Life Cycle. In: *Hepatitis C Viruses - Genomes and Molecular Biology*. Edited by Tan S. Norfolk (UK): Horizon Bioscience; 2006.
4. Moradpour D, Penin F, Rice C. Replication of hepatitis C virus. *Nat Rev Microbiol* 2007,**5**:453-463.
5. Sharma S. Hepatitis C virus: molecular biology & current therapeutic options. *Indian J Med Res* 2010,**131**:17-34.
6. Thurner C, Witwer C, Hofacker I, Stadler P. Conserved RNA secondary structures in Flaviviridae genomes. *J Gen Virol* 2004,**85**:1113-1124.
7. NCBI. Hepatitis C virus genotype 1, complete genome. In. http://www.ncbi.nlm.nih.gov/nuccore/NC_004102, acedido a 12/07/2011.
8. Otto G, Puglisi J. The pathway of HCV IRES-mediated translation initiation. *Cell* 2004,**119**:369-380.
9. Bartschlager R, Frese M, Pietschmann T. Novel insights into hepatitis C virus replication and persistence. *Adv Virus Res* 2004,**63**:71-180.
10. Rehmann B. Hepatitis C virus versus innate and adaptive immune responses: a tale of coevolution and coexistence. *J Clin Invest* 2009,**119**:1745-1754.
11. Shesheer Kumar M, Venkateswara Rao K, Mohammed Habeebullah C, Dashavantha Reddy V. Expression of alternate reading frame protein (F1) of hepatitis C virus in *Escherichia coli* and detection of antibodies for F1 in Indian patients. *Infect Genet Evol* 2008,**8**:374-377.
12. Trimoulet P, Bernard P, de Ledinghen V, Oui B, Chene G, Saint-Marc Girardin M, *et al.* Quantitation of hepatitis C virus RNA in plasma and peripheral blood mononuclear cells of patients with chronic hepatitis treated with interferon-alpha. *Dig Dis Sci* 2000,**45**:175-181.
13. Laskus T, Radkowski M, Wang LF, Nowicki M, Rakela J. Uneven distribution of hepatitis C virus quasispecies in tissues from subjects with end-stage liver disease: confounding effect of viral adsorption and mounting evidence for the presence of low-level extrahepatic replication. *J Virol* 2000,**74**:1014-1017.
14. Navas S, Martín J, Quiroga JA, Castillo I, Carreño V. Genetic diversity and tissue compartmentalization of the hepatitis C virus genome in blood mononuclear cells, liver, and serum from chronic hepatitis C patients. *J Virol* 1998,**72**:1640-1646.
15. Cavalheiro NP, Filgueiras T, Melo C, Morimitsu S, de Araújo E, Tengan F, *et al.* Detection of HCV by PCR in serum and PBMC of patients with hepatitis C after treatment. *Braz J Infect Dis* 2007,**11**:471-474.
16. Pileri P, Uematsu Y, Campagnoli S, Galli G, Falugi F, Petracca R, *et al.* Binding of hepatitis C virus to CD81. *Science* 1998,**282**:938-941.
17. Scarselli E, Ansuini H, Cerino R, Roccasecca R, Acali S, Filocamo G, *et al.* The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J* 2002,**21**:5017-5025.
18. Evans M, von Hahn T, Tscherne D, Syder A, Panis M, Wölk B, *et al.* Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature* 2007,**446**:801-805.
19. Sabahi A. Hepatitis C Virus entry: the early steps in the viral replication cycle. *Virol J* 2009,**6**:117.
20. Lavillette D, Bartosch B, Nourrisson D, Verney G, Cosset FL, Penin F, *et al.* Hepatitis C virus glycoproteins mediate low pH-dependent membrane fusion with liposomes. *J Biol Chem* 2006,**281**:3909-3917.
21. Lukavsky P, Otto G, Lancaster A, Sarnow P, Puglisi J. Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function. *Nat Struct Biol* 2000,**7**:1105-1110.

22. Spahn C, Kieft J, Grassucci R, Penczek P, Zhou K, Doudna J, *et al.* Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science* 2001,**291**:1959-1962.
23. Otto GA, Lukavsky PJ, Lancaster AM, Sarnow P, Puglisi JD. Ribosomal proteins mediate the hepatitis C virus IRES-HeLa 40S interaction. *RNA* 2002,**8**:913-923.
24. Varaklioti A, Vassilaki N, Georgopoulou U, Mavromara P. Alternate translation occurs within the core coding region of the hepatitis C viral genome. *J Biol Chem* 2002,**277**:17713-17721.
25. Egger D, Wölk B, Gosert R, Bianchi L, Blum H, Moradpour D, *et al.* Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J Virol* 2002,**76**:5974-5984.
26. André P, Perlemuter G, Budkowska A, Bréchet C, Lotteau V. Hepatitis C virus particles and lipoprotein metabolism. *Semin Liver Dis* 2005,**25**:93-104.
27. Suzuki T, Ishii K, Aizaki H, Wakita T. Hepatitis C viral life cycle. *Adv Drug Deliv Rev* 2007,**59**:1200-1212.
28. Neumann A, Lam N, Dahari H, Gretch D, Wiley T, Layden T, *et al.* Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science* 1998,**282**:103-107.
29. Bukh J, Miller R, Purcell R. Biology and genetic heterogeneity of hepatitis C virus. *Clin Exp Rheumatol* 1995,**13 Suppl 13**:S3-7.
30. Domingo E, Escarmís C, Sevilla N, Moya A, Elena SF, Quer J, *et al.* Basic concepts in RNA virus evolution. *FASEB J* 1996,**10**:859-864.
31. Robertson B, Myers G, Howard C, Brettin T, Bukh J, Gaschen B, *et al.* Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization. International Committee on Virus Taxonomy. *Arch Virol* 1998,**143**:2493-2503.
32. Simmonds P, Holmes E, Cha T, Chan S, McOmish F, Irvine B, *et al.* Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *J Gen Virol* 1993,**74 (Pt 11)**:2391-2399.
33. Simmonds P. Genetic diversity and evolution of hepatitis C virus--15 years on. *J Gen Virol* 2004,**85**:3173-3188.
34. Kumar U, Brown J, Monjardino J, Thomas H. Sequence variation in the large envelope glycoprotein (E2/NS1) of hepatitis C virus during chronic infection. *J Infect Dis* 1993,**167**:726-730.
35. Simmonds P, Bukh J, Combet C, Deléage G, Enomoto N, Feinstone S, *et al.* Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology* 2005,**42**:962-973.
36. HCV_Sequence_Database. Old and new HCV genotypes in the database. In. <http://hcv.lanl.gov/content/sequence/HCV/classification/genotable.html>, acedido a 04/11/2010.
37. Murphy DG, Willems B, Deschênes M, Hilzenrat N, Mousseau R, Sabbah S. Use of sequence analysis of the NS5B region for routine genotyping of hepatitis C virus with reference to C/E1 and 5' untranslated region sequences. *J Clin Microbiol* 2007,**45**:1102-1112.
38. Kuiken C, Hraber P, Thurmond J, Yusim K. The hepatitis C sequence database in Los Alamos. *Nucleic Acids Res* 2008,**36**:D512-516.
39. HCV_Sequence_Database. Distribution of HCV sequences: World. In. http://hcv.lanl.gov/components/sequence/HCV/geo/geo.comp?organism=HCV®ion=world®ion=&form=all&show=Show&prev_region=eur, acedido a 15/08/2011.
40. HCV_Sequence_Database. Distribution of HCV sequences: Europe. In. http://hcv.lanl.gov/components/sequence/HCV/geo/geo.comp?organism=HCV®ion=eur®ion=&form=all&show=Show&prev_region=world, acedido a 15/08/2011.
41. HCV_Sequence_Database. Distribution of HCV sequences: Portugal. In. http://hcv.lanl.gov/components/sequence/HCV/geo/geo.comp?organism=HCV®ion=eur®ion=pt&form=all&show=Show&prev_region=eur, acedido a 15/08/2011.
42. Kalinina O, Norder H, Mukomolov S, Magnius LO. A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg. *J Virol* 2002,**76**:4034-4043.
43. Moreau I, Hegarty S, Levis J, Sheehy P, Crosbie O, Kenny-Walsh E, *et al.* Serendipitous identification of natural intergenotypic recombinants of hepatitis C in Ireland. *Virology* 2006,**3**:95.

44. Kageyama S, Agdamag DM, Alesna ET, Leaño PS, Heredia AM, Abellanos-Tac-An IP, *et al.* A natural inter-genotypic (2b/1b) recombinant of hepatitis C virus in the Philippines. *J Med Virol* 2006,**78**:1423-1428.
45. Legrand-Abravanel F, Claudinon J, Nicot F, Dubois M, Chapuy-Regaud S, Sandres-Saune K, *et al.* New natural intergenotypic (2/5) recombinant of hepatitis C virus. *J Virol* 2007,**81**:4357-4362.
46. Noppornpanth S, Lien TX, Poovorawan Y, Smits SL, Osterhaus AD, Haagmans BL. Identification of a naturally occurring recombinant genotype 2/6 hepatitis C virus. *J Virol* 2006,**80**:7569-7577.
47. Colina R, Casane D, Vasquez S, García-Aguirre L, Chunga A, Romero H, *et al.* Evidence of intratypic recombination in natural populations of hepatitis C virus. *J Gen Virol* 2004,**85**:31-37.
48. Ross RS, Verbeeck J, Viazov S, Lemey P, Van Ranst M, Roggendorf M. Evidence for a complex mosaic genome pattern in a full-length hepatitis C virus sequence. *Evol Bioinform Online* 2008,**4**:249-254.
49. Moreno MP, Casane D, López L, Cristina J. Evidence of recombination in quasispecies populations of a Hepatitis C Virus patient undergoing anti-viral therapy. *Virology* 2006,**3**:87.
50. Sentandreu V, Jiménez-Hernández N, Torres-Puente M, Bracho MA, Valero A, Gosalbes MJ, *et al.* Evidence of recombination in inpatient populations of hepatitis C virus. *PLoS One* 2008,**3**:e3239.
51. Lavanchy D. The global burden of hepatitis C. *Liver Int* 2009,**29 Suppl 1**:74-81.
52. Hagan H, Snyder N, Hough E, Yu T, McKeirnan S, Boase J, *et al.* Case-reporting of acute hepatitis B and C among injection drug users. *J Urban Health* 2002,**79**:579-585.
53. Alter M, Kruszon-Moran D, Nainan O, McQuillan G, Gao F, Moyer L, *et al.* The prevalence of hepatitis C virus infection in the United States, 1988 through 1994. *N Engl J Med* 1999,**341**:556-562.
54. Shepard C, Finelli L, Alter M. Global epidemiology of hepatitis C virus infection. *Lancet Infect Dis* 2005,**5**:558-567.
55. National Travel Health Network and Centre. Hepatitis C. In. http://www.nathnac.org/travel/factsheets/hep_c.htm; 2008, acedido a 08/01/2011.
56. Marinho R, Moura M, Gira J, Ferrinho P. Epidemiological aspects of hepatitis C in Portugal. *J Gastroenterol Hepatol* 2001,**16**:1076-1077.
57. Instituto Nacional de Estatística. Indicadores Sociais - 2009. In. http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=105186688&PUBLICACOESmodo=2; 2009, acedido a 13/01/2011.
58. Roche. O que é a Hepatite C? In. <http://www.roche.pt/hepatites/hepatitec/>, acedido a 11/01/2011.
59. Pawlotsky J, Tsakiris L, Roudot-Thoraval F, Pellet C, Stuyver L, Duval J, *et al.* Relationship between hepatitis C virus genotypes and sources of infection in patients with chronic hepatitis C. *J Infect Dis* 1995,**171**:1607-1610.
60. McOmish F, Yap P, Dow B, Follett E, Seed C, Keller A, *et al.* Geographical distribution of hepatitis C virus genotypes in blood donors: an international collaborative survey. *J Clin Microbiol* 1994,**32**:884-892.
61. Nguyen M, Keffe E. Prevalence and treatment of hepatitis C virus genotypes 4, 5, and 6. *Clin Gastroenterol Hepatol* 2005,**3**:S97-S101.
62. Esteban J, Sauleda S, Quer J. The changing epidemiology of hepatitis C virus infection in Europe. *J Hepatol* 2008,**48**:148-162.
63. Smuts H, Kannemeyer J. Genotyping of hepatitis C virus in South Africa. *J Clin Microbiol* 1995,**33**:1679-1681.
64. Wong D, Tong L, Lim W. High prevalence of hepatitis C virus genotype 6 among certain risk groups in Hong Kong. *Eur J Epidemiol* 1998,**14**:421-426.
65. WHO. Hepatitis C Virus. In. http://www.who.int/vaccine_research/diseases/viral_cancers/en/index2.html#disease_burden; 2010, acedido a 10/01/2011.
66. Vieira A, Freire R, Mangualde J, Pinho C, Fernandes V, Alves A, *et al.* Hepatite C - casuística da consulta de hepatologia de um hospital distrital. *J Port Gastroenterol* 2007,**14**:134-140.
67. Sêco ABL, Gomes N, Patrício N, Meira T, Marques G, Vacas S. Distribution of hepatitis C virus genotypes in patients from Centro Hospitalar Lisboa Norte (CHLN) between

- December 2009 and April 2011. In. 14th Annual Meeting of the European Society for Clinical Virology; 2011.
68. Wasley A, Alter M. Epidemiology of hepatitis C: geographic differences and temporal trends. *Semin Liver Dis* 2000;**20**:1-16.
 69. Sutton A, Hope V, Mathei C, Mravcik V, Sebakova H, Vallejo F, *et al.* A comparison between the force of infection estimates for blood-borne viruses in injecting drug user populations across the European Union: a modelling study. *J Viral Hepat* 2008;**15**:809-816.
 70. Alter H, Purcell R, Shih J, Melpolder J, Houghton M, Choo Q, *et al.* Detection of antibody to hepatitis C virus in prospectively followed transfusion recipients with acute and chronic non-A, non-B hepatitis. *N Engl J Med* 1989;**321**:1494-1500.
 71. Pomper G, Wu Y, Snyder E. Risks of transfusion-transmitted infections: 2003. *Curr Opin Hematol* 2003;**10**:412-418.
 72. Bortolotti F, Iorio R, Jorio R, Resti M, Cammà C, Marcellini M, *et al.* Epidemiological profile of 806 Italian children with hepatitis C virus infection over a 15-year period. *J Hepatol* 2007;**46**:783-790.
 73. Mitsui T, Iwano K, Masuko K, Yamazaki C, Okamoto H, Tsuda F, *et al.* Hepatitis C virus infection in medical personnel after needlestick accident. *Hepatology* 1992;**16**:1109-1114.
 74. Vandelli C, Renzo F, Romanò L, Tisminetzky S, De Palma M, Stroffolini T, *et al.* Lack of evidence of sexual transmission of hepatitis C among monogamous couples: results of a 10-year prospective follow-up study. *Am J Gastroenterol* 2004;**99**:855-859.
 75. Wasmuth J. Hepatitis C - Epidemiology, transmission and natural history. In: *Hepatology*. Edited by Mauss S, Berg T, Rockstroh J, Sarrazin C, Wedemeyer H: Flying Publisher; 2009. pp. 19-29.
 76. Lauer G, Walker B. Hepatitis C virus infection. *N Engl J Med* 2001;**345**:41-52.
 77. Chevaliez S, Pawlotsky J. Hepatitis C virus: virology, diagnosis and management of antiviral therapy. *World J Gastroenterol* 2007;**13**:2461-2466.
 78. Chung R. Acute hepatitis C virus infection. *Clin Infect Dis* 2005;**41 Suppl 1**:S14-17.
 79. RCM_Pharma. Hepatite B: números têm diminuído em Portugal. In. <http://www.rcmpharma.com/news/10737/51/Hepatite-B-numeros-tem-diminuido-em-Portugal.html>; 2010, acedido a 13/01/2011.
 80. McHutchison J, Gordon S, Schiff E, Shiffman M, Lee W, Rustgi V, *et al.* Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. *N Engl J Med* 1998;**339**:1485-1492.
 81. Manns M, McHutchison J, Gordon S, Rustgi V, Shiffman M, Reindollar R, *et al.* Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet* 2001;**358**:958-965.
 82. Feld J, Hoofnagle J. Mechanism of action of interferon and ribavirin in treatment of hepatitis C. *Nature* 2005;**436**:967-972.
 83. Potter C, Phair J, Vodinelich L, Fenton R, Jennings R. Antiviral, immunosuppressive and antitumour effects of ribavirin. *Nature* 1976;**259**:496-497.
 84. Fried M, Shiffman M, Reddy K, Smith C, Marinos G, Gonçalves FJ, *et al.* Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med* 2002;**347**:975-982.
 85. Welzel T, Morgan T, Bonkovsky H, Naishadham D, Pfeiffer R, Wright E, *et al.* Variants in interferon-alpha pathway genes and response to pegylated interferon-Alpha2a plus ribavirin for treatment of chronic hepatitis C virus infection in the hepatitis C antiviral long-term treatment against cirrhosis trial. *Hepatology* 2009;**49**:1847-1858.
 86. McHutchison J. Understanding hepatitis C. *Am J Manag Care* 2004;**10**:S21-29.
 87. Antaki N, Craxi A, Kamal S, Moucari R, Van der Merwe S, Haffar S, *et al.* The neglected hepatitis C virus genotypes 4, 5 and 6: an international consensus report. *Liver Int* 2010;**30**:342-355.
 88. Zeuzem S, Andreone P, Pol S, Lawitz E, Diago M, Roberts S, *et al.* Telaprevir for retreatment of HCV infection. *N Engl J Med* 2011;**364**:2417-2428.
 89. Scott J, Gretch D. Molecular diagnostics of hepatitis C virus infection: a systematic review. *JAMA* 2007;**297**:724-732.
 90. Lemon M, Alter M, Walker CY, M. *Fields - Virology*. 5^a ed. USA: Lippincott Williams & Wilkins; 2007.

91. Terrault N, Pawlotsky J, McHutchison J, Anderson F, Kraiden M, Gordon S, *et al.* Clinical utility of viral load measurements in individuals with chronic hepatitis C infection on antiviral therapy. *J Viral Hepat* 2005,**12**:465-472.
92. Lange C, Sarrazin C. Acute and chronic hepatitis C - Diagnostic tests. In: *Hepatology*. Edited by Mauss S, Berg T, Rockstroh J, Sarrazin C, Wedemeyer H: Flying Publisher; 2009. pp. 171-182.
93. Bowden D, Berzsenyi M. Chronic hepatitis C virus infection: genotyping and its clinical role. *Future Microbiol* 2006,**1**:103-112.
94. Smith DB, Mellor J, Jarvis LM, Davidson F, Kolberg J, Urdea M, *et al.* Variation of the hepatitis C virus 5' non-coding region: implications for secondary structure, virus detection and typing. The International HCV Collaborative Study Group. *J Gen Virol* 1995,**76 (Pt 7)**:1749-1761.
95. Chen Z, Weck KE. Hepatitis C virus genotyping: interrogation of the 5' untranslated region cannot accurately distinguish genotypes 1a and 1b. *J Clin Microbiol* 2002,**40**:3127-3134.
96. Chinchai T, Labout J, Noppornpanth S, Theamboonlers A, Haagmans BL, Osterhaus AD, *et al.* Comparative study of different methods to genotype hepatitis C virus type 6 variants. *J Virol Methods* 2003,**109**:195-201.
97. Verbeeck J, Stanley MJ, Shieh J, Celis L, Huyck E, Wollants E, *et al.* Evaluation of Versant hepatitis C virus genotype assay (LiPA) 2.0. *J Clin Microbiol* 2008,**46**:1901-1906.
98. Bukh J, Miller RH, Purcell RH. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin Liver Dis* 1995,**15**:41-63.
99. Calado RA, Rocha MR, Parreira R, Piedade J, Venenno T, Esteves A. Hepatitis C virus subtypes circulating among intravenous drug users in Lisbon, Portugal. *J Med Virol* 2011,**83**:608-615.
100. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* 2007,**23**:2947-2948.
101. Hall T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 1999,**41**:95-98.
102. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007,**24**:1596-1599.
103. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 1980,**16**:111-120.
104. Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 1985,**39**:783-791.
105. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, *et al.* Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 1999,**73**:152-160.
106. Thompson WD, Walter SD. Kappa and the concept of independent errors. *J Clin Epidemiol* 1988,**41**:969-970.
107. Lee JH, Roth WK, Zeuzem S. Evaluation and comparison of different hepatitis C virus genotyping and serotyping assays. *J Hepatol* 1997,**26**:1001-1009.
108. Stuyver L, Wyseur A, van Arnhem W, Hernandez F, Maertens G. Second-generation line probe assay for hepatitis C virus genotyping. *J Clin Microbiol* 1996,**34**:2259-2266.
109. Weck K. Molecular methods of hepatitis C genotyping. *Expert Rev Mol Diagn* 2005,**5**:507-520.
110. Dev AT, McCaw R, Sundararajan V, Bowden S, Sievert W. Southeast Asian patients with chronic hepatitis C: the impact of novel genotypes and race on treatment outcome. *Hepatology* 2002,**36**:1259-1265.
111. Bouchardeau F, Cantaloube J, Chevaliez S, Portal C, Razer A, Lefrère J, *et al.* Improvement of hepatitis C virus (HCV) genotype determination with the new version of the INNO-LiPA HCV assay. *J Clin Microbiol* 2007,**45**:1140-1145.
112. Cantaloube JF, Laperche S, Gallian P, Bouchardeau F, de Lamballerie X, de Micco P. Analysis of the 5' noncoding region versus the NS5b region in genotyping hepatitis C virus isolates from blood donors in France. *J Clin Microbiol* 2006,**44**:2051-2056.
113. Laperche S, Lunel F, Izopet J, Alain S, Dény P, Duverlie G, *et al.* Comparison of hepatitis C virus NS5b and 5' noncoding gene sequencing methods in a multicenter study. *J Clin Microbiol* 2005,**43**:733-739.

114. Nolte FS, Green AM, Fiebelkorn KR, Caliendo AM, Sturchio C, Grunwald A, *et al.* Clinical evaluation of two methods for genotyping hepatitis C virus based on analysis of the 5' noncoding region. *J Clin Microbiol* 2003,**41**:1558-1564.
115. Corbet S, Bukh J, Heinsen A, Fomsgaard A. Hepatitis C virus subtyping by a core-envelope 1-based reverse transcriptase PCR assay with sequencing and its use in determining subtype distribution among Danish patients. *J Clin Microbiol* 2003,**41**:1091-1100.
116. Amoroso P, Rapicetta M, Tosti ME, Mele A, Spada E, Buonocore S, *et al.* Correlation between virus genotype and chronicity rate in acute hepatitis C. *J Hepatol* 1998,**28**:939-944.
117. Zein NN, Rakela J, Krawitt EL, Reddy KR, Tominaga T, Persing DH. Hepatitis C virus genotypes in the United States: epidemiology, pathogenicity, and response to interferon therapy. Collaborative Study Group. *Ann Intern Med* 1996,**125**:634-639.
118. Davidson F, Simmonds P, Ferguson JC, Jarvis LM, Dow BC, Follett EA, *et al.* Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' non-coding region. *J Gen Virol* 1995,**76 (Pt 5)**:1197-1204.
119. Simmonds P, Mellor J, Sakuldamrongpanich T, Nuchaprayoon C, Tanprasert S, Holmes EC, *et al.* Evolutionary analysis of variants of hepatitis C virus found in South-East Asia: comparison with classifications based upon sequence similarity. *J Gen Virol* 1996,**77 (Pt 12)**:3013-3024.
120. Worobey M, Holmes EC. Evolutionary aspects of recombination in RNA viruses. *J Gen Virol* 1999,**80 (Pt 10)**:2535-2543.
121. Tallo T, Norder H, Tefanova V, Krispin T, Schmidt J, Ilmoja M, *et al.* Genetic characterization of hepatitis C virus strains in Estonia: fluctuations in the predominating subtype with time. *J Med Virol* 2007,**79**:374-382.
122. Kurbanov F, Tanaka Y, Avazova D, Khan A, Sugauchi F, Kan N, *et al.* Detection of hepatitis C virus natural recombinant RF1_2k/1b strain among intravenous drug users in Uzbekistan. *Hepatol Res* 2008,**38**:457-464.
123. Kurbanov F, Tanaka Y, Chub E, Maruyama I, Azlarova A, Kamitsukasa H, *et al.* Molecular epidemiology and interferon susceptibility of the natural recombinant hepatitis C virus strain RF1_2k/1b. *J Infect Dis* 2008,**198**:1448-1456.
124. Morel V, Descamps V, François C, Fournier C, Brochot E, Capron D, *et al.* Emergence of a genomic variant of the recombinant 2k/1b strain during a mixed Hepatitis C infection: a case report. *J Clin Virol* 2010,**47**:382-386.
125. Kalinina O, Norder H, Magnius LO. Full-length open reading frame of a recombinant hepatitis C virus strain from St Petersburg: proposed mechanism for its formation. *J Gen Virol* 2004,**85**:1853-1857.
126. Aitken C, McCaw R, Jardine D, Bowden S, Higgs P, Nguyen O, *et al.* Change in hepatitis C virus genotype in injecting drug users. *J Med Virol* 2004,**74**:543-545.
127. Tuveri R, Rothschild C, Pol S, Reijasse D, Persico T, Gazengel C, *et al.* Hepatitis C virus genotypes in French haemophiliacs: kinetics and reappraisal of mixed infections. *J Med Virol* 1997,**51**:36-41.
128. Qian KP, Natov SN, Pereira BJ, Lau JY. Hepatitis C virus mixed genotype infection in patients on haemodialysis. *J Viral Hepat* 2000,**7**:153-160.
129. Mitchell PS, Sloan LM, Majewski DW, Rys PN, Heimgartner PJ, Rosenblatt JE, *et al.* Comparison of line probe assay and DNA sequencing of 5' untranslated region for genotyping hepatitis C virus: description of novel line probe patterns. *Diagn Microbiol Infect Dis* 2002,**42**:175-179.
130. Al Balwi MA. Prevalence of mixed hepatitis C virus (HCV) genotypes among recently diagnosed dialysis patients with HCV infection. *Saudi J Kidney Dis Transpl* 2011,**22**:712-716.

8. ANEXOS

8. ANEXOS

Tabela 8.1 – Dados demográficos (sexo e idade) e resultados laboratoriais (ensaio LiPA, valores de RNA de VHC e classificação de VHC com base nas regiões C/E1 e NS5B) das amostras clínicas analisadas no presente estudo. M – masculino; F – feminino; (-) informação não disponível.

Nº amostra clínica	Sexo	Idade	Resultado LiPA	RNA de VHC (UI/mL)	Resultado C/E1	Resultado NS5B
C1	M	33	2a/2c	9940000	1b/2k	1b/2k
C2	F	-	1b	847	1a	-
C3	M	-	1a	2870	1a	1a
C4	F	-	2	5790000	2	2
C5	M	-	2a/2c	29600	2a	2a
C6	M	24	2b	508000	2b	2b
C7	M	64	2a/2c	1670000	2	-
C8	F	50	2a/2c	853780	-	2
C9	M	38	1a	1050000	1a	1a
C10	M	-	2a/2c	520	2a	2a
C11	M	-	2	50200000	2b	2b
C12	M	37	1a	316000	1a	1a
C13	F	42	1b	45400	1b	1b
C14	M	31	1a	15300000	1a	1a
C15	M	-	2b	23469333	2b	2b
C16	F	44	4	57500	-	4d
C17	M	35	1a	11400000	1a	1a
C18	M	35	3a	4054893	3a	3a
C19	M	22	4a/4c/4d	470	-	4d
C20	M	-	1	26500000	1a	1a
C21	M	-	3a	24200000	3a	3a
C22	M	-	4a/4c/4d	108594	4d	4d
C23	M	-	4a/4c/4d	8692	1a	-
C24	M	32	1	943949	1a	1a
C25	M	51	1a	20700000	1a	1a
C26	M	33	4a/4c/4d	40800	1b	4a
C27	M	40	4a/4c/4d	264536	4a	4a
C28	M	-	1a	10000000	1a	1a
C29	M	-	4a/4c/4d	395773	4d	4d
C30	M	35	1b	2860000	1b	1b
C31	M	33	1	7400000	1a	1a
C32	M	43	4a/4c/4d	653000	4a	4a
C33	M	-	4	5091115	3a	4a
C34	M	39	4a/4c/4d	237000	4a	-
C35	M	36	1	724000	1a	1a
C36	M	49	4	1040000	4a	4a

C37	M	43	4	3910000	4a	4a
C38	M	43	4a/4c/4d	204000	4a	4a
C39	M	-	1	7580000	1a	1a
C40	M	44	3a	13300000	3a	3a
C41	M	-	1b	13600000	1b	1b
C42	M	55	1	8500000	1a	1a
C43	M	-	1a	9590000	1a	1a
C44	M	-	1b	1230000	1b	1b
C45	M	-	1b	28200000	1b	1b
C46	M	32	1a	18600000	1a	1a
C47	M	30	4a/4c/4d	3870000	4a	4a
C48	M	36	4	187000	4d	4d
C49	M	-	4	587000	4a	4a
C50	M	-	1a	10100000	1a	1a
C51	M	39	4a/4c/4d	1670000	4a	4a
C52	M	57	1b	4960000	1b	1b
C53	M	31	1a	21800000	1a	1a
C54	M	43	1b	5130000	1b	1b
C55	M	32	1	15100000	1a	1a
C56	M	-	3a	15800000	3a	3a
C57	M	-	3a	3330000	3a	3a
C58	M	30	4	3650	-	4d
C59	M	-	1b	14800000	1a	1a
C60	M	-	4	1230000	4d	4d
C61	M	42	1	10800000	1a	1a
C62	M	43	3a	1940000	3a	3a
C63	M	-	1b	11900000	1a	1a
C64	M	-	3a	19200000	3a	3a
C65	M	51	2b	19700000	2b	2b
C66	M	41	1a	7740000	1a	1a
C67	M	-	3a	4250000	3a	3a
C68	M	-	2a/2c	3350000	2a	2a
C69	M	40	3a	3240	3a	3a
C70	M	-	1	1710	1a	1a
C71	M	-	3	9720	3a	3a
C72	F	-	2	35000	2a	2a