

**UNIVERSIDADE DE LISBOA**  
**FACULDADE DE CIÊNCIAS**  
**DEPARTAMENTO DE BIOLOGIA VEGETAL**



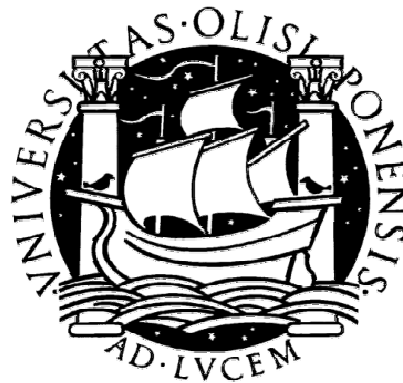
***Prevalência e caracterização genética do Vírus da Imunodeficiência  
Humana em indivíduos atendidos no Hospital de Cumura em  
Bissau***

**Filipa Vaz Sena**

Mestrado em Microbiologia Aplicada

2010

**UNIVERSIDADE DE LISBOA**  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE BIOLOGIA VEGETAL



***Prevalência e caracterização genética do Vírus da Imunodeficiência  
Humana em indivíduos atendidos no Hospital de Cumura em  
Bissau***

**Filipa Vaz Sena**

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

Mestrado em Microbiologia Aplicada

2010

## AGRADECIMENTOS

A todas as pessoas que de alguma forma contribuíram para a realização deste trabalho, gostaria de expressar os meus sinceros agradecimentos, em particular:

- ❖ À Doutora Elizabeth Pádua, pela oportunidade de realizar este estudo sob a sua supervisão, pelo apoio e conhecimentos transmitidos e sobretudo pela confiança e compreensão demonstrada ao longo destes meses.
- ❖ À Professora Doutora Filomena Caeiro por ter aceitado a co-orientação deste projecto a partir da Faculdade de Ciências da Universidade de Lisboa, pela sua disponibilidade, simpatia e pelos conselhos finais que me transmitiu.
- ❖ À Doutora Suzana David, por ter viabilizado a integração do estudo molecular do VIH-1 num projecto alargado à análise da “Variabilidade Genética do *M. tuberculosis* para Identificação de Marcadores Moleculares de Virulência” financiado pela “*Luso-American Collaborative Response Award on Tuberculosis program*” da Fundação Luso Americana.
- ❖ Ao Doutor Vítor Henriques, médico residente do Hospital de Cumura em Bissau, por ter disponibilizado as amostras e toda a informação demográfica e clínica necessária para a realização deste estudo. Ao Dr. Alfredo Mané e Dr. Armando Sanca pela colheita das amostras. Ao Professor Doutor Abílio Antunes, do Instituto de Higiene e Medicina Tropical, pela sua colaboração no projecto com a Guiné-Bissau.
- ❖ À colega Patrícia Francisco pela ajuda preciosa sobre os programas bioinformáticos, pela partilha de conhecimentos e prontidão em ajudar.
- ❖ À Doutora Eleonora Paixão pelo auxílio na realização dos testes estatísticos, assim como pela disponibilidade para responder a todas as minhas dúvidas.
- ❖ Às técnicas do Laboratório da SIDA, Catarina e Ivone, pela ajuda laboratorial e pelo constante sentido de humor.
- ❖ À amiga e colega Filipa Sutre, pela constante entreaajuda, pelas sugestões e pela preocupação demonstrada ao longo destes meses.
- ❖ Ao Pedro, pela paciência e apoio, pela companhia e constante boa disposição, o meu muito obrigado.
- ❖ À minha tia Margarida por me ter ajudado nos pormenores desta tese mas acima de tudo pela amizade de sempre.
- ❖ Aos meus pais e irmão, um agradecimento especial, pelo apoio incondicional em todos os aspectos, por confiarem em mim e no meu trabalho e pela compreensão demonstrada em todas as minhas ausências.

## RESUMO

O Vírus da Imunodeficiência Humana é caracterizado por uma enorme capacidade de recombinação genética com implicações para o diagnóstico, desenvolvimento de fármacos e produção de vacinas. O crescente número de variantes virais em circulação no mundo coloca questões sobre eventuais efeitos biológicos, nomeadamente diferentes taxas de transmissibilidade, virulência ou progressão para a doença.

O presente estudo pretendeu determinar a prevalência da infecção VIH-1 e VIH-2 e da co-infecção com membros do complexo *Mycobacterium tuberculosis* num grupo de indivíduos atendidos no Hospital de Cumura na Guiné-Bissau. Pretendeu também, amplificar por *Nested* PCR fragmentos virais correspondentes à região V3-V5 de *env* e ao gene *nef* de VIH-1 para proceder à classificação dos vírus por inferência filogenética e análise *bootscanning* utilizando diferentes ferramentas informáticas. Foi também realizada uma análise das sequências aminoacídicas de ambas as regiões com o intuito de observar o grau de conservação ou disrupção dos domínios estruturais e funcionais de VIH-1.

Assim, foi detectada uma prevalência de 15,8% para o VIH-1 e de 4,2% para o VIH-2 no grupo estudado, o que está em acordo com o declínio da infecção VIH-2 descrita no país. Não foram detectados subtipos puros e a maioria dos vírus foram classificados como formas genéticas A/G. Foram também identificadas formas complexas A/G com envolvimento de outros subtipos virais, tais como o B, C, J e H, indicando maior diversidade genética relativamente ao que foi descrito na Guiné-Bissau.

A análise da região V3-V5 e da proteína Nef revelou uma conservação dos motivos funcionais, reflectindo a sua importância ao longo da infecção. Foram observados alguns polimorfismos “naturais” que podem estar associados a formas virais não-B, tais como, subtipo G e CRF01\_AE e CRF02\_AG.

Estudos adicionais *in vitro*, juntamente com informação clínica dos indivíduos infectados, podem determinar se as diferenças entre CRF/subtipos virais terão repercussões na progressão da infecção.

**PALAVRAS-CHAVE:** VIH-1, VIH-2, Diversidade genética, Subtipos virais, CRF, *env*, *nef*, Motivos funcionais

## **ABSTRACT**

The Human Immunodeficiency Virus is characterized by an enormous capacity for genetic recombination with implications for diagnosis, drug development and vaccine production. The growing number of new virus variants circulating in the world raises questions about possible biological effects, including different rates of transmissibility, virulence or disease progression.

This study aimed to determine the prevalence of HIV-1 and HIV-2 and co-infection with *Mycobacterium tuberculosis* complex members of a group of individuals treated in the Cumura Hospital in Guinea-Bissau. Also aimed to amplify viral fragments, by *Nested* PCR, corresponding to the V3-V5 region of *env* and the *nef* gene of HIV-1 to classify the viruses by phylogenetic inference and *bootscanning* analysis, using different tools. Was also undertaken an analysis of aminoacidic sequences of both regions in order to observe the degree of conservation or disruption of structural and functional domains of HIV-1.

Thus, we found a prevalence of 15.8% for HIV-1 and 4.2% for HIV-2 in the study group, which is according with the decline of HIV-2 reported in the country. Pure subtypes were not detected and most forms were classified as viral genetic forms A/G. We also identified complex forms A/G with involvement of other subtypes, such as B, C, H and J, indicating a greater genetic diversity in relation to what was described in Guinea-Bissau.

The analysis of the V3-V5 region and Nef protein revealed a conservation of functional motifs, reflecting their importance along the infection. Some "natural" polymorphisms were observed that may be associated with non-B viral forms, often observed in sequences of subtype G and in CRF02\_AG, and CRF01\_AE.

Further *in vitro* studies, together with clinical information of the infected individuals, can determine whether the differences between CRF/subtypes will affect the progression of infection.

**KEY WORDS:** HIV-1, HIV-2, genetic diversity, viral subtypes, CRF, *env*, *nef*, functional motifs

# ÍNDICE GERAL

|  |     |
|--|-----|
| <b>AGRADECIMENTOS</b>  | i   |
| <b>RESUMO</b>  | ii  |
| <b>ABSTRACT</b>  | iii |
| <b>ÍNDICE GERAL</b>  | iv  |
| <b>LISTA DE ABREVIATURAS E SIGLAS</b>  | vi  |
| <br>   |     |
| <b>I. INTRODUÇÃO</b>   | 1   |
| 1. Vírus da Imunodeficiência Humana  | 1   |
| 1.1. Origem  | 1   |
| 1.2. Classificação, Morfologia e Organização Genómica                        | 2   |
| 1.3. Ciclo de Replicação   | 3   |
| 1.4. Infecção e Desenvolvimento da Imunodeficiência                          | 4   |
| 1.5. Variabilidade Genética do VIH   | 5   |
| 1.5.1. Causas e Consequências da Variabilidade                               | 5   |
| 1.5.2. Classificação das Estirpes  | 5   |
| 1.5.3. Epidemiologia molecular no Mundo                                      | 6   |
| 1.5.3.1. Contributos na Transmissão e disseminação da infecção               | 7   |
| 1.5.3.2. Infecção VIH em Portugal e na Guiné-Bissau                          | 8   |
| 1.6. Impacto da Diversidade Genética na Pandemia VIH                         | 9   |
| 1.6.1. Patogenicidade, Transmissibilidade e Progressão para Doença           | 9   |
| 1.6.2. Diagnóstico e Monitorização da Infecção                               | 10  |
| 1.6.3. Mutações de Resistência e Implicações no Desenvolvimento de Vacinas   | 11  |
| 1.7. Co-infecção com membros do complexo <i>Mycobacterium tuberculosis</i>   | 11  |
| 1.8. Domínios Estruturais e Funcionais das Proteínas Env e Nef               | 12  |
| 1.8.1. Glicoproteína do Invólucro Viral gp120                                | 12  |
| 1.8.2. Proteína Nef  | 12  |
| <br>   |     |
| <b>II. OBJECTIVOS</b>  | 14  |
| <br>   |     |
| <b>III. MATERIAL E MÉTODOS</b>   | 15  |
| 1. Análise da População Alvo   | 15  |
| 2. Desenho do estudo   | 15  |
| 3. Preparação das Amostras   | 17  |
| 3.1. Extração de ADN a partir de sangue total imobilizado em papel de filtro | 17  |
| 3.2. Extração de ADN em colunas a partir de PBMC                             | 18  |
| 3.3. Quantificação Espectrofotométrica do material genético                  | 18  |
| 4. Amplificação da Região Genómica LTR de VIH-1 e de VIH-2                   | 18  |
| 5. Amplificação da Região V3-V5 de <i>env</i> e do gene <i>nef</i> de VIH-1  | 19  |
| 6. Purificação e Sequenciação dos Produtos de PCR                            | 21  |

|             |   |           |
|-------------|---|-----------|
| 7.          | Análise de Sequências Nucleotídicas .....                                     | 22        |
| 7.1.        | Obtenção das Sequências Consenso .....  | 22        |
| 7.2.        | Sequências de Referência .....  | 22        |
| 7.3.        | Inferência Filogenética .....   | 22        |
| 7.4.        | Análise da Hipermutação .....   | 23        |
| 7.5.        | Detecção de Recombinação Genética .....                                       | 23        |
| 7.6.        | Cálculo de Distâncias Nucleotídicas .....                                     | 23        |
| 8.          | Análise de Sequências Aminoacídicas .....                                     | 23        |
| 9.          | Tratamento Estatístico .....  | 24        |
| <b>IV.</b>  | <b>RESULTADOS</b> .....   | <b>25</b> |
| 1.          | Detecção de VIH em Extraídos de ADN por Procedimentos Distintos .....         | 25        |
| 2.          | Diagnóstico da Infecção VIH na amostra em estudo .....                        | 25        |
| 3.          | Caracterização dos casos de infecção VIH-1 ou VIH-2 .....                     | 26        |
| 4.          | Amplificação e Sequenciação das regiões genómicas de VIH-1 .....              | 26        |
| 5.          | Análise Molecular da Região V3-V5 de <i>env</i> de VIH-1 .....                | 27        |
| 6.          | Análise Molecular da Região <i>nef</i> de VIH-1 .....                         | 32        |
| 7.          | Análise Molecular Conjunta das Regiões V3-V5 de <i>env</i> e <i>nef</i> ..... | 36        |
| 8.          | Distâncias Genéticas das Sequências Virais .....                              | 38        |
| <b>V.</b>   | <b>DISCUSSÃO</b> .....  | <b>39</b> |
| <b>VI.</b>  | <b>CONCLUSÃO</b> .....  | <b>44</b> |
| <b>VII.</b> | <b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....                                       | <b>45</b> |

## LISTA DE ABREVIATURAS E SIGLAS

- ADN** – Ácido Desoxirribonucleico
- ARN** – Ácido Ribonucleico
- ARNm** – Ácido Ribonucleico mensageiro
- BLAST** – *Basic Local Alignment and Search Tool*
- CA** – Proteína da Cápside
- CD4** – Receptor celular, do inglês *cluster of differentiation type 4*
- CMTB** – Complexo *Mycobacterium tuberculosis*
- CRF** – Formas recombinadas em circulação, do inglês *circulating recombinant form*
- HLA** – Antígeno leucocitário humano, do inglês *human leukocyte antigen*
- IN** – Integrase
- INSA** – Instituto Nacional de Saúde
- LTR** – Regiões repetitivas terminais longas, do inglês *Long Terminal Repeat*
- MA** – Proteína da Matriz
- MEGA** – *Molecular Evolutionary Genetic Analysis*
- MHC** – Complexo Principal de Histocompatibilidade, do inglês *Major Histocompatibility Complex*
- NC** – Proteína da Nucleocápside
- NSI** – Variantes virais não indutoras de sincícios, do inglês *Non-Syncytium Inducing*
- Pb** – Pares de bases
- PCR** – Reação em cadeia da polimerase de ADN, do inglês *Polymerase Chain Reaction*
- PMBC** – *Peripheral Blood Mononuclear Cells*
- PR** – Protease
- SI** – Variantes virais indutoras de sincícios, do inglês *Syncytium Inducing*
- SIDA** – Síndrome da Imunodeficiência Adquirida
- SPSS** – *Statistical Package for Social Sciences*
- SU** – Proteína de Superfície
- TM** – Proteína Transmembranar
- TR** – Transcriptase Reversa
- UNAIDS** – *Joint United Nations Programme on HIV/AIDS*
- USA** – Estados Unidos da América, do inglês *United States of America*
- URF** – Formas virais recombinantes únicas, do inglês *unique recombinant forms*
- VIH** – Vírus da Imunodeficiência Humana
- VIH-1** – Vírus da Imunodeficiência Humana do tipo 1
- VIH-2** – Vírus da Imunodeficiência Humana do tipo 2
- VIS** – Vírus da Imunodeficiência Símia
- VIS<sub>cpz</sub>** – Vírus da Imunodeficiência Símia dos chimpanzés
- VIS<sub>gor</sub>** – Vírus da Imunodeficiência Símia dos *Gorillas sp.*
- VIS<sub>sm</sub>** – Vírus da Imunodeficiência Símia de *Sooty mangabeys*

# I.INTRODUÇÃO

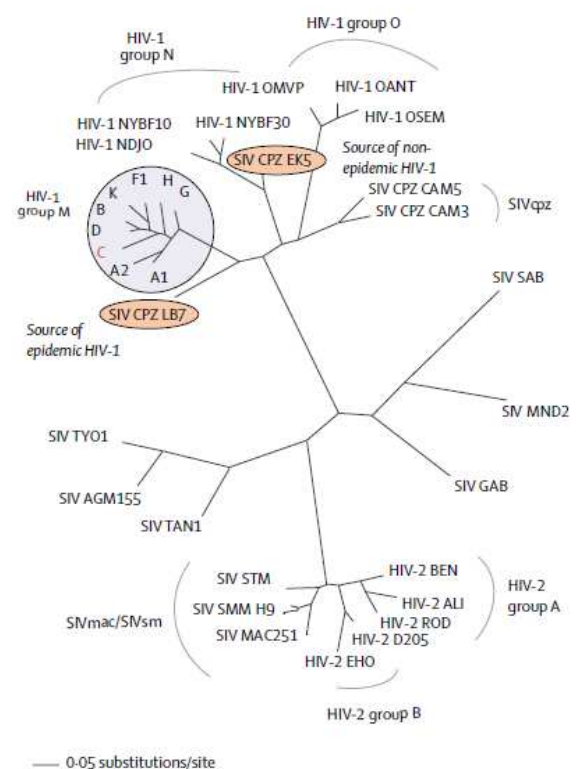
## 1. VÍRUS DA IMUNODEFICIÊNCIA HUMANA

O Vírus da Imunodeficiência Humana (VIH) é o agente etiológico da SIDA (Síndrome da Imunodeficiência Adquirida) que, segundo o último relatório da UNAIDS já provocou, a nível mundial e desde o início da pandemia a morte a 25 milhões de indivíduos. Em 2009, o número estimado de novas infecções foi de 2,7 milhões, existindo actualmente cerca de 33 milhões de pessoas infectadas no globo<sup>1</sup>. O aumento contínuo do número de indivíduos que vivem infectados com VIH mais tempo, reflecte o elevado número de novas infecções que vão surgindo, e também, o impacto benéfico da terapia anti-retrovírica no atraso da progressão para SIDA.

Existem dois tipos de vírus causadores de SIDA sendo designados de VIH-1 e VIH-2<sup>2</sup>. O primeiro foi isolado em 1983, pela equipa de Luc Montagnier<sup>3</sup> e o segundo, em 1986, pela mesma equipa, com a colaboração de investigadores portugueses<sup>4</sup>. Os dois tipos de vírus diferem entre si, em aspectos da sua organização genómica, nas relações filogenéticas com outros vírus, apresentando cerca de 55% de divergência quando as suas sequências nucleotídicas são comparadas entre si<sup>5</sup>. Enquanto o VIH-1 é responsável por uma pandemia a nível global, o VIH-2 é responsável por epidemias localizadas sobretudo em países da África Ocidental, apresentando ainda, prevalências muito reduzidas em alguns países Europeus (por exemplo em Portugal), e em países Asiáticos como a Índia<sup>6</sup>.

### 1.1. ORIGEM

Estudos filogenéticos permitiram mostrar que os dois tipos de VIH presentes na população humana surgiram, a partir de vírus da imunodeficiência dos símios (VIS) que infectam primatas não humanos Africanos<sup>7</sup>. VIH-1 encontra-se mais próximo de VIS<sub>cpz</sub>, que infecta chimpanzés da subespécie *Pan troglodytes troglodytes*<sup>8</sup>, enquanto VIH-2 encontra-se filogeneticamente mais próximo de VIS<sub>sm</sub>, que infecta *Cercocebus torquatus atys*<sup>9</sup>, mais conhecido como o macaco africano *sooty mangabey*, endémico na África ocidental. Neste último, a sua origem está claramente definida, uma vez que a área afectada (desde a costa do Senegal até ao Gana) coincide com a distribuição geográfica desta espécie de macacos, sendo simultaneamente uma região onde se pratica a caça frequente destes animais, favorecendo o contacto com o ser humano e potenciando a transmissão inter-espécies. Para além da prevalência no hospedeiro natural, da sobreposição geográfica e via de transmissão conhecida, a



**Figura 1** Relação filogenética dos lentivírus no homem e em primatas<sup>10</sup>.

semelhança na organização genômica também comprova a hipótese descrita, uma vez que o gene *vpu*, só existe na linhagem de VIH-1 e VIS<sub>opz</sub>, enquanto o gene *vpx* é exclusivo da linhagem de VIH-2 e VIS<sub>sm</sub><sup>11</sup>.

## 1.2. CLASSIFICAÇÃO, MORFOLOGIA E ORGANIZAÇÃO GENÔMICA

Ambos os tipos de VIH pertencem à família *Retroviridae*, sub-família *Orthoretrovirinae* e ao género *Lentivirus* (<http://ncbi.nlm.nih.gov/ICTVdb/index.htm>). As partículas virais do VIH apresentam uma forma esférica, com um diâmetro de aproximadamente 100 nm, sendo constituídas pela cápside viral, no centro, revestida por uma matriz e rodeada por um invólucro<sup>12</sup>. Este consiste numa bicamada lipídica, com glicoproteínas de superfície (gp120, SU) ancoradas a glicoproteínas transmembranares (gp41, TM); a matriz, essencial para integridade do virião, é proteica (p17, MA), assim como a cápside (p24, CA), em forma de cone truncado. No seu interior, além do genoma de ácido ribonucleico (ARN) estabilizado por proteínas da nucleocápside (p7) num complexo ribonucleoproteico, estão ainda presentes as enzimas protease (PR), transcriptase reversa (TR) e integrase (IN), assim como as proteínas Nef, Vif, Vpr e p6.

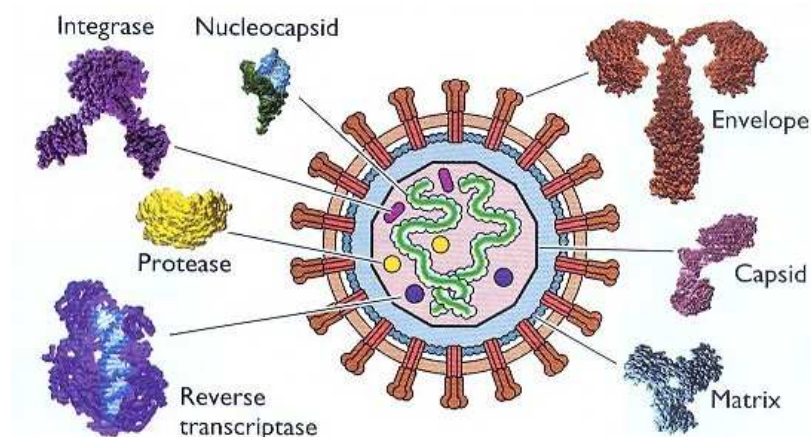


Figura 2 Representação da partícula VIH e respectivas proteínas<sup>2</sup>.

O VIH possui um genoma de cerca de 10 Kb constituído por duas cadeias simples de ARN de polaridade positiva. Como todos os Retrovírus, o genoma do VIH contém três genes principais – *gag*, *pol* e *env* – que codificam precursores proteicos que, uma vez clivados, originam respectivamente as proteínas estruturais (MA, CA, NC e p6), enzimáticas (PR, TR e IN) e do invólucro (SU e TM). Para além destes, possui ainda seis genes que codificam proteínas com funções regulatórias – Tat e Rev – e proteínas ditas acessórias, uma vez que não são essenciais para a replicação *in vitro*, mas cujo papel *in vivo* é importante – *vif*, *vpr*, *vpu* (apenas em VIH-1) ou *vpx* (só em VIH-2) e *nef*<sup>13</sup>.

### 1.3. CICLO DE REPLICAÇÃO

O ciclo de replicação do VIH, tal como o de todos os lentivírus de primatas, pode ser dividido em duas fases denominadas de fase precoce e de fase tardia<sup>13</sup>. Na fase precoce, que corresponde à infecção da célula, existe adsorção do vírus à célula alvo, que é mediada pela interacção entre a proteína de superfície (SU) e uma molécula CD4, presente na membrana plasmática dos linfócitos T-auxiliares e monócitos/macrófagos<sup>14</sup>. A presença da molécula CD4, apesar de necessária, não é por si só suficiente para que a infecção ocorra, sendo necessária a presença de co-receptores de quimiocinas, CCR5 e CXCR4<sup>14</sup>. O vírus entra na célula por fusão directa com a membrana celular, com a consequente libertação da cápside no citoplasma. As partículas víricas são descapsidadas, formando um complexo nucleoproteico<sup>15</sup>. A transcriptase reversa é activada, iniciando-se a síntese do ADN (ácido desoxirribonucleico) complementar a partir do primer tARN<sup>lys</sup>, que se encontra ligado ao ARN genómico vírico no local denominado PBS (*primer binding site*), localizado logo após o final do LTR 5'. Por fim, a integrase viral medeia a integração do ADN no genoma celular com formação do provírus<sup>15</sup>.

A fase tardia inicia-se com a transcrição do provírus e engloba tanto a expressão génica como a formação da partícula viral. As proteínas reguladoras víricas Tat e Rev vão ter, respectivamente, como função recrutar a maquinaria celular para controlar a expressão dos genes e a replicação viral. Assim, da transcrição vão resultar moléculas de ARN mensageiro, com diferentes graus de *splicing*, que são exportadas para o citoplasma com o auxílio de Rev, originando tanto o genoma presente nos novos viriões como moléculas que serão traduzidas nos precursores das várias proteínas virais<sup>15</sup>. A partícula vírica é formada ao nível da membrana plasmática<sup>2</sup>. O novo virião forma-se por gemulação, adquirindo o invólucro através da passagem pela membrana celular e a sua maturação ocorre, no espaço extracelular, por intermédio da protease viral que promove rearranjos intramoleculares, essenciais à sua infecciosidade<sup>14,16</sup>.

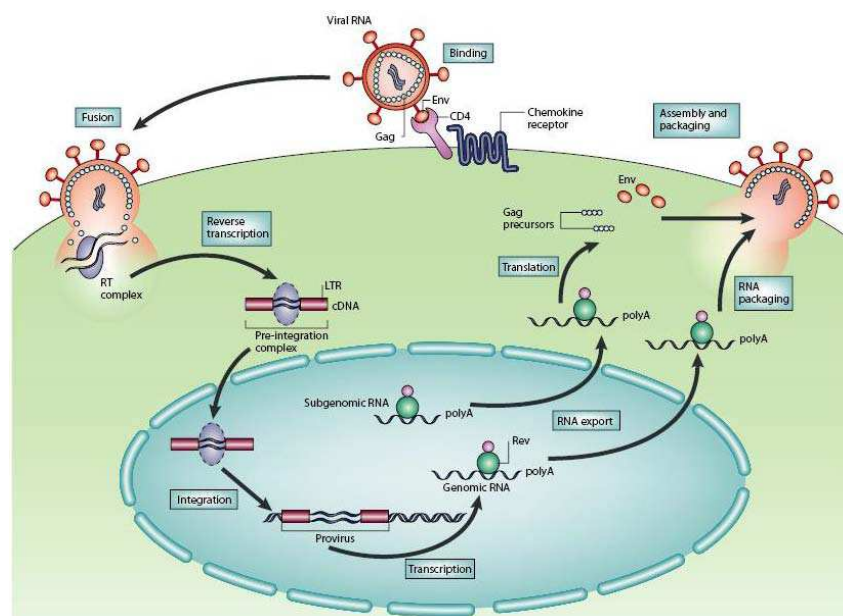


Figura 3 Representação do ciclo replicativo de VIH<sup>17</sup>.

#### 1.4. ESTABELECIMENTO DA INFECÇÃO VIH E DESENVOLVIMENTO DA IMUNODEFICIÊNCIA

O VIH é transmitido através de relações sexuais não protegidas, via sanguínea, por contacto com sangue e/ou material contaminado (utilização de material de injeção não esterilizado), ou por via vertical, da mãe infectada para o filho, durante a gravidez, no parto ou através da amamentação<sup>2,18</sup>. A infecção VIH é caracterizada por três fases distintas: fase aguda, fase crónica ou de latência clínica e a fase final de SIDA<sup>19</sup>. Aproximadamente, 3 a 6 semanas após exposição e infecção primária por VIH, cerca de metade dos infectados apresenta síndrome clínica aguda, com manifestações semelhantes à mononucleose infecciosa, autolimitada, que dura, geralmente, uma a três semanas. Estas manifestações associam-se à subida da virémia e são acompanhadas pela quebra transitória dos níveis de linfócitos T CD4<sup>+</sup>, T CD8<sup>+</sup> e B, no sangue periférico<sup>19</sup>. A seroconversão, isto é, a produção de anticorpos contra o VIH, ocorre geralmente 4 a 6 semanas após a infecção. Nesta fase primária, a população viral é relativamente homogénea, constituída por variantes R5 (utilizam o co-receptor CCR5) que infectam preferencialmente macrófagos (tropismo M)<sup>2,20</sup>.

A fase de infecção crónica está associada a uma fase assintomática, onde a latência é apenas clínica, uma vez que esta é caracterizada pelo enfraquecimento continuado do sistema imunitário e uma replicação viral persistente. A sua duração é variável, de acordo com as características do vírus, do indivíduo e do uso ou não de terapêutica anti-retrovírica, caracterizando-se no geral por níveis de virémia mais reduzidos e pelo aumento da heterogeneidade da população viral<sup>19</sup>.

A terceira e última fase, correspondente ao estágio de SIDA, que é caracterizada por um aumento da replicação viral e por uma contagem de células T CD4<sup>+</sup> abaixo de 200 células/mm<sup>3</sup>. Neste estágio, as variantes X4, que utilizam o co-receptor CXCR4, tornam-se predominantes, surgindo uma população viral novamente homogénea com tropismo T (de linfócitos T)<sup>20</sup>. A progressão da doença conduz a um risco crescente de reactivação ou aparecimento de infecções oportunistas que, se não forem prevenidas e controladas, podem levar à morte<sup>2,19</sup>.

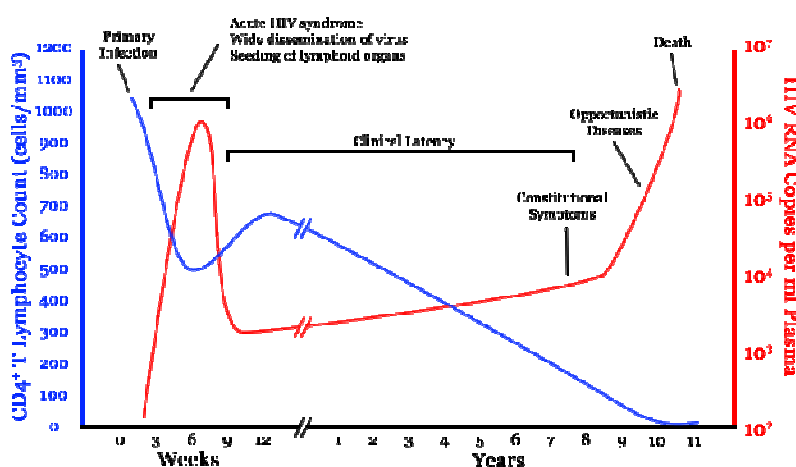


Figura 4 Evolução da carga viral (vermelho) e da contagem de células T CD4<sup>+</sup> (azul) ao longo do tempo em indivíduos infectados pelo VIH e não tratados<sup>21</sup>.

## 1.5. VARIABILIDADE GENÉTICA DO VIH

### 1.5.1 Causas e Consequências da Variabilidade

Uma das principais características do VIH é a sua elevada variabilidade genética, que resulta principalmente de duas propriedades da enzima transcriptase reversa<sup>22,23</sup>. A primeira está relacionada com o elevado número de erros efectuados pela enzima durante a síntese de ADN viral, devido à inexistência de actividade exonucleásica 3' → 5', o que origina uma elevada taxa de mutação. A segunda consiste num processo de recombinação, uma vez que a TR, aquando da transcrição reversa para ADN, alterna entre as duas moléculas de ARN. Estes eventos podem resultar na formação de um genoma mosaico, caso a célula tenha sido infectada por diferentes vírus contendo informação genómica de origem distinta<sup>24</sup>. As características da TR, aliadas à elevada taxa de replicação das partículas virais (cerca de  $4,4 \times 10^{10}$  viriões produzidos diariamente) geram a imensa diversidade do VIH<sup>25</sup>. Para além disso, também ocorrem inúmeras substituições, deleções e inserções, embora a frequência destes erros seja difícil de estimar, sendo a taxa de mutação distinta e influenciada pelas diferentes pressões selectivas para os vários genes virais. O processo de recombinação constitui uma força importante na evolução do VIH, não só no indivíduo em si como também na epidemia global desta infecção<sup>6</sup>.

Uma consequência da complexidade de populações virais observada em hospedeiros infectados por VIH-1, é de que estas populações são capazes de responder rápida e eficientemente às perturbações do ambiente onde se replicam, porque oferecem um grande espectro de mutantes sobre os quais a selecção natural pode actuar. Quando existem alterações no ambiente, como por exemplo, na presença de anti-retrovíricos, a existência de mutantes mais aptos a replicar nestas novas condições faz com que a população derivada desses mutantes ganhe aptidão (fitness) e aumente as oportunidades de sobreviver. No caso do VIH-1, a pressão selectiva exercida pelo sistema imunitário resulta na adaptação do vírus a novas células alvo e na manutenção de uma infecção persistente<sup>26</sup>.

### 1.5.2 Classificação das Estirpes

Actualmente a classificação adoptada baseia-se na análise do genoma completo de amostras de VIH colhidas de diferentes regiões geográficas<sup>26</sup>.

As estirpes de VIH-1 podem ser classificadas em 4 grupos designados M, O, N e P<sup>27,28</sup>.

Relativamente ao grupo M conhecem-se até ao momento, 9 subtipos geneticamente distintos, designados por subtipos de A a D, de F a H, J, K, e ainda também o recente proposto subtipo L<sup>27</sup>. Os subtipos A e F subdividem-se em sub-subtipos, respectivamente, A1 a A4 e F1 a F2. O grupo M inclui a esmagadora maioria dos vírus responsáveis pela pandemia VIH/SIDA.

De VIH-2, conhecem-se actualmente 8 grupos ou também designados subtipos, geneticamente distintos, identificados de A a H, sendo os grupos/subtipos A e B responsáveis pela epidemia VIH/SIDA e os restantes representam casos de infecção esporádicos na população.

Como existem áreas geográficas afectadas com mais do que um subtipo ou variante VIH-1 em circulação na população podem ocorrer infecções num mesmo indivíduo por vírus de diferentes subtipos. A presença destes vírus numa célula pode conduzir à recombinação entre genomas e

originar um vírus “híbrido”. A maioria dos vírus “híbridos” são defectivos e apenas as formas viáveis e infecciosas são transmitidas a um novo hospedeiro. Quando uma determinada composição genómica é encontrada num único indivíduo, esta é classificada como URF, “formas recombinadas únicas”<sup>29</sup>. As formas virais que possuem características que lhes conferem um fitness que lhes permite superar pressões selectivas, são facilmente disseminadas na população e tornam-se CRF, “formas recombinadas em circulação”<sup>6</sup>. A recombinação entre diferentes CRF resulta numa segunda geração de recombinantes, normalmente denominadas de formas “complexas” (cpx), sendo a diversidade intragénica elevada e apresentando uma contínua capacidade de disseminação na população<sup>6</sup>.

Actualmente os 9 subtipos e as 48 CRF de VIH-1 identificadas na base de dados *Los Alamos*<sup>30</sup> apresentam uma distribuição específica, em todo o mundo<sup>31</sup> mas esta tem variado ao longo do tempo.

### 1.5.3 Epidemiologia molecular no Mundo

As estirpes de VIH-1 pertencentes ao grupo M distribuem-se segundo um padrão geográfico em constante evolução, embora a maior diversidade genética da infecção VIH-1 possa ser observada na região da África Sub-Sahariana onde todos os subtipos são reportados<sup>32</sup>.

Tendo em conta a prevalência molecular da infecção VIH-1, existem quatro subtipos que dominam a epidemia mundial, os subtipos A, B, C e D, e também duas CRF, as CRF01\_AE e CRF02\_AG<sup>33</sup>.

A distribuição do subtipo A está concentrada na parte Este do continente Africano e nas antigas Repúblicas Soviéticas. O subtipo B está difundido globalmente dominando epidemias no continente Americano, Europa Ocidental e Austrália. O subtipo C, que é responsável por mais de 50% de todas as infecções mundiais, é prevalente a Sul e Este do continente Africano e também na Índia e na China. O subtipo D circula no Leste do continente Africano. A CRF01\_AE e o subtipo B co-circulam no Sudeste Asiático e a CRF02\_AG juntamente com outros recombinantes, domina a infecção descrita na África Ocidental e Central. Na América do Sul, a epidemia consiste numa mistura de estirpes do subtipo B e recombinantes BF, existindo uma pequena proporção de infecções causadas pelo subtipo C. Na Ásia, os subtipos B, C e as estirpes recombinantes BC dominam a epidemia<sup>34</sup>.

As estirpes de VIH-1 do grupo O e N provêm principalmente dos Camarões, país onde surgiram casos classificados no grupo P<sup>28</sup>. As infecções por VIH-2 predominam na África Ocidental<sup>35</sup>.

As actuais URF descritas desempenham um importante papel e podem ser um forte contributo para a evolução da pandemia. De facto, cerca de 30% dos casos que se identificaram na África do Leste são URF, entre as quais, a URF\_AC e a URF\_AD, que foram descritas em países como a Tanzânia, Zâmbia, Uganda e Quénia<sup>36</sup>.

Um problema que surge é o facto de certos subtipos que apresentavam uma prevalência rara estarem a emergir e a disseminar-se como formas geneticamente recombinadas. Exemplos disso são a CRF12\_BF e a CRF14\_BG que representam novas linhagens dos subtipos F e G mas agora na forma de estirpes recombinantes. O subtipo F tem sido detectado na África Central, América do Sul e Europa Oriental. O subtipo G e a CRF\_AG foram observados na África Ocidental e Europa Central. O

subtipo H só foi encontrado na África Central, o subtipo J, apenas, na América Central e o subtipo K, na República Democrática do Congo e nos Camarões<sup>37</sup>.

Em suma, o subtipo C ou as formas recombinantes de VIH-1 que contêm pelo menos o gene *env* do subtipo C contribuem para 50% das infecções de VIH-1 no mundo, parcialmente como resultado das epidemias recentes na África do Sul, América do Sul e Ásia. Outros subtipos importantes como o A, B, G e D foram responsáveis por, respectivamente 12, 10, 6 e 3% das infecções por VIH-1 no mundo<sup>37</sup>. Os restantes subtipos (F, H, J, K), todos juntos são considerados responsáveis por apenas 0,9% das infecções. Ambas as CRF, designadas de CRF01\_AE e CRF02\_AG, contribuem para 5% das infecções e a CRF03\_AB representa apenas 0,1% das infecções a nível mundial. Todas as outras formas recombinantes representam cerca de 8 % do total das infecções mundiais<sup>37</sup>.

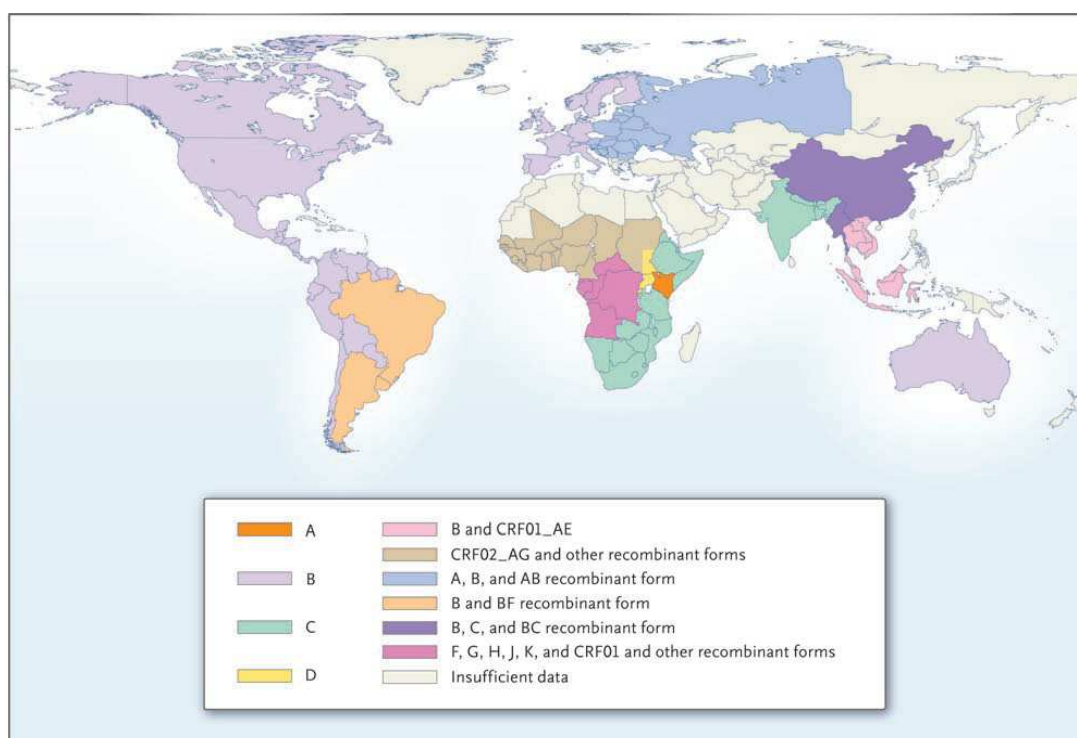


Figura 5 Distribuição Mundial dos subtipos de VIH-1 e das CRFs<sup>29</sup>.

### 1.5.3.1. Contributos na Transmissão e disseminação da infecção

O padrão de distribuição descrito sugere, indirectamente, que a disseminação do VIH-1 fora de África ocorreu através de alguns indivíduos infectados. A introdução inicial do vírus numa área não afectada é caracterizada geralmente por um determinado factor de transmissão que conduz a uma rápida disseminação de um único subtipo ou CRF num grupo de risco definido. Por exemplo, inicialmente o subtipo B, que predomina na América do Norte, Europa Ocidental e Austrália, afectava principalmente homossexuais, e o subtipo A, que está associado a infecções em utilizadores de drogas injectáveis (UDI) era prevalente na Rússia. Em contraste, na Tailândia ocorreram duas epidemias em simultâneo com o subtipo B e a CRF01\_AE, mas a disseminação dos vírus ocorreu em dois grupos de comportamento de risco diferentes. Os UDI foram infectados pelo subtipo B, enquanto indivíduos heterossexuais foram predominantemente infectados pela CRF01\_AE<sup>38</sup>.

Alguns grupos de indivíduos com comportamentos de risco, particularmente viajantes, contribuem para o início de epidemias locais ou para o aumento da diversidade do VIH-1. Estes grupos incluem, em particular, imigrantes, UDI, turistas, camionistas, militares e pescadores. Em países como a Rússia, China, Índia, Brasil e África do Sul, as migrações internas nos países contribuem para a disseminação e a diversidade da infecção do VIH-1<sup>38</sup>.

A contribuição das migrações dos indivíduos para a disseminação do VIH-1 pelo mundo está bem ilustrada pelo primeiro caso de infecção por VIH-1 na Europa, identificado numa família Norueguesa. O pai, um pescador infectado pelo VIH-1 do grupo O no princípio dos anos 60, provavelmente através do contacto heterossexual em África, transmitiu a infecção à sua mulher que por sua vez transmitiu à filha. Todos faleceram em 1976. Esta observação sugere que, dependendo do comportamento dos indivíduos, a disseminação do VIH-1 pode estar limitada a poucos indivíduos e pode não conduzir necessariamente a uma epidemia local<sup>38</sup>.

Como a maioria dos estudos englobam os países ocidentais onde o subtipo B é predominante, a visão global da contribuição das deslocações das populações para a disseminação do vírus deriva da prevalência dos subtipos não-B nesses mesmos países. Embora as infecções de subtipos não-B sejam pouco frequentes na América do Norte, um estudo em Nova Iorque identificou infecções não-B em alguns cidadãos que nunca viajaram, sugerindo que a transmissão do vírus também ocorre nos EUA<sup>39</sup>.

Assim, a distribuição geográfica das variantes do VIH-1 (grupos, subtipos e CRFs) é heterogénea, diferindo em extensão e temporalmente. A dinâmica populacional das variantes é influenciada pela intensa circulação de indivíduos infectados, apesar de existirem indícios de que factores virológicos possam contribuir para o estabelecimento de epidemias, favorecendo alguns subtipos em detrimento de outros. Além disso, a caracterização molecular das estirpes prevalentes nos países afectados pela pandemia permite que se acompanhe retrospectivamente a trajectória mundial do vírus, reconstruindo-se a introdução de variantes nas diferentes populações, e se possa monitorizar as reintroduções do VIH-1 em grupos de comportamento de risco de infecção<sup>38</sup>.

### 1.5.3.2. Infecção VIH em Portugal e na Guiné-Bissau

O relatório de 31 de Agosto de 2010 do Núcleo de Vigilância Epidemiológica do Departamento de Doenças Infecciosas (DDI-URVE) do Instituto Nacional de Saúde (INSA) refere a existência em Portugal, até à presente data, de 38 608 casos de infecção por VIH/SIDA nos diferentes estádios de infecção. Como elemento comum a todos os estádios, verifica-se que o maior número de casos notificados (“casos acumulados”) corresponde à infecção por transmissão sexual (heterossexual) constituindo 41,8% de todas as notificações. O número de casos associados à infecção em indivíduos referindo consumo de drogas por via endovenosa ou “toxicodependentes” representa o segundo grupo com 40,9% dos registos, revelando que a situação em Portugal se inverteu nos últimos anos uma vez que, este grupo corresponde à categoria de transmissão que mais tem contribuído para o número de casos notificados de VIH no país. A transmissão sexual (homossexual masculina) apresenta 13,6% dos casos e as restantes formas de transmissão correspondem a 4,4% do total.

Desde o início da epidemia até final de Agosto de 2010, cerca de 3,4% dos casos de VIH/SIDA notificados, em Portugal, foram devidos à infecção por VIH-2, sendo a prevalência deste tipo de vírus a mais elevada da Europa. Devido às ligações sociais e económicas com ex-colónias de África (Angola, Moçambique e Guiné-Bissau), Ásia (Índia) e América do Sul (Brasil), Portugal possui um perfil único de infecção que o diferencia do resto da Europa. As estirpes de VIH-1 mais comuns em Portugal têm sido classificadas nos subtipos B (41,7%) e G (29,4%), contudo, ainda que menos frequentes, são reportados outros vírus classificados em outros subtipos virais. São exemplos os vírus dos subtipos C (7,2%), F (2,2%), A (1,7%), D (1,1%) e J (0,6%)<sup>39</sup>. Alguns estudos referem ainda prevalências na ordem dos 16% do total de infecções de formas recombinantes, nomeadamente CRF02\_AG e CRF14\_BG<sup>40, 41,42</sup>, bem como um aumento significativo de casos de infecção pelo subtipo G.

No que diz respeito à Guiné-Bissau, segundo alguns dados estatísticos,<sup>1</sup> viviam infectados com VIH/SIDA neste país cerca de 16 000 indivíduos e o valor estimado da prevalência da infecção VIH nos indivíduos dos 15 aos 49 anos em 2007 era de 1,8%. Cerca de 3,8% da população adulta está infectada com VIH/SIDA e milhares de mulheres grávidas são seropositivas, elevando o risco de transmissão do vírus aos recém-nascidos. A Guiné-Bissau teve ao longo de muitos anos a maior taxa de prevalência de infecção por VIH-2 comparativamente ao resto do mundo. A partir da década de 90, a prevalência de VIH-1 tem aumentado rapidamente observando-se um decréscimo progressivo da infecção por VIH-2 ao longo dos anos. Um dos raros estudos publicados<sup>43</sup>, sobre a epidemia molecular de VIH-1 na Guiné-Bissau, revelou que a maioria das sequências de VIH-1 neste país, está estreitamente relacionada com os isolados IbNG, DJ263 e DJ264, partilhando um padrão de



Figura 6 Localização da Guiné-Bissau<sup>45</sup>.

recombinação A/G comum. Para além disso, estes investigadores descreveram que esta forma recombinada é característica nos países da África Ocidental e que é diferente das estirpes subtipo A que circulam na África Oriental. Isto revela a prevalência de um subtipo recombinante A/G IbNG-like [CRF AG (IbNG)] na Guiné-Bissau. Os dados disponíveis indicam que as formas virais dominantes de VIH-2 incluem apenas o subtipo A<sup>44</sup>.

## 1.6. IMPACTO DA DIVERSIDADE GENÉTICA DO VIH NA PANDEMIA

### 1.6.1. Patogenicidade, Transmissibilidade e Progressão para Doença

Vários estudos sugeriram que os diversos subtipos virais podem apresentar características biológicas distintas, tais como diferentes taxas de transmissibilidade e de infecciosidade conduzindo a uma progressão mais lenta ou mais rápida para a doença<sup>34,35</sup>. Num estudo publicado em 2006 foi

descrito que, indivíduos do Uganda não tratados e infectados pelo subtipo D, desenvolveram SIDA mais cedo do que indivíduos infectados pelo subtipo A. Os autores do estudo sugeriram que o subtipo D é mais virulento pois apresenta maior eficácia na ligação e infecção das células<sup>46</sup>. Este resultado está em acordo com um outro estudo realizado em 2007, que mostrou que indivíduos infectados com o subtipo D possuem 2x maior risco de morte, comparativamente a indivíduos infectados com o subtipo A<sup>47</sup>. Também um estudo na Tailândia sugeriu que indivíduos não tratados e infectados com a CRF01\_AE progrediam mais rapidamente para SIDA, do que indivíduos infectados com o subtipo B<sup>48</sup>. Outros estudos revelaram que certos subtipos/CRF estão predominantemente associados a vias de transmissão específicas. Em particular, o subtipo B é transmitido preferencialmente por via homossexual também pela via sanguínea através do uso de material de injeção contaminado (consumidores de drogas), enquanto o subtipo C e a CRF01\_AE estão associados à transmissão por via heterossexual. Contudo, permanece ainda a controvérsia sobre a existência ou não de causas virológicas para as diferenças observadas<sup>49,50</sup>. Estudos mais recentes descrevem que a transmissão perinatal é mais comum ocorrer no subtipo D comparativamente a infecções com o subtipo A. No entanto, outros investigadores mostraram não levantar qualquer associação entre os subtipos e esta via de transmissão<sup>51</sup>. O assunto continua em controvérsia.

### 1.6.2. Diagnóstico e Monitorização da Infecção

A diversidade genética do VIH tem tido um elevado impacto na sensibilidade dos métodos utilizados na detecção da infecção. Presentemente, para o diagnóstico das infecções pelo VIH recorre-se a testes serológicos, através dos quais se detectam anticorpos anti-VIH<sup>32</sup>. Actualmente, os testes comerciais permitem a detecção dos 2 tipos de vírus e da maioria das estirpes virais pertencentes aos grupos M e O do VIH-1, sendo sensíveis e específicos no diagnóstico de infecções estabelecidas em adultos. Geralmente, são utilizados testes imunoenzimáticos para rastreio, vulgarmente apelidados de EIA, e nas amostras reactivas é recomendada a confirmação do resultado por testes imunoenzimáticos mais específicos, designados por Western-Blot. Recentemente, também existem disponíveis no mercado ensaios que detectam simultaneamente anticorpos e antígenos contra o VIH, de forma a encurtar os períodos de janela serológica. No entanto, em casos pontuais e para algumas infecções causadas por subtipos não-B, os resultados baseados numa potencial reacção cruzada podem ser dúbios ou inconclusivos, sobretudo porque os antígenos utilizados nos ensaios são baseados na informação relativa ao subtipo B<sup>52</sup>. Nas situações em que os testes serológicos não podem ser aplicados, como por exemplo, no diagnóstico precoce de infecções em recém-nascidos de mães infectadas (idade inferior a 12-15 meses de vida), a estratégia consiste numa abordagem molecular com pesquisa directa do vírus<sup>32,52</sup>.

Assim, recorrendo à técnica PCR e utilizando conjuntos de *primers* genéricos de regiões conservadas, é possível amplificar, *in vitro*, todos os subtipos de VIH<sup>53</sup>. Por seu lado, a detecção de RNA VIH-1 (carga vírica) é uma das abordagens experimentais que mais tem beneficiado do desenvolvimento e aperfeiçoamento dos ensaios moleculares, sendo particularmente importante para

a monitorização e consequente avaliação da eficiência da terapêutica anti-retrovírica de indivíduos infectados em tratamento.

### 1.6.3. Mutações de Resistência e Implicações no Desenvolvimento de Vacinas

Nos países desenvolvidos, onde predominam as infecções com os vírus do subtipo B, as terapias combinadas de alta eficácia, designadas por HAART (do inglês, *Highly Active Antiretroviral Therapy*) têm originado uma enorme redução da transmissão, morbidade e mortalidade associadas à infecção pelo VIH-1<sup>34,35</sup>. No uso combinado de dois grandes grupos de inibidores da RT, os análogos de nucleósidos (NRTI, do inglês, *Nucleoside Reverse Transcriptase Inhibitors*) e os não análogos de nucleósidos (NNRTI, do inglês *Non Nucleoside Reverse Transcriptase Inhibitors*)<sup>54</sup>, como também outras classes de drogas como inibidores da protease viral, inibidores de fusão e inibidores da integrase, permitiram uma diminuição dos níveis de carga vírica na infecção, e consequentemente, restringiram as probabilidades de risco de transmissão do vírus, permitindo o restabelecimento do sistema imunitário do indivíduo infectado<sup>55</sup>. Em contrapartida, à baixa adesão à terapêutica anti-retrovírica está associada a emergência de mutações de resistência e variantes víricas. Gradualmente, os subtipos designados de “puros” estão a ser substituídos por formas de vírus geneticamente recombinadas, potencialmente de maior virulência e transmissibilidade. Por isso, permanece a necessidade de ser determinado se estas variantes têm alguma vantagem biológica sobre os seus subtipos parentais, ou se existem outras razões para a sua fácil propagação, pelo mundo. De facto, as formas virais recombinadas podem ter um sério impacto na eficácia da terapia anti-retrovírica. Além disso, a possível emergência de uma segunda ou terceira geração de recombinantes, não só tornará a epidemia de VIH-1 mais complexa, como também poderá afectar gravemente uma estratégia futura na concepção de uma vacina<sup>37</sup>.

### 1.7. CO-INFECÇÃO COM MEMBROS DO COMPLEXO *Mycobacterium tuberculosis*

A Organização Mundial de Saúde (OMS) estima que um terço da população mundial está infectado com *Mycobacterium tuberculosis*, indicando 8 milhões de novos casos de tuberculose e 2 milhões de mortes por ano<sup>56</sup>.

No mundo, estima-se que existam aproximadamente 10 milhões de pessoas co-infectadas com VIH e *M. tuberculosis*, e que mais de 90% destes indivíduos, residam em países em desenvolvimento. A OMS indica que o risco de contrair tuberculose é 20 a 37 vezes maior entre os indivíduos infectados pelo VIH do que na população geral, dependendo da prevalência do VIH na população. Desta forma, a tuberculose é a infecção oportunista mais comum em todo o mundo, sendo a causa de morte mais vulgar em pacientes com SIDA<sup>57</sup>.

Estudos feitos na Guiné-Bissau, sobre os tipos de micobactérias mais frequentes no país, revelaram a presença de membros do complexo *M. tuberculosis* (CMTB) e também do complexo *M. avium*<sup>57</sup>. Os membros CMTB que causam tuberculose no homem, incluem *M. tuberculosis*, *M. bovis* (incluindo a estirpe vacinal BCG), *M. africanum*, *M. microti*, *M. canneti*, *M. pinnitedii* e *M. catrae*.

O objectivo de controlar e, eventualmente, eliminar a tuberculose no mundo exige um contínuo esforço de forma a responder às necessidades de prevenção e tratamento das pessoas em maior risco, incluindo aquelas que estão infectadas com o VIH. Os esforços para eliminar a tuberculose são, portanto, essenciais para reduzir também o número global de infecção pelo VIH e mortes por SIDA em todo o mundo<sup>56</sup>.

### 1.8. DOMÍNIOS ESTRUTURAIS E FUNCIONAIS DAS PROTEÍNAS ENV E NEF

#### 1.8.1. Glicoproteína do Invólucro Viral gp120

A estrutura do *core* da glicoproteína SU é composta por dois grandes domínios, interno e externo, contendo cinco regiões conservadas (C1-C5) e cinco domínios altamente variáveis (V1-V5)<sup>59</sup>.

Após a fase inicial de ligação de gp120 à molécula CD4 e, dado que esta não é suficiente para assegurar a entrada do virião na célula-alvo, há uma alteração da conformação de gp120 que promove a sua interacção com o co-receptor CCR5 ou CXCR4 (respectivamente o receptor  $\alpha$  e  $\beta$  de quimiocinas)<sup>14,20</sup>. A região que está associada à escolha de determinado co-receptor corresponde ao *loop* da região V3<sup>60</sup>. Duas cisteínas (respectivamente nas posições 1 e 35) são essenciais para a conformação do *loop* e encontram-se geralmente conservadas<sup>60</sup>. A carga da região V3 influencia a escolha de qual o co-receptor a ser utilizado pois, um menor número de cargas positivas está associado ao uso do co-receptor CCR5, enquanto um maior número de cargas positivas, ao uso do co-receptor CXCR4<sup>62</sup>. Nas posições 11 e 25 do V3 *loop*, em isolados que não têm a capacidade de formar sincícios (NSI), estão conservados aminoácidos com cargas neutras e negativas (respectivamente S/G e E/D), ao passo que na maior parte dos isolados que são capazes de induzir a fusão de várias células (SI), os aminoácidos possuem cargas positivas<sup>63</sup>.

Foi também identificada uma pequena região consenso (GPG), bem como um motivo mais abrangente (S/GXXXGPGXXXXXXE/D) conservado nos isolados NSI, cuja informação do conteúdo aminoacídico pode prever o co-receptor da variante viral a ser utilizado<sup>63</sup>.

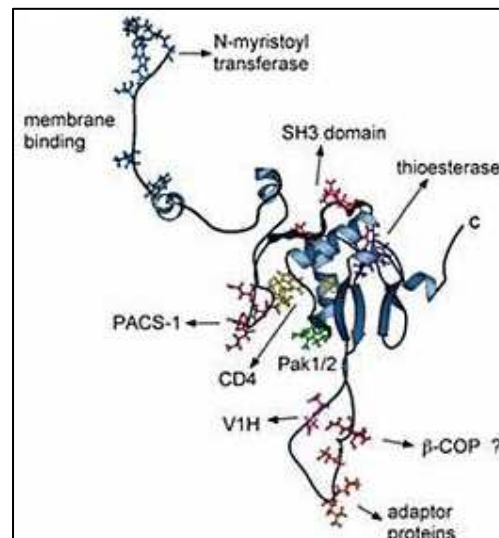
Assim, a gp120 torna-se um alvo preferencial da resposta imunitária, através de diversos locais de reconhecimento para anticorpos neutralizantes, linfócitos T auxiliares e citotóxicos (CTL, Cytotoxic T Lymphocytes)<sup>64</sup>. No entanto, as regiões mais conservadas entre as diferentes variantes estão menos expostas e são altamente glicosiladas, de forma a evitar o seu reconhecimento pelo hospedeiro, permitindo o escape do vírus à resposta imunitária<sup>65</sup>.

#### 1.8.2. Proteína Nef

A proteína Nef é abundantemente produzida durante as etapas iniciais de expressão génica viral<sup>66,67</sup>. Após tradução, é alvo de miristilação do terminal amina e clivagem pela protease viral. A miristilação ocorre no resíduo de glicina do motivo MGxxxS e é essencial para a associação da proteína com a membrana celular, bem como para a realização de quase todas as suas funções

biológicas<sup>68</sup>. Quanto à clivagem, ocorre já na fase de maturação da partícula viral entre os resíduos 57 e 58, separando a proteína em dois domínios, âncora e core<sup>69</sup>.

Nef é uma proteína versátil, que depende da conservação de domínios estruturais específicos (Figura 8) para desempenhar funções, que incluem a regulação negativa de CD4, a regulação de moléculas MHC classe I e II, interferência com vias de sinalização intracelulares e vias apoptóticas celulares e aumento da infecciosidade das partículas virais. A título de exemplo, um resíduo de tirosina encontrado nas caudas citoplasmáticas de moléculas MHC I dos haplótipos HLA-A e HLA-B, mas não em HLA-C, é determinante para a sua regulação negativa pela proteína Nef<sup>70</sup>. Para além do sinal de miristilação e dos motivos FPD<sub>121</sub> e EE<sub>154</sub>, envolvidos na regulação de MHC I (e também na regulação negativa de moléculas CD4), intervêm o cluster ácido EEEE<sub>62</sub>, que se liga a PACS-1 (uma molécula que controla o tráfego do endossoma para o Golgi) e o motivo de prolina<sup>71,72,73,74</sup>. Desta forma, as moléculas MHC I (e CD4) são internalizadas de uma forma mais rápida na presença de Nef, direccionadas para o complexo de Golgi e posteriormente degradadas nos endossomas<sup>70</sup>. Esta regulação tem uma elevada importância na infecção, dado que as moléculas MHC I apresentam antígenos virais na superfície celular, que permitem o reconhecimento e a posterior eliminação, por parte dos linfócitos T citotóxicos<sup>75</sup>. Nef, ao acelerar a internalização destas moléculas irá permitir um escape (ainda que não completamente) à resposta imunitária do hospedeiro, ao mesmo tempo que fornece uma vantagem selectiva para a manutenção e replicação viral *in vivo*<sup>76</sup>.



**Figura 7** Motivos funcionais da proteína Nef de VIH-1<sup>68</sup>.

## II.OBJECTIVOS

A variabilidade genética do VIH é reconhecida como um potencial problema para o diagnóstico e tratamento do VIH/SIDA, assim como para o desenvolvimento de vacinas globalmente efectivas, o que torna imprescindível o conhecimento e a monitorização da dinâmica de distribuição dos diferentes subtipos e variantes virais em circulação na população humana.

No estudo proposto, integrado num projecto de maior amplitude e alargado também à análise da “Variabilidade Genética do *Mycobacterium tuberculosis* para Identificação de Marcadores Moleculares de Virulência”, pretende-se caracterizar uma população de indivíduos atendidos no Hospital de Cumura na Guiné-Bissau, tendo como objectivo principal determinar a prevalência da infecção VIH-1 e VIH-2, identificar e caracterizar molecularmente as formas VIH-1 em circulação neste grupo populacional.

Os objectivos específicos deste trabalho incluem:

- ✓ O diagnóstico da infecção por VIH-1 e VIH-2 através da amplificação das respectivas regiões LTR para identificação dos casos de infecção no grupo.
- ✓ Determinar a prevalência da infecção por VIH-1 e VIH-2 e da eventual co-infecção com membros do complexo *M. tuberculosis* no grupo estudado.
- ✓ Nos casos positivos VIH-1, caracterizar molecularmente a região C2V3V5 de *env* e o gene *nef* recorrendo à construção de árvores filogenéticas e a uma análise por *bootscanning*, de forma a identificar as formas virais presentes na população residente em Bissau.
- ✓ Calcular as distâncias nucleotídicas entre as sequências obtidas em função de cada região genómica e de cada forma de vírus.
- ✓ Proceder a uma análise das sequências aminoacídicas das regiões V3-V5 de Env e Nef com especial relevo para a conservação ou disrupção de motivos nas proteínas que potencialmente possam estar associadas a funções do vírus.

Com o desenvolvimento deste projecto, pretende-se obter um melhor conhecimento sobre a epidemiologia molecular da infecção VIH em indivíduos residentes na Guiné-Bissau, onde são escassos estudos realizados nesta área.

### **III. MATERIAL E MÉTODOS**

#### **1. ANÁLISE DA POPULAÇÃO ALVO**

A população envolvida neste estudo representa uma amostra de conveniência obtida a partir de indivíduos que foram atendidos no Hospital de Cumura em Bissau durante o ano de 2009. Amostras de sangue total foram colhidas e imobilizadas em papel de filtro, posteriormente enviadas ao laboratório e conservadas a -80°C, até a sua utilização. A população foi caracterizada de acordo com dados disponibilizados para a implementação do presente estudo. Foi também disponibilizada informação clínica referente à infecção ou não dos indivíduos por membros do complexo *Mycobacterium tuberculosis*. A participação no estudo foi autorizada pelos indivíduos envolvidos através de consentimento informado, situação gerida pelos clínicos do Hospital de Cumura. Dados e amostras foram fornecidas na forma confidencial e não relacionável.

Assim, neste estudo foram incluídos 165 indivíduos, 83 (51%) pertencentes ao sexo masculino e 80 (48%) pertencentes ao sexo feminino, com intervalo de idades dos 2 aos 73 anos, com uma média de 36,6 anos e uma mediana de 32 anos de idade.

Os indivíduos foram também agrupados de acordo com a área geográfica em que residiam, verificando-se que a maioria dos indivíduos (69%, 115) residia em Bissau. No que respeita ao estudo da infecção relativamente ao VIH, o diagnóstico era desconhecido em 107 indivíduos. Do total da população, apenas tinham sido testados 58 indivíduos, dos quais 16 estavam infectados pelo VIH-1 e/ou VIH-2 e os restantes 42, tinham resultados de pesquisa de anticorpos anti-VIH negativos. Relativamente à infecção por membros do complexo *M. tuberculosis*, quase metade dos indivíduos (47%, 78) estava infectada por estes microrganismos.

#### **2. DESENHO DO ESTUDO**

##### **a) Estratégia para implementação de um método alternativo de extracção de ADN**

Com o objectivo de testar o método *Chelex* recorreu-se a uma amostra de conveniência constituída por 40 indivíduos infectados por VIH-1, todos sob terapia HAART e com níveis de RNA VIH-1 (carga vírica) indetectáveis. Por um lado, e para cada caso, as colheitas de sangue foram efectuadas para um suporte de papel de filtro (duas gotas de sangue), e por outro lado, para tubos contendo anticoagulante. Cada uma das amostras foi submetida a extracção de ADN usando respectivamente um método diferente alternativo (*Chelex*, gotas imobilizadas em papel de filtro) e um método tradicional (coluna, sangue em tubo). Após a extracção do ADN e respectiva quantificação foi realizada a amplificação pela técnica de *Nested* PCR da região LTR de VIH-1. Ambos os procedimentos foram efectuados em paralelo, como indicado no fluxograma descrito na Figura 8.

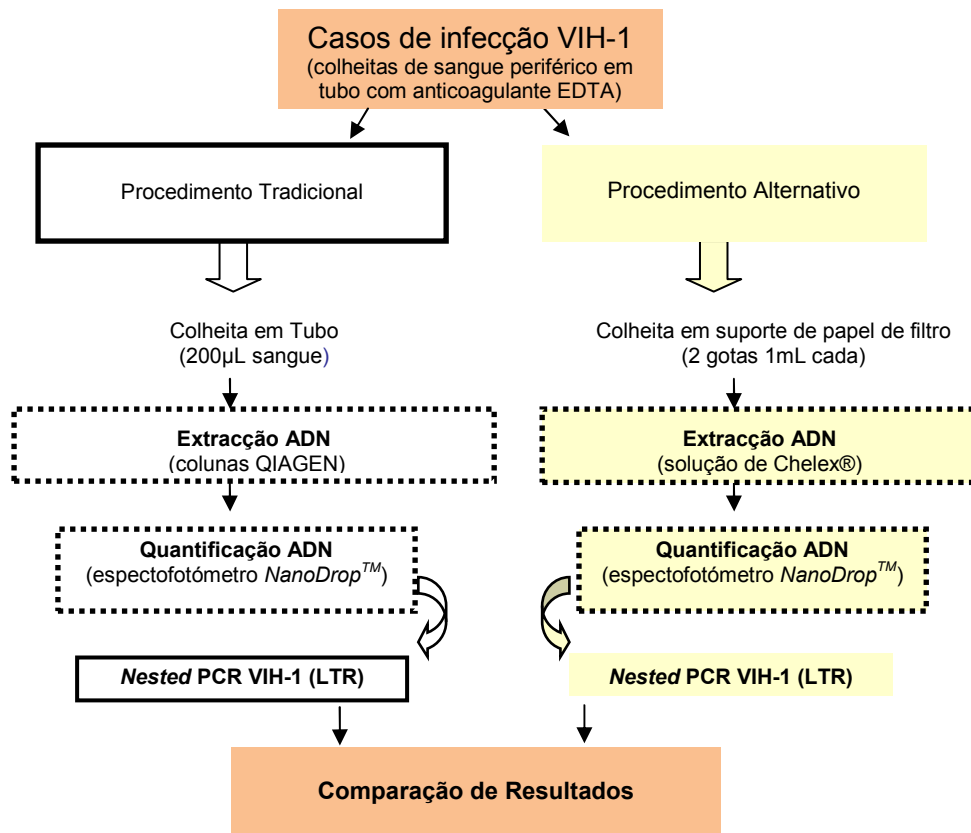


Figura 8 Representação esquemática dos dois procedimentos de extracção de ADN realizados em paralelo.

**b) Estratégia para análise da população alvo de estudo**

A estratégia metodológica adoptada para o estudo da população alvo baseou-se inicialmente na informação que foi disponibilizada pelos clínicos do Hospital de Cumura em Bissau. Todas as amostras correspondentes a indivíduos com informação conhecida de que estavam infectados pelo Vírus da Imunodeficiência Humana (VIH-1 e/ou VIH-2), e também, todas as amostras que não tinham sido testadas para anticorpos anti-VIH foram processadas, segundo o fluxograma descrito na Figura 9.

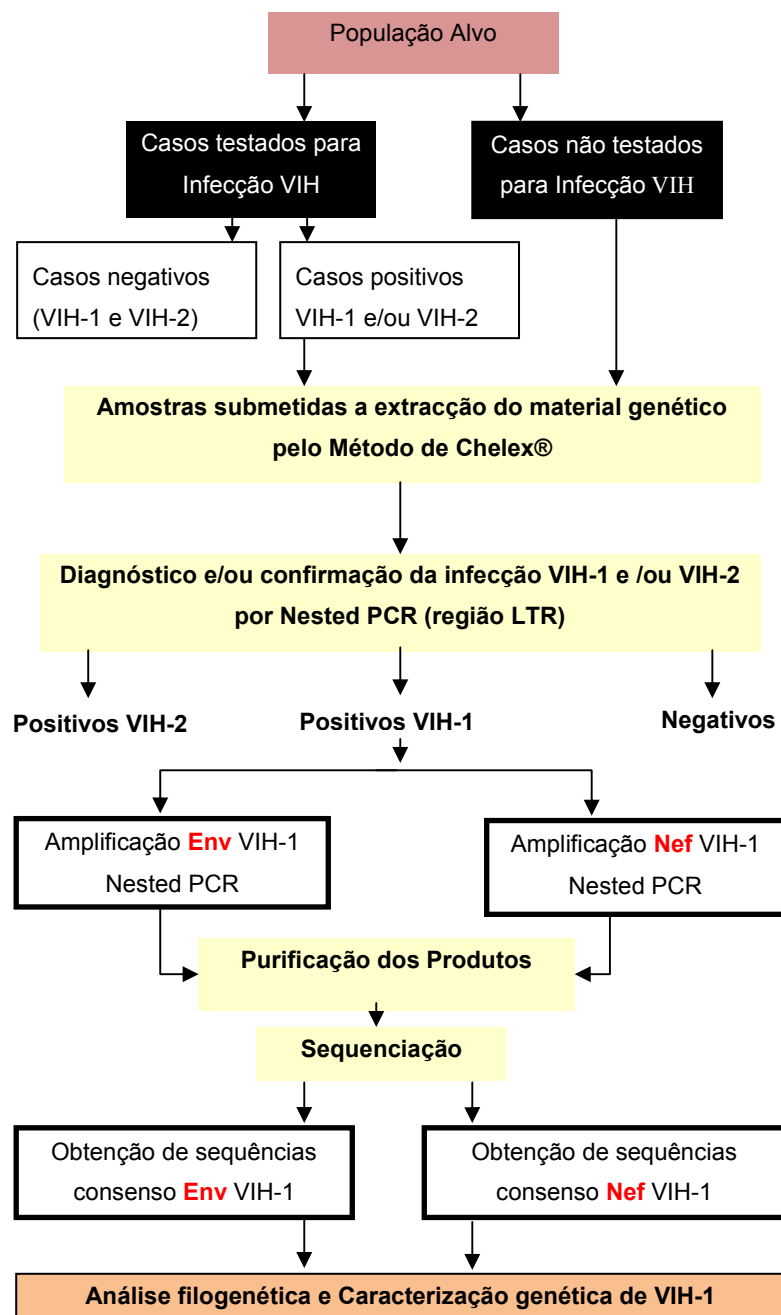


Figura 9 Diagrama da estratégia metodológica adoptada na população em estudo.

### 3. PREPARAÇÃO DAS AMOSTRAS

#### 3.1. Extracção de ADN a partir de sangue total imobilizado em papel de filtro

A extracção do ADN foi feita a partir de colheitas de cerca de 1ml de sangue imobilizado em papel de filtro em duplicado. Numa primeira fase, fez-se o recorte do disco (correspondente à gota de sangue) de aproximadamente 1,5cm de diâmetro e sua colocação em microtubo de 1,5ml. Adicionou-se 1ml de tampão PBS 1x concentrado a pH 7,4 e 50µl de solução de saponina a 10% (Alfa Aesar, A

*Johnson Matthey Company, Germany*) com incubação a 4°C durante a noite. Numa segunda fase, os microtubos foram centrifugados durante 5 segundos a 13000rpm (*Eppendorf Centrifuge 5415 D*) à temperatura ambiente e a solução resultante de cor avermelhada foi aspirada. Os discos foram submetidos a uma lavagem, com o mesmo tampão, e incubados a 4°C durante cerca de 30 minutos, seguida de nova centrifugação em iguais condições. Após concluída esta etapa, a solução foi aspirada para um novo microtubo, desta vez, apresentando uma cor amarelada. Adicionou-se 50µl de água bidestilada e 50µl de solução de Chelex® a 20% (*Bio – Rad Laboratories*) com posterior incubação a 95°C durante 10 minutos agitando os microtubos no vórtex de 2 em 2 minutos. Foi efectuada nova centrifugação durante 5 minutos a 13000rpm (*Eppendorf Centrifuge 5415 D*), tendo sido transferida a totalidade da solução para novos tubos, que por sua vez foram novamente centrifugados durante 10 minutos a 13000rpm. O sobrenadante (branco amarelado) foi transferido para novos tubos que foram armazenados a -20°C.

#### 3.2. Extração de ADN em colunas a partir de PBMC

Para a extração do ADN a partir de células mononucleadas de sangue periférico (PBMC) foram utilizadas colunas e reagentes incluídos no *QIAmp DNA Mini Kit* (QIAGEN, Valencia, USA) tendo sido seguidas as instruções do fabricante.

#### 3.3. Quantificação Espectrofotométrica do material genético

Os extraídos de ADN foram quantificados, através da absorvância a 260nm no espectrofotómetro *NanoDrop™ 1000 Spectrophotometer V3.7* (*Thermo Fisher Scientific Wilmington, DE 19810 U.S.A.*). Os valores foram controlados pela razão de absorvância 260/280nm, indicativa do grau de pureza do ADN obtido, em que a absorvância a 280nm quantifica a presença de proteínas. Após cálculo das concentrações foi estimada a quantidade de ADN a utilizar nas reacções de amplificação.

### 4. AMPLIFICAÇÃO DA REGIÃO GENÓMICA LTR DE VIH-1 E DE VIH-2

O diagnóstico de infecção VIH-1 e VIH-2 foi realizado através da técnica de *Nested PCR* (do inglês *Polymerase Chain Reaction*) com amplificação da região LTR, *Long Terminal Repeat*, especificamente do VIH-1 e VIH-2. As sequências LTR de VIH-1 foram amplificadas utilizando o conjunto de *primers* externos e internos descritos por Semple *et al.* (1991)<sup>77</sup>, resultando na obtenção de um fragmento de cerca de 124 pb. O conjunto de *primers* usados para a amplificação das sequências LTR VIH-2 foi descrito por Berry *et al.* (1994)<sup>78</sup>, originando um fragmento de cerca de 141pb.

Na amplificação foi utilizada a enzima *AmpliTaQ Gold®* (*Applied Biosystems, Foster City, USA*) e as misturas e condições de amplificação, encontram-se descritas nas tabelas 1 e 2. A quantidade de ADN adicionada na primeira reacção de amplificação variou entre os 5 µl e os 15µl de

acordo com as concentrações de cada amostra, enquanto na segunda reacção foram transferidos 2µl de produto amplificado.

**Tabela 1** Misturas de reacção para amplificação da região LTR de VIH-1 e de VIH-2.

| Misturas de Reacção                                     | VIH-1           |                 | VIH-2           |                 |
|---|-----------------|-----------------|-----------------|-----------------|
|   | 1ª Amplificação | 2ª Amplificação | 1ª Amplificação | 2ª Amplificação |
| 10x Tampão PCR  | 5 µl            | 2,5 µl          | 5 µl            | 2,5 µl          |
| 25mM MgCl <sub>2</sub>                                  | 3 µl            | 1,25 µl         | 4 µl            | 2 µl            |
| 1,25mM dNTPs  | 8 µl            | 4 µl            | 8 µl            | 4 µl            |
| 10µM Primers(C,N) de VIH-1<br>5µM Primers(C,N) de VIH-2 | 2 µl            | 1 µl            | 2,5 µl          | 1,25 µl         |
| 5U/µl AmplitaqGold                                      | 0,25 µl         | 0,25 µl         | 0,25 µl         | 0,125 µl        |
| Volume Final  | 50 µl           | 25 µl           | 50 µl           | 25 µl           |

**Tabela 2** Programa de amplificação da região LTR de VIH-1 e VIH-2.

| Programas de Amplificação VIH-1 / VIH-2 | 1ª amplificação |           | 2ª amplificação |            | Nº ciclos        |
|---|-----------------|-----------|-----------------|------------|------------------|
|   | Temperatura     | Tempo     | Temperatura     | Tempo      |                  |
| <b>Desnaturação inicial</b>             | 95°C            | 9m        | 95°C            | 9m         | 1 ciclo          |
| <b>Desnaturação</b>                     | 95°C            | 30s       | 95°C            | 15s / 30s* | <b>35 ciclos</b> |
| <b>Hibridação</b>                       | 50°C / 53°C*    | 1m / 45s* | 50°C            | 45s/ 30s*  |                  |
| <b>Síntese</b>                          | 72°C            | 1m/ 45s*  | 72°C            | 45s/ 30s*  |                  |
| <b>Extensão final</b>                   | 72°C            | 7m        | 72°C            | 7m         | 1 ciclo          |

\* No lado esquerdo da barra está informação para VIH-1 e no lado direito para VIH-2. m= minuto; s= segundo.

## 5. AMPLIFICAÇÃO DA REGIÃO V3-V5 DE ENV E NEF DE VIH-1

As sequências *env* e *nef* foram amplificadas por PCR, utilizando respectivamente os *primers* descritos por Delwart *et al.* (1993)<sup>79</sup> e Parreira *et al.* (2005)<sup>41</sup>. Relativamente à região *env*, a primeira amplificação foi efectuada com os *primers* externos ED3 e ED14 (que correspondem respectivamente às posições nucleotídicas 5957-5986 e 7932-7961 no isolado HXB2) e na segunda amplificação utilizaram-se os *primers* internos ED5 e ED12 (posições 6557-6582 e 7782-7811 em HXB2). Foi ainda realizada uma terceira amplificação com os primers ES7 e ES8 (posições 6983-7021 e 7648-7686 em HXB2), originando um fragmento final com cerca de 699pb, correspondente à região alvo pretendida de C2V3V5 de *env* de VIH-1.

Na amplificação do gene *nef* foram utilizados na primeira reacção de PCR os *primers* externos NefO1 e NefO2 (posições 8513-8633 e 9488-9508 em HXB2). Na segunda amplificação

foram usados os *primers* internos Nef1 e Nef2 (posições 8697-8717 e 9448-9468 em HXB2) que originaram um fragmento final com cerca de 771pb, correspondente à sequência total do gene *nef*.

Para a maioria das amostras, recorreu-se para a amplificação, ao sistema *puReTaq™ Ready-To-Go™ PCR Beads* (GE Healthcare, England) com concentração de *primers* de 0,25µM (no caso do gene *env*) e de 0,4µM (no caso do gene *nef*). Em alguns casos de amostras de difícil amplificação, alterou-se a mistura de reacção no sistema acima referido, aumentando a concentração dos *primers* e acrescentando 1µl de MgCl<sub>2</sub> a 25mM nas misturas de reacção descritas na tabela 3. A partir de produtos de PCR do 1º *round* (*primers* externos) recorreu-se também à tentativa de amplificação de um fragmento intermédio utilizando um *primer* externo (directo ou complementar) e um *primer* interno (complementar ou directo) seguindo-se uma 3ª reacção de PCR utilizando *primers* internos.

As misturas de reacção correspondentes às diferentes enzimas utilizadas, *AmpliTaq Gold®* (Applied Biosystems, Foster City, USA) e *Expand High Fidelity* (Roche), bem como os diferentes programas de amplificação encontram-se descritos nas tabelas 3, 4 e 5. A quantidade de ADN adicionada, às misturas de reacção na primeira amplificação, variou entre os 5µl, 10µl, e 15µl de acordo com a concentração da amostra. Para as amplificações seguintes na segunda e terceira amplificações foi transferido 2µl de produto inicialmente amplificado.

Tabela 3 Enzimas utilizadas nas misturas de reacção para amplificação de *env* e *nef* VIH-1.

| Enzima puReTaq™  | Enzima AmpliTaq Gold         | Enzima Expand High Fidelity       |
|--|------------------------------|-----------------------------------|
| puReTaq™<br>Ready-To-Go™<br>PCR Bead<br><br>Adição:<br>1µl MgCl <sub>2</sub> a 25 mM<br>Primers 0,5 µM | Tampão de PCR                | Tampão de PCR                     |
|  | 1 x                          | 1 x                               |
|  | MgCl <sub>2</sub> 2 mM       | MgCl <sub>2</sub> 2 mM            |
|  | dNTPs 0,2 mM                 | dNTPs 0,2 mM                      |
|  | Primers 0,25 µM              | Primers 0,25 µM                   |
|  | Enzima AmpliTaq Gold 0,625 U | Enzima Expand High Fidelity 1,3 U |

Tabela 4 Programa de amplificação usado para misturas de reacção com enzimas puReTaq™ e AmpliTaq Gold.

| Programa de Amplificação |                      | Enzima puReTaq™<br>e<br>Enzima AmpliTaq Gold |       |
|--------------------------|----------------------|--|-------|
|                          |                      | Temperatura                                  | Tempo |
| 1 ciclo                  | Desnaturação inicial | 95°C   | 5m    |
| 45 ciclos                | Desnaturação         | 95°C   | 45s   |
|                          | Hibridação           | 50°C   | 1m    |
|                          | Síntese              | 72°C   | 1m30s |
| 1 ciclo                  | Extensão final       | 72°C   | 7m    |

m= minuto; s= segundo

**Tabela 5** Programa de amplificação usado para misturas de reacção com enzima Expand High fidelity.

| Programa de Amplificação |                      | Enzima Expand High fidelity |       |
|--------------------------|----------------------|-----------------------------|-------|
|                          |                      | Temperatura                 | Tempo |
| 1 ciclo                  | Desnaturação inicial | 94°C                        | 2m    |
|                          | Desnaturação         | 94°C                        | 30s   |
| 10 ciclos                | Hibridação           | 48°C                        | 1m    |
|                          | Síntese              | 72°C                        | 1m30s |
| 35 ciclos                | Desnaturação         | 94°C                        | 30s   |
|                          | Hibridação           | 50°C                        | 1m    |
|                          | Síntese              | 72°C                        | 1m30s |
| 1 ciclo                  | Extensão final       | 72°C                        | 7m    |

m= minuto; s= segundo.

Os produtos de *PCR* obtidos foram visualizados sob luz ultravioleta, após integração de *SYBR® Safe* (Invitrogen Life Science, California, USA) ou *Gel Red* (*Gel Red™ Nucleic Acid Gel stain*, Hayward, USA) directamente no gel de agarose a 1,3%. Os marcadores moleculares aplicados nos géis de agarose foram *DNA Molecular Weight Marker VIII* (Roche, Roche Diagnostics, Germany) e também *1Kb Plus DNA Ladder* (Invitrogen Life Science, California, USA) no sentido de identificar a banda correspondente ao fragmento amplificado pretendido.

## 6. PURIFICAÇÃO E SEQUENCIAÇÃO DOS PRODUTOS DE *PCR*

De forma a garantir a eliminação do excesso de *primers* e de dNTPs, os produtos obtidos a partir da segunda e terceira reacções de amplificação, foram purificados recorrendo ao kit *QIAquick® PCR Purification* (QIAGEN, Valencia, CA), seguindo as indicações do fabricante. Efectuou-se para alguns casos, após corrida electroforética em gel de agarose, uma extracção do ADN correspondente ao fragmento amplificado pretendido (banda única), através do kit *MinElute® Gel Extraction* (QIAGEN, Valencia, CA), seguindo também as instruções do fabricante. Com o objectivo de confirmar a presença e a concentração dos produtos, uma alíquota de produto purificado foi sujeita a uma electroforese em gel de agarose a 1% nas condições já descritas anteriormente.

Os produtos purificados foram sequenciados recorrendo ao método de terminação em cadeia, utilizando os *primers* internos e o kit *Big Dye™ Terminator Cycle Sequencing Ready Reaction* (PE Applied Biosystems, Foster City, USA). A sequenciação cíclica incluiu 25 ciclos com desnaturação a 96°C durante 30 segundos, hibridação a 50°C durante 15 segundos e um passo de síntese a 60°C durante 4 minutos. Os produtos obtidos foram entregues à Unidade de Tecnologia e Inovação do INSA, onde foi realizado o processo de precipitação e electroforese capilar para obtenção das respectivas sequências.

## 7. ANÁLISE DE SEQUÊNCIAS NUCLEOTÍDICAS

### 7.1. Obtenção das Sequências Consenso

As sequências consenso para cada amostra e para cada região genômica (*env* e *nef*) foram construídas a partir do conjunto de sequências nucleotídicas obtidas com o auxílio do programa *ClustalX 2.0.10*<sup>80</sup>. As sequências foram alinhadas a partir do *BioEdit 7.0.9.0*<sup>81</sup>, com pequenos ajustes manuais. A cada sequência consenso foi atribuído um sistema de numeração iniciado pela letra “G” seguido pelo número de amostra e pelas letras “E” (de *env*) ou “N” (de *nef*) de acordo com a região genômica. As sequências consenso obtidas neste estudo foram submetidas à base de dados *GenBank* e esperam respectivos números de acesso.

### 7.2. Sequências de Referência

As sequências de referência usadas neste estudo foram retiradas da base de dados de VIH de *Los Alamos*<sup>30</sup> e utilizadas na construção das árvores filogenéticas e na análise de *bootscanning*. Os respectivos números de acesso são os seguintes: **sub-subtipo A1** (AF004885, AB253421, DQ676872); **sub-subtipo A2** (AF286238, AF286237); **sub-subtipo A3** (AY531630, AY521629); **sub-subtipo A4** (AM000053, AM000054); **subtipo B** (K03455, AY173951); **subtipo C** (U46016, AY772699); **subtipo D** (K03454, AY371157, U88824); **sub-subtipo F1** (AF077336, AF005494); **subtipo G** (AF061641, U88826); **subtipo H** (AF190127, AF005496); **subtipo J** (AF082394, AF082395, EF614151); **subtipo K** (AJ249235, AJ249239); **CRF02\_AG** (AY271690, L39106, AB286856, AB286857, AB286859, AB286862, DQ168578, DQ168577, AF063223, AF063224); **CRF03\_AB** (AF414006, AF193276); **CRF06\_cpx** (AY535659, AF064699); **CRF09\_cpx** (AJ866553, AY093605); **CRF11\_cpx** (AY371151, AY371150, AF492624, AF492623, AJ291718); **CRF13\_cpx** (AF460972, AF460974); **CRF14\_BG** (AF450096, AF450097) e o *outgroup* **SIVcpzMT145** (DQ373066).

### 7.3. Inferência Filogenética

Todas as sequências consenso relativas às duas regiões, *env* e *nef*, foram submetidas inicialmente à análise por *BLAST*, através da base de dados de VIH de *Los Alamos*<sup>30</sup>. Para análise filogenética, as sequências de cada indivíduo foram alinhadas, através do programa *ClustalX 2.0.10*<sup>80</sup> com sequências de referência, representativas de todos os subtipos e formas genéticas importantes na infecção VIH/SIDA. De forma a fazer alguns ajustes manuais e remover gaps recorreu-se ao programa *BioEdit 7.0.9.0*<sup>81</sup>. As árvores filogenéticas foram construídas recorrendo ao método de junção de vizinhos (*Neighbor-Joining*)<sup>82</sup> e as distâncias genéticas foram calculadas com base no modelo de evolução *Kimura*<sup>83</sup> a dois parâmetros. A robustez da topologia das árvores foi avaliada através da execução de testes de *bootstrapping*, com base em 1000 amostragens aleatórias, cujos

agrupamentos com valores de *bootstrap* iguais ou superiores a 70% foram considerados válidos. Todos os métodos referidos foram realizados através do programa MEGA 4<sup>84</sup>.

#### 7.4. Análise da Hipermutação

Recorreu-se à ferramenta *Hypermut 2.0*<sup>84</sup> com o objectivo de investigar a ocorrência de eventos de hipermutação, num conjunto de sequências cujos ramos na árvore filogenética eram longos. O fenómeno de hipermutação é induzido pela enzima celular APOBEC e baseia-se numa elevada taxa de transições G→A, e por isso o programa descreve a natureza e o contexto das substituições de nucleótidos num conjunto de sequências relativamente a uma sequência de referência. A importância desta análise está relacionada com o facto de os vírus hipermutantes apresentarem uma taxa de evolução bastante elevada, porém são incapazes de se replicar.

#### 7.5. Detecção de Recombinação Genética

A análise de sequências por *bootscanning* tem como objectivo investigar a ocorrência de eventos de recombinação e identificar quais as formas genéticas envolvidas, sendo aplicada a casos cuja localização na árvore filogenética origina uma classificação controversa.

Numa primeira fase foi utilizada a ferramenta bioinformática *Viral Genotyping* (<http://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi>), com o intuito de determinar quais as potenciais formas genéticas envolvidas nos eventos de recombinação. De seguida, recorreu-se ao programa *SimPlot 3.5.1*.<sup>85</sup> para analisar tanto as sequências de interesse como as sequências de referência representativas de várias formas virais. A análise *Simplot*, permitiu determinar o grau de semelhança entre as sequências, as quais foram analisadas posteriormente por *bootscanning*, no sentido de se obter informação acerca dos possíveis eventos de recombinação e também do tamanho dos fragmentos que formam o respectivo mosaico genético.

#### 7.6. Distâncias Nucleotídicas

Com o auxílio do programa Mega 4<sup>84</sup> foram construídas matrizes, tendo em conta as duas regiões genómicas estudadas, V3-V5 de *env* e *nef*, com o objectivo de determinar as distâncias intra e inter-subtipo na população de sequências em análise correspondentes aos indivíduos infectados em estudo.

### 8. ANÁLISE DE SEQUÊNCIAS AMINOACÍDICAS

As sequências consenso em estudo para a região V3-V5 e para o gene *nef* foram traduzidas com base no código genético universal, utilizando o programa *Gene Runner 3.05* (*Hastings Software, Inc., Hastings, USA*). Posteriormente as sequências consenso de cada indivíduo e para cada região

genómica, foram alinhadas com o programa *Clustal X 2.0.10*<sup>80</sup> e foi construído o respectivo consenso, com recurso ao programa *BioEdit 7.0.9.0*<sup>81</sup>.

Foi dada especial atenção à conservação ou disrupção dos domínios estruturais e funcionais descritos na literatura como importantes na infecção VIH-1.

Para cada sequência analisada, foi calculado, com recurso à ferramenta *N-Glycosite*<sup>86</sup>, disponível na base de dados de VIH de *Los Alamos*<sup>30</sup>, o número total de potenciais locais de N-glicosilação, caracterizados pelo motivo NxS/T, presentes na região V3-V5 de Env.

## 9. ANÁLISE ESTATÍSTICA

Para a realização do tratamento estatístico recorreu-se ao programa PASW Statistics (*Statistical Package for the Social Sciences*) versão 18.0 e ao apoio do Departamento de Epidemiologia do INSA para esta área.

Numa primeira abordagem, foram utilizadas estatísticas descritivas, tais como contagens e percentagens. Representaram-se alguns indicadores de forma gráfica (gráfico de barras). Para todos os testes estatísticos efectuados foi considerado um nível de significância de 5%.

Foi estudada uma eventual associação entre algumas variáveis consideradas, nomeadamente a infecção VIH com a tuberculose recorrendo ao Teste Qui-Quadrado ( $\chi^2$ ). Como alternativa utilizou-se o Teste Exacto de Fisher, quando os valores esperados nas células da tabela eram inferiores a 5 ou no caso de se tratar de tabelas 2x2. Calculou-se também a medida de Risco Relativo (RR) e os seus respectivos intervalos de confiança a 95%.

## **IV.RESULTADOS**

### **1. DETECÇÃO DE VIH EM EXTRAÍDOS DE ADN POR PROCEDIMENTOS DISTINTOS**

Foram comparados resultados obtidos da amplificação de um grupo de amostras VIH-1 positivas (amostra de conveniência) extraídas pelo método de extracção tradicional (coluna *QIAmp DNA Mini Kit*) e segundo o método de *Chelex* com o objectivo de conhecer o desempenho de um método alternativo de extracção de ADN a utilizar em amostras de sangue colhido em suporte de papel de filtro e incluídas no presente estudo. Verificou-se com este método alternativo e para as mesmas amostras (n=40), um maior sucesso de amplificação relativamente ao procedimento tradicional. De facto, nestas amostras de difícil amplificação, obteve-se resultados positivos para a detecção de VIH-1 em 75% (n=30) das amostras extraídas pelo método de *Chelex*, comparativamente, a 52% (n=21) das amostras extraídas pelo procedimento tradicional.

### **2. DIAGNÓSTICO DA INFECÇÃO VIH NA AMOSTRA EM ESTUDO**

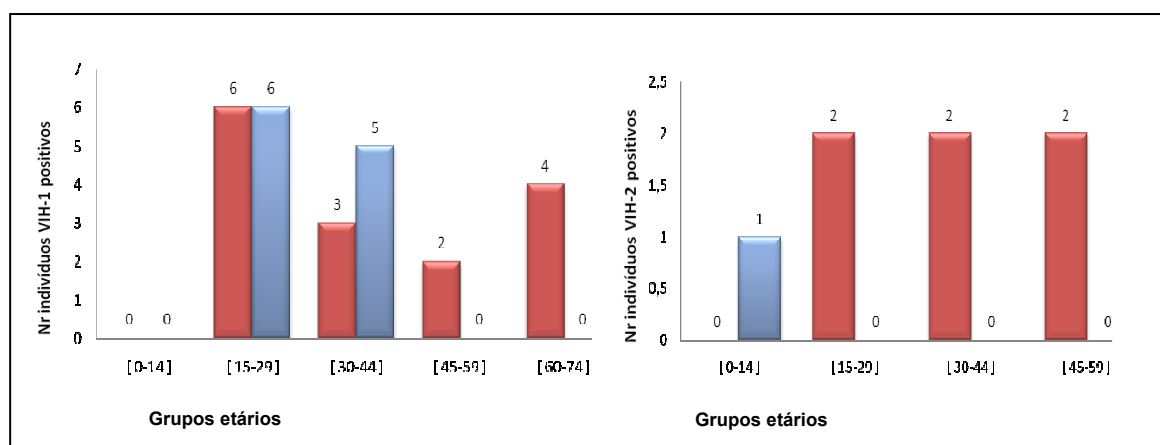
A extracção de ADN foi implementada no grupo de amostras em estudo proveniente da Guiné-Bissau após a verificação de que o método alternativo apresentava os requisitos necessários. Assim, foi realizada a extracção do ADN celular e posteriormente, a pesquisa de ADN proviral de VIH-1 e/ou VIH-2 pela técnica de nested PCR em 123 amostras, correspondentes a casos que nunca tinham sido testados para VIH (n=107) e também a casos cujo diagnóstico positivo da infecção já tinha sido previamente realizado na Guiné-Bissau (n=16). As amostras com resultados VIH negativos (n=42), cujos testes de diagnóstico foram realizados na Guiné-Bissau, não foram submetidas a qualquer procedimento laboratorial prático e assumidas como sendo casos livres desta infecção. Assim, a prevalência de infecção VIH-1 na população estudada (n=165) foi de 15,8% (n=26) e a prevalência de VIH-2 foi de 4,2% (n=7). De acordo com informação disponibilizada referente à tuberculose, 15 indivíduos infectados por VIH estavam também infectados com membros do complexo *M.tuberculosis*, o que corresponde a uma prevalência de co-infecção de 9,1% na população alvo. A análise da eventual associação entre a infecção VIH e a infecção por membros do CMTB não mostrou ser estatisticamente significativa ( $\chi^2=0,055$ ;  $p=0,848$ ). Também foi realizada uma estimativa de risco que não revelou ser estatisticamente significativa entre a infecção TB e o risco de aquisição de infecção VIH, RR=1,2 (IC95%, 0,7; 1.6).

Relativamente aos indivíduos não infectados por VIH-1 e/ou VIH-2 (n=132), 53% pertencia ao sexo masculino (n=71), 45% pertencia ao sexo feminino (n=59) e em 2 indivíduos (2%) não era conhecida esta informação. A média de idades encontrada foi de 36,6 anos e a mediana de 32 anos. Bissau era a residência da maioria dos indivíduos (n=90; 67%) e cerca de 48% dos indivíduos (n=63) encontrava-se infectado por membros do CMTB.

### 3. CARACTERIZAÇÃO DOS CASOS COM INFECÇÃO VIH-1 OU VIH-2

Relativamente aos casos de infecção VIH-1 (n=26), verificou-se que 15 indivíduos pertenciam ao sexo feminino (n=15; 58%) e 11 indivíduos pertenciam ao sexo masculino (n=11; 42%) e que a média de idades encontrada foi de 35,3 anos e a mediana de 31 anos.

A maioria dos indivíduos (n=19; 72%) vivia na cidade de Bissau. Apenas 7 indivíduos viviam, cada um deles, nas cidades Cacheu, Bandim, Biombo, Mansôa, Nhacra, Quebo e Gabú (n=7; 28%). Metade dos casos com diagnóstico de infecção VIH-1 (n=13; 50%), estavam co-infectados por membros do CMTB.

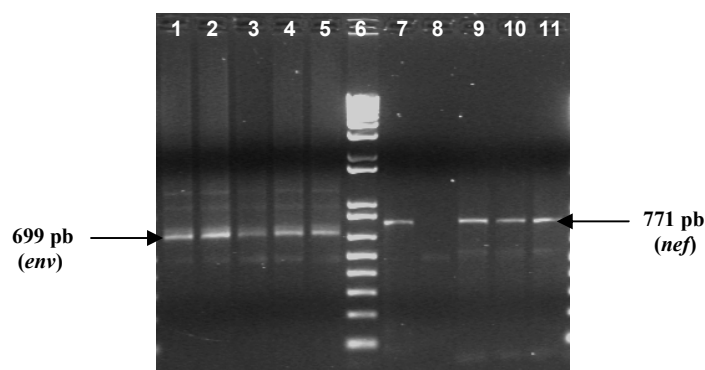


**Figura 10** Distribuição dos indivíduos VIH-1 positivos (n=26) à esquerda e VIH-2 positivos (n=7) à direita por grupo etário e género. Barras rosas- feminino; Barras azuis- masculino.

Do total da população estudada, apenas em 7 indivíduos foi confirmada a infecção VIH-2, sendo 6 indivíduos do sexo feminino (86%) e 1 do sexo masculino (14%). A média de idades encontrada foi de 34 anos e a mediana de 31 anos. Seis indivíduos residiam na cidade de Bissau e 1 indivíduo vivia em Oio. A co-infecção VIH-2 e tuberculose foi confirmada em 2 casos (29%).

### 4. AMPLIFICAÇÃO E SEQUENCIAÇÃO DAS REGIÕES GENÓMICAS DE VIH-1

Todas as amostras VIH-1 positivas (n=26) foram sujeitas a um *nested* PCR com o objectivo de amplificar a região C2V3V5 de *env* e o gene *nef*. A amplificação das regiões genómicas alvo foi conseguida com sucesso em 88,5% (n=23) dos casos para *env* e em 69,2% (n=18) dos casos para *nef*. Na figura 11 estão exemplificados alguns casos de produtos amplificados de ambas as regiões analisadas.



**Figura 11** Visualização de produtos de PCR da região *env* (à esquerda) e do gene *nef* (à direita). Nos poços 1 a 5 foram aplicados produtos correspondentes a um fragmento de 699 pb, no poço 6 foi aplicado o marcador de massa molecular *1Kb Plus DNA Ladder* (Invitrogen Life Science, California, USA), nos poços de 7 a 11 foram aplicados produtos correspondentes a um fragmento de 771 pb e no poço 8, a aplicação corresponde a uma amostra que não amplificou.

Após a amplificação, todos os produtos de PCR foram purificados e sequenciados. Contudo, nem para todos os casos foi obtido um cromatograma de leitura molecular viável. Assim, foi apenas possível construir 21 sequências consenso para a região *env* e 11 sequências consenso para o gene *nef*, cada sequência associada a um indivíduo em estudo. Posteriormente, todas as sequências foram submetidas à base de dados *GenBank* ([www.ncbi.nlm.nih.gov/genbank/](http://www.ncbi.nlm.nih.gov/genbank/)).

## 5. ANÁLISE MOLECULAR DA REGIÃO V3-V5 DE *ENV* DE VIH-1

Numa primeira fase, as sequências relativas à região *env* assim como as sequências relativas ao gene *nef* foram submetidas a uma análise prévia por BLAST, recorrendo à base de dados de VIH de *Los Alamos*. Esta primeira abordagem permitiu orientar a escolha de algumas sequências de referência eventualmente mais indicadas para utilizar nas posteriores análises a efectuar. Foram seleccionadas sequências de referência correspondentes a todos os subtipos conhecidos e a formas recombinadas de importância na infecção VIH/SIDA. Para além das variantes indicadas pela análise por BLAST, foi tomada especial atenção a sequências de referência de casos de infecção ocorridos na Guiné-Bissau. A sequência *SIVcpzMT145*, foi utilizada como *outgroup* na construção das árvores filogenéticas.

### a) Sequências nucleotídicas

As sequências nucleotídicas obtidas para a região V3-V5 de *env* variaram entre os 513 pb e os 623 pb, onde as diferenças no comprimento, foram resultantes do grau de definição dos picos obtidos na parte final dos cromatogramas analisados.

Na construção da árvore filogenética foram comparados fragmentos com 540pb correspondentes às 21 sequências em estudo e 41 sequências de referência seleccionadas. Os agrupamentos representados pelas sequências de referência correspondentes a subtipos distintos foram suportados por valores de *bootstrap* iguais ou superiores a 70% e apresentaram-se de forma clara e definida na árvore como se pode observar na Figura 14.

O agrupamento correspondente ao subtipo A apresentou-se bem definido na árvore, suportado por um valor de 70% de *bootstrap* em que sequências de referência A1, A2, A3 e A4 agruparam de forma bastante robusta, todas apoiadas por valores de *bootstrap* elevados (respectivamente, 70%, 100%, 100% e 99%). Dentro do agrupamento A, as sequências G7E, G188E, G10E, G13E e G14E, formaram um *cluster* isolado (100% de *bootstrap*), bem como as sequências G81E e G82E (100% de *bootstrap*). As demais sequências (G80E, G40E, G50E, G96E, G61E, G194E, G38E, G12E, G2E, G11E, G9E, G79E) distribuíram-se de uma forma mais heterogênea dentro do agrupamento. Nenhuma das sequências em causa agrupou dentro do cluster definido para cada sequência de referência (A1, A2, A3 ou A4) indicando a potencial presença de sequências não puras, levantando-se a hipótese de ocorrência de eventos de recombinação intragénica ao nível desta região. A título de exemplo, no caso da sequência G10E, a informação obtida através da ferramenta *Viral Genotyping* indicou o subtipo B e a CRF02\_AG como as potenciais formas genéticas envolvidas. Consequentemente, na construção do ficheiro a ser analisado no programa *Simplot*<sup>85</sup>, foram incluídas sequências adicionais importantes na pandemia VIH/SIDA. A análise por *bootscanning* revelou a ocorrência de um único evento de recombinação, confirmando a presença das duas formas virais previamente indicadas, verificando-se que a sequência G10E era constituída por um primeiro segmento, com cerca de 250 pb, classificado como eventual subtipo B e por um segundo segmento com cerca de 190 pb, classificado como eventual CRF02\_AG (Figura 12). Curiosamente, para a sequência G2E que se encontrava num sub-agrupamento distinto da G10E, também foram obtidos resultados semelhantes relativamente ao envolvimento do subtipo B e da CRF02\_AG. O primeiro segmento foi classificado como CRF02\_AG, com cerca de 270 pb e o segundo segmento classificado como subtipo B, com aproximadamente 170 pb (Figura 13).

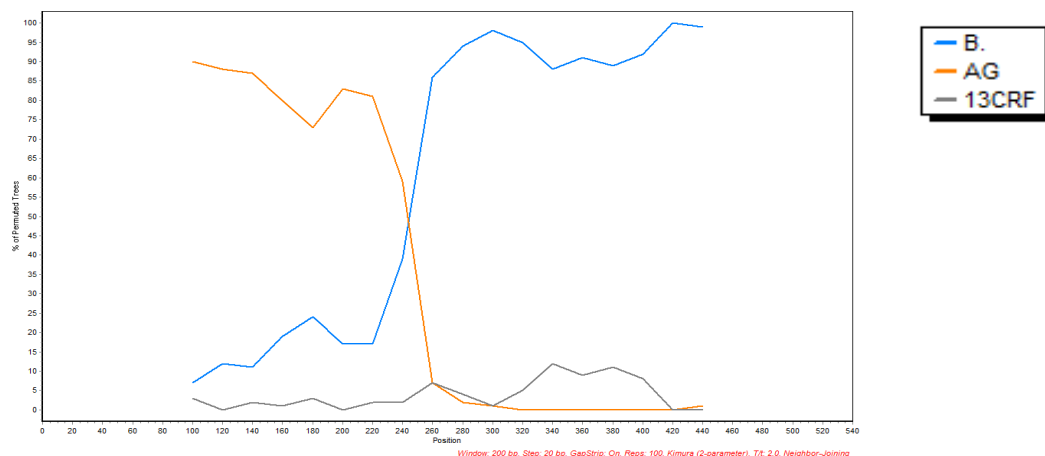
Para todas as restantes sequências foi feita a mesma abordagem (ferramenta *Viral Genotyping* e análise *bootscanning*) que revelou associação com a CRF02\_AG (dados não apresentados).

No que diz respeito às sequências G88E e G196E, estas agruparam dentro do *cluster* definido para o subtipo G, suportadas por um elevado valor de *bootstrap*, correspondente a 90%.

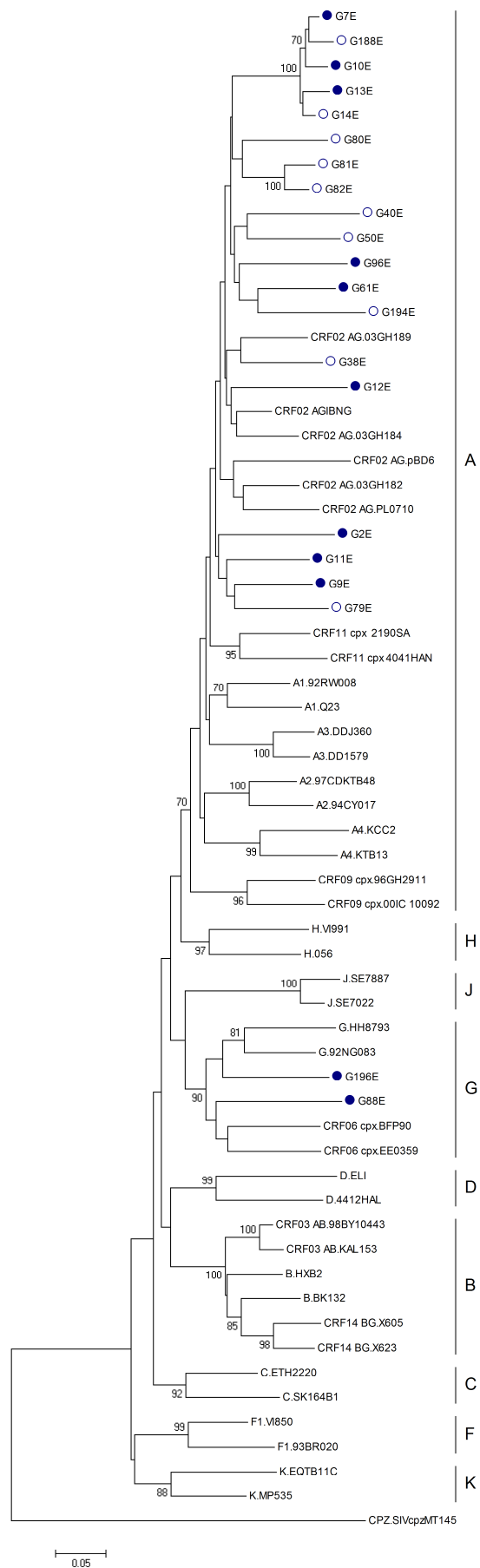
Apesar da presença possível de outros subtipos no genoma, a caracterização das sequências correspondente à região *env* permitiu, assim, classificar 19 (90%) delas como agrupando com o subtipo A e apenas 2 (10%) sequências como agrupando com o subtipo G.



**Figura 12** Análise por *bootscanning* da região V3-V5 da sequência G10E. A vermelho estão indicados os parâmetros utilizados e os Subtipos e CRFs de referência incluídos na análise encontram-se mencionados na respectiva legenda.



**Figura 13** Análise por *bootscanning* da região V3-V5 da sequência G2E. A vermelho estão indicados os parâmetros utilizados e os Subtipos e CRFs de referência incluídos na análise encontram-se mencionados na respectiva legenda.



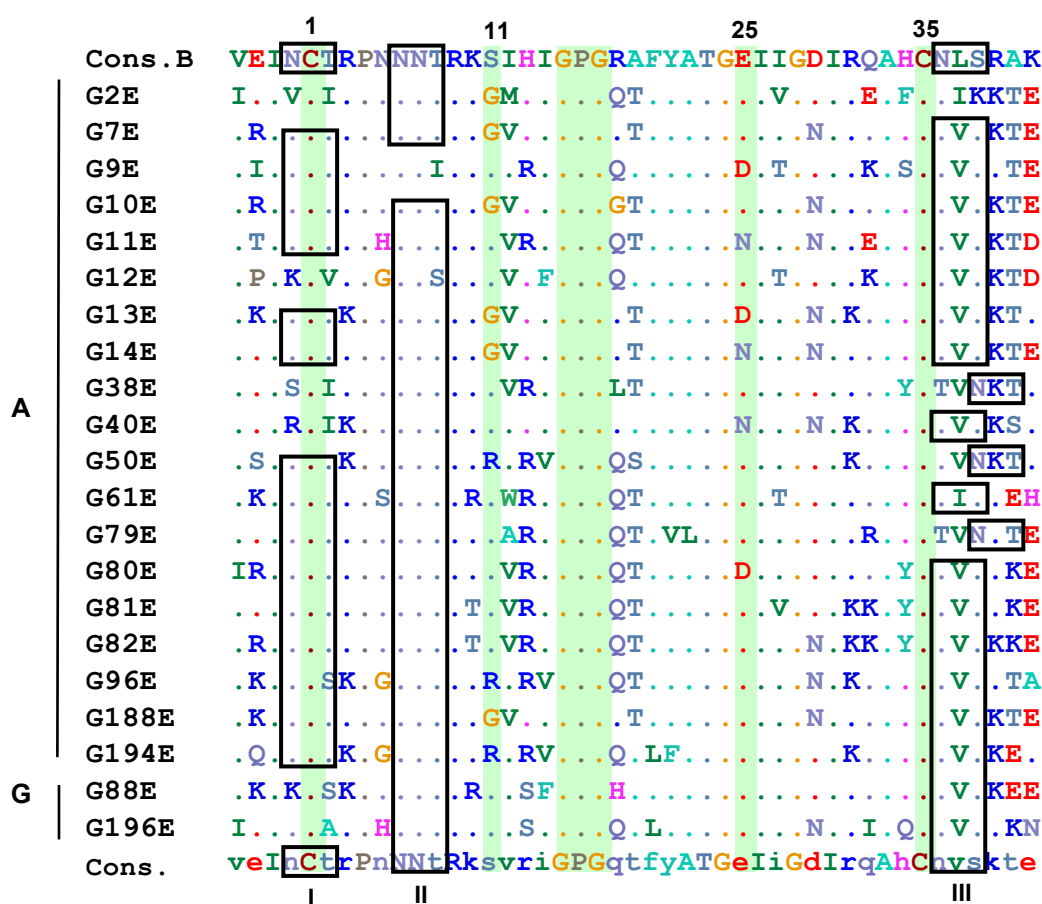
**Figura 14** Árvore filogenética com base na região V3-V5 de *env*. As sequências em estudo encontram-se assinaladas com círculos, correspondendo os escuros a casos de co-infecção VIH/TB (●).

## b) Sequências Aminoacídicas

A análise aminoacídica foi realizada com base no alinhamento múltiplo das sequências aminoacídicas da região V3, que inclui sequências de 35 resíduos e que se encontra representado na Figura 15.

O motivo estrutural composto por duas cisteínas, C<sub>1</sub> e C<sub>35</sub>, responsável pela conformação em *loop* da região V3<sup>87</sup>, apresentou-se totalmente conservado. Da mesma forma, o motivo GPG, também se encontrou conservado em todas as sequências estudadas.

Na análise global da região V3-V5 de Env, verificou-se a existência de alguns polimorfismos. A título de exemplo observou-se a existência de um resíduo de glutamina (Q) na posição 18 na maioria dos casos envolvidos no estudo (n=13). Observou-se também os resíduos G, H e L em três sequências e o resíduo R nas 5 restantes sequências.



**Figura 15** Alinhamento parcial da região C2V3V5. As sequências foram comparadas com uma sequência do subtipo B (Cons. B, em cima), disponível na base de dados de VIH *Los Alamos*<sup>30</sup>. O consenso das sequências estudadas (Cons., em baixo) encontra-se também representado. Os pontos indicam semelhança entre as sequências. As sequências estão ordenadas de acordo com a sua classificação viral (à esquerda). Os retângulos assinalam os locais de N-glicosilação (I, II e III) e a sombreamento estão os motivos analisados (C<sup>1</sup>, GPG e C<sup>35</sup>).

Analisando os potenciais locais de N-glicosilação (NxS/T), verificou-se uma elevada conservação no conjunto das sequências em estudo (Figura 15).

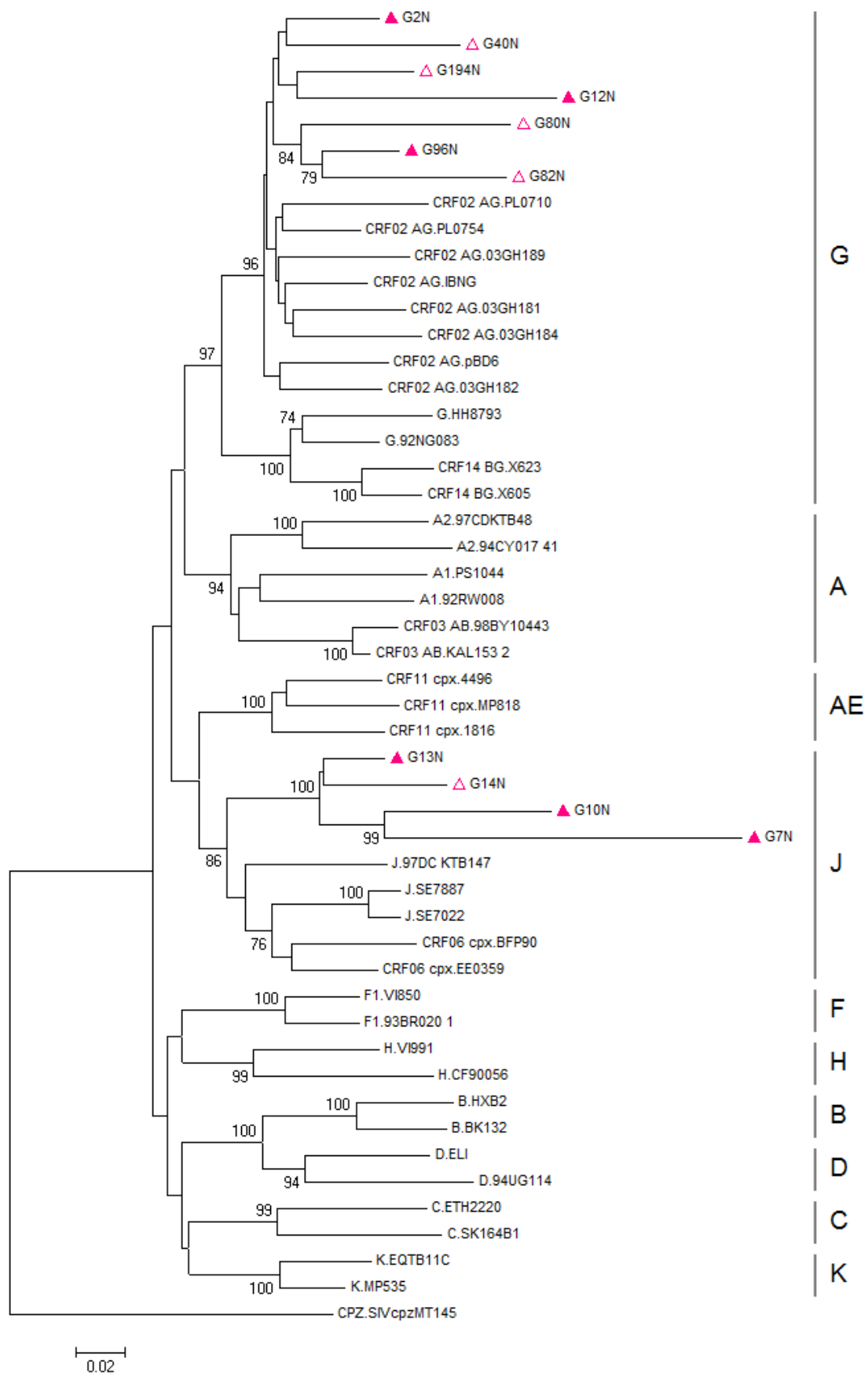
Com exceção de 6 casos (G2E, G12E, G38E, G40E, G88E e G196E), o local de N-glicosilação centrado na C<sub>1</sub> (I) foi observado na maioria dos casos. Um segundo local de N-glicosilação constituído pelo motivo NNT (II), apresentou uma elevada conservação. Apenas a sequência G12E apresentou a forma NNS. Na sequência G9E não foi observada a conservação deste motivo (NNI). O terceiro potencial local de N-glicosilação (III), apresentou-se conservado na forma NVS ou NIS para 17 sequências, na forma NKT para duas sequências e na forma NRT para uma sequência. Na sequência G2E não se observaram dois dos três locais de N-glicosilação (I e III) relativamente ao que seria de esperar.

## 6. ANÁLISE MOLECULAR DA REGIÃO NEF DE VIH-1

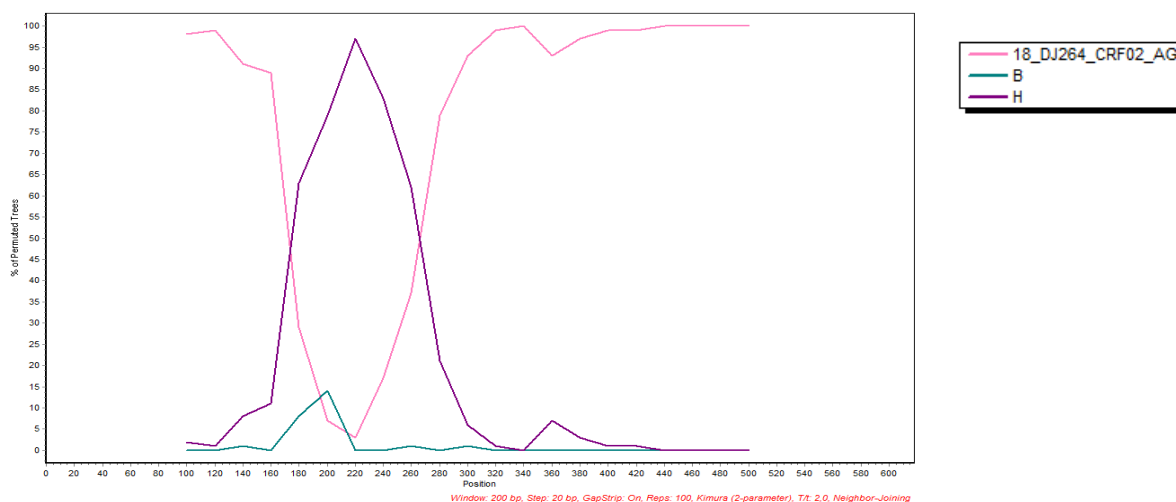
### a) Sequências nucleotídicas

As sequências obtidas para o gene *nef* variaram entre os 573 pb e os 633 pb, dependente da leitura dos cromatogramas. A árvore filogenética foi construída com base na comparação de fragmentos com 621 pb das 11 sequências em estudo e de 39 sequências de referência previamente seleccionadas (Figura 16). Todos os agrupamentos relativos aos subtipos e formas geneticamente recombinadas, representados pelas respectivas sequências de referência, apresentaram-se robustos e bem delimitados na árvore filogenética, suportados por valores de *bootstrap* iguais ou superiores a 74%. A topologia da árvore revelou a existência de longos ramos representando distâncias nucleotídicas elevadas, nomeadamente observado no caso da sequência G7N. Deste modo, investigou-se a eventual ocorrência de hipermutação, em todas as sequências analisadas. Porém os resultados indicaram que nenhuma sequência evidenciou estar hipermutada ( $p\text{-value} > 0,05$ ) sendo mantidas no estudo.

O subtipo G apresentou um agrupamento bastante robusto, suportado por um valor de 97% de *bootstrap*. Dentro deste, as sequências G2N, G40N, G194N, G12N, G80N, G96N e G82N, agruparam com sequências de referência para a CRF02\_AG, apoiadas por valor de *bootstrap* de 96%. Foi investigada a ocorrência de eventos de recombinação para estas sequências pois nenhuma delas agrupou dentro do cluster definido para sequências de referências do subtipo G. Numa primeira abordagem, e a para a sequência G2N foi mais uma vez utilizada a ferramenta *Viral Genotyping* que revelou o subtipo H e a CRF02\_AG como as formas genéticas provavelmente envolvidas nos eventos de recombinação. A análise por *bootscanning* confirmou a presença das duas formas previamente indicadas, mostrando três eventos de recombinação, uma vez que o primeiro segmento de G2N é semelhante a CRF02\_AG (cerca de 173 pb), o segundo segmento pertencente ao subtipo H (aproximadamente 92 pb) e o terceiro segmento semelhante novamente a CRF02\_AG com cerca de 265 pb (Figura 17).



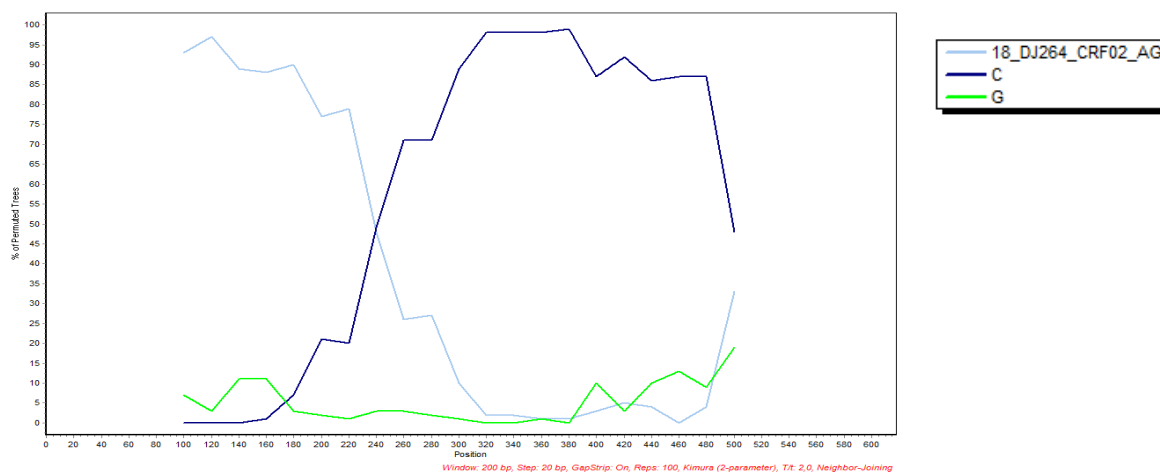
**Figura 16** Árvore filogenética com base na região *nef*. As sequências em estudo encontram-se assinaladas por triângulos, em que os escuros correspondem aos casos de co-infecção VIH/TB (▲).



**Figura 17** Análise por *bootscanning* da região *nef* da sequência G2N. A vermelho estão indicados os parâmetros utilizados e os Subtipos e CRFs de referência incluídos na análise encontram-se descritos na legenda.

A existência de eventos de recombinação também foi investigada na sequência G194N. Como se pode observar na Figura 18, esta sequência é constituída por um primeiro segmento com aproximadamente 245 pb, classificado como CRF02\_AG e por um segundo segmento, com sensivelmente 255 pb, semelhante ao subtipo C.

Para todas as restantes sequências foi feita a mesma abordagem (ferramenta *Viral Genotyping* e análise *bootscanning*) que revelou associação com a CRF02\_AG (dados não apresentados).



**Figura 18** Análise por *bootscanning* da região *nef* da sequência G194N. A vermelho estão indicados os parâmetros utilizados e os Subtipos e CRFs de referência incluídos na análise encontram-se descritos na legenda.

O agrupamento correspondente ao subtipo J, apresentou-se bem definido, suportado por um valor de 86% de *bootstrap*. Neste agrupamento, as sequências G13N, G14N, G10N e G7N, formaram um cluster isolado das restantes sequências de referência e com um valor de *bootstrap* de 100%.

Apesar da presença possível de outros subtipos na constituição do genoma viral, a análise geral para a região *nef* permitiu, assim, classificar 7 (64%) sequências como agrupando com o subtipo G e 4 (36%) sequências como agrupando com o subtipo J.

### b) Sequências Aminoacídicas

A Figura 19 representa um alinhamento parcial da região Nef contendo a maioria dos motivos estruturais e funcionais da proteína, descritos na literatura, com importância biológica na infecção VIH<sup>68</sup>. Este alinhamento foi comparado com uma sequência consenso do subtipo B do VIH-1.

O sinal de mistilação (MGxxxS<sub>1</sub>), importante na associação de Nef com a membrana plasmática, apresentou-se totalmente conservado assim como o resíduo de metionina (M<sub>20</sub>)<sup>68</sup>.

Imediatamente a jusante, no motivo da hélice  $\alpha$  descrito na literatura como RRAE<sub>21</sub><sup>88</sup>, foi observado em todos os casos na forma RQNP. O motivo WL<sub>57</sub> revelou estar totalmente conservado. Relativamente ao motivo ácido EEEE<sub>62</sub>, foi observado na forma EGEE para todas as sequências.

O domínio de tetraprolina PxxPxxPxRP<sub>69</sub> foi encontrado conservado na totalidade das sequências, assim como o alvo de fosforilação pela cinase PKC, caracterizado pela sequência RPMTYK<sub>77</sub>.

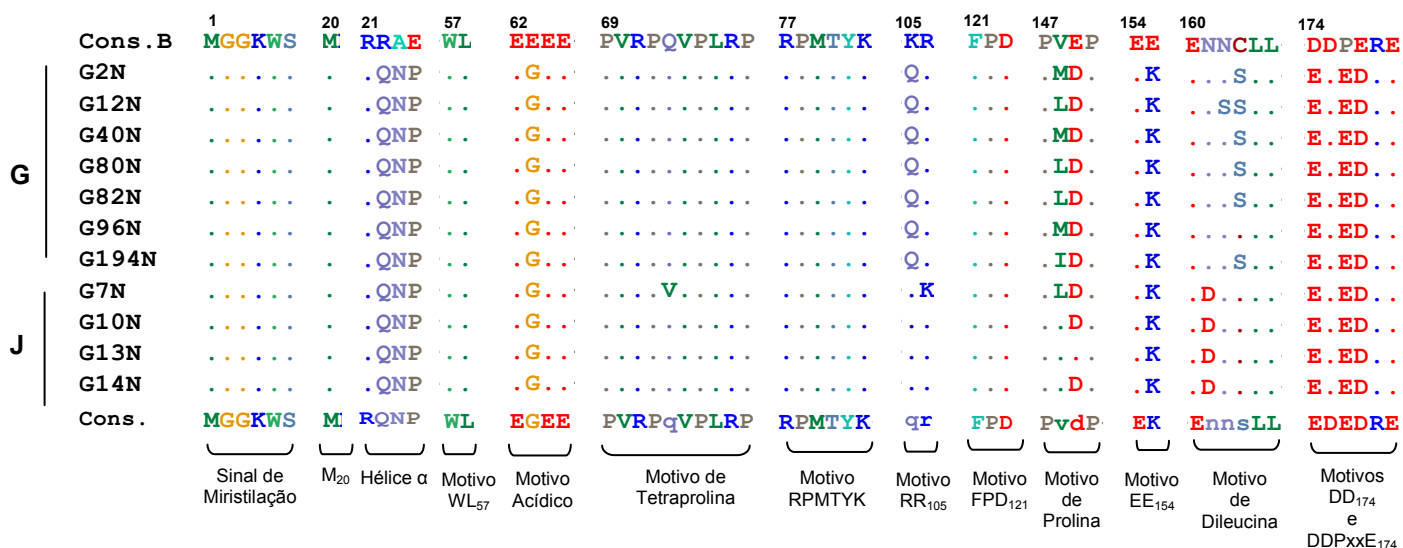
No que se refere ao motivo RR<sub>105</sub>, implicado na ligação à PAK, verificou-se uma semi-conservação, uma vez que o primeiro resíduo do motivo apresentou uma alteração R<sub>105</sub>→Q, em 7 das 11 sequências analisadas.

O motivo FDP<sub>121</sub>, envolvido na ligação à tioesterase humana e o motivo de prolina, PxxP<sub>147</sub>, situado numa zona mais próxima do terminal carboxílico de Nef, apresentaram 100% de conservação.

O motivo diacídico EE<sub>154</sub> não se encontrou conservado em nenhum dos casos, uma vez que se verificou a alteração do segundo aminoácido E<sub>155</sub>→K em todas as sequências, traduzindo-se na modificação de um resíduo ácido para básico.

O domínio E/DxxxLL<sub>160</sub> foi mais um motivo que se apresentou totalmente conservado em todos os casos.

No motivo DDPxxE<sub>174</sub>, envolvido na ligação à cinase c-Raf1<sup>89</sup> foram detectadas alterações como, D<sub>174</sub>→E e P<sub>176</sub>→E.



**Figura 19** Alinhamento parcial contendo os motivos funcionais da proteína Nef. As sequências foram comparadas com uma sequência do subtipo B (Cons. B, em cima), retirada da base de dados de VIH. O consenso das sequências estudadas (Cons., em baixo) encontra-se também representado. Os pontos indicam semelhança entre as sequências. As sequências estão ordenadas de acordo com a sua classificação viral (à esquerda).

## 7. ANÁLISE MOLECULAR CONJUNTA DAS REGIÕES V3-V5 DE ENV E NEF

Tendo em conta as classificações virais obtidas para as duas regiões genómicas em estudo procedeu-se a uma análise conjunta dos dados.

Esta análise compreendeu os casos em que foi possível obter em simultâneo informação molecular para as duas regiões genómicas e estes encontram-se descritos na tabela 6.

Todas as sequências virais apresentaram uma classificação discordante relativamente às duas regiões analisadas. Assim, observou-se a existência de vírus classificados como geneticamente recombinados entre subtipos A e G e também entre os subtipos A e J. Embora tenha sido possível inicialmente associar as sequências em estudo a um subtipo, os resultados obtidos sugerem que estas sequências podem apresentar um genoma mosaico, provavelmente resultado de uma recombinação entre uma forma recombinante AG com envolvimento de outros subtipos virais (H, J, C, B) e não necessariamente na sua forma pura.

A análise por *bootscanning* nas sequências G2N e G194N revelou a presença de outros subtipos, tais como o subtipo H e o subtipo C e nas sequências G2E e G10E, o subtipo B.

Tabela 6 Classificação por Filogenia e *bootscanning* das sequências, com base nas duas regiões genômicas.

| Classificação das Sequências |                  |                  |
|------------------------------|------------------|------------------|
| Amostra                      | Região V3_V5     | Gene <i>nef</i>  |
| G2                           | A <sup>1,2</sup> | G <sup>2,4</sup> |
| G7                           | A <sup>1,2</sup> | J                |
| G9                           | A <sup>2</sup>   | -                |
| G10                          | A <sup>1,2</sup> | J                |
| G11                          | A <sup>2</sup>   | -                |
| G12                          | A <sup>2</sup>   | G <sup>2</sup>   |
| G13                          | A <sup>1,2</sup> | J                |
| G14                          | A <sup>1,2</sup> | J                |
| G38                          | A <sup>2</sup>   | -                |
| G40                          | A <sup>2</sup>   | G <sup>2</sup>   |
| G50                          | A <sup>2</sup>   | -                |
| G61                          | A <sup>2</sup>   | -                |
| G79                          | A <sup>2,3</sup> | -                |
| G80                          | A <sup>2</sup>   | G <sup>2</sup>   |
| G81                          | A <sup>2</sup>   | -                |
| G82                          | A <sup>2</sup>   | G <sup>2</sup>   |
| G88                          | G                | -                |
| G96                          | A <sup>2</sup>   | G <sup>2</sup>   |
| G188                         | A <sup>1,2</sup> | -                |
| G194                         | A <sup>2</sup>   | G <sup>2,5</sup> |
| G196                         | G                | -                |

Todas as sequências do subtipo G (região *nef*) agruparam na árvore filogenética com sequências CRF02\_AG;

<sup>1</sup> *Bootscanning* revelou envolvimento com o subtipo B; <sup>2</sup> *Bootscanning* revelou envolvimento com a CRF02\_AG; <sup>3</sup> *Bootscanning* revelou envolvimento com o subtipo A1; <sup>4</sup> *Bootscanning* revelou envolvimento com o subtipo H; <sup>5</sup> *Bootscanning* revelou envolvimento com o subtipo C.

## 8. DISTÂNCIAS GENÉTICAS DAS SEQUÊNCIAS VIRAIS

As distâncias nucleotídicas foram calculadas atendendo ao conjunto das sequências obtidas, tendo em conta a sua classificação em cada uma das duas regiões genómicas estudadas. Na tabela 7 encontram-se representados os valores percentais médios das distâncias intra- e inter-subtipos por região genómica analisada.

**Tabela 7** Distâncias nucleotídicas<sup>1</sup> intra- e inter-subtipos nas regiões, V3-V5 de *env* e *nef*.

| <i>env</i> |         |      | <i>nef</i> |         |   |      |
|------------|---------|------|------------|---------|---|------|
|            | Subtipo | A    | G          | Subtipo | G | J    |
| V3-V5      | A       | 19,9 |            | nef     | G | 14,9 |
|            | G       | 30,8 | 24,5       |         | J | 25,8 |

<sup>1</sup>Valores médios em percentagem.

Verificou-se que a região V3-V5 de *env*, apresentou valores de distâncias nucleotídicas médias superiores comparativamente ao gene *nef* e, independentemente do subtipo viral em causa, o que está de acordo com outros estudos que descrevem a elevada evolução deste gene em resultado da maior pressão selectiva comparativamente a *nef*<sup>90,91</sup>.

Relativamente à região V3-V5 de *env*, verifica-se que a distância média intra-subtipo G (24,5) foi superior quando comparada com o subtipo A (19,9), enquanto no gene *nef*, a distância média intra-subtipo G (14,9) foi inferior comparativamente ao subtipo J (16,3).

## V.DISSCUSSÃO

O Vírus da Imunodeficiência Humana apresenta uma enorme variabilidade genética, o que diminui a eficácia dos anti-retrovíricos e dificulta a produção de vacinas. O crescente número de variantes genómicas de VIH-1 observado em circulação na população mundial coloca questões sobre a potencial diferença na transmissibilidade e na virulência dos genótipos virais, e pode conduzir a dificuldades no diagnóstico da infecção. Desta forma, é importante conhecer o padrão de epidemiologia molecular da infecção VIH/SIDA na população e prever as possíveis consequências da sua alteração no tempo.

No presente estudo foram incluídos 165 indivíduos de naturalidade africana, residentes na Guiné-Bissau, 83 (51%) pertencentes ao sexo masculino e 80 (48%) pertencentes ao sexo feminino, com intervalo de idades dos 2 aos 73 anos e em que cerca de metade dos indivíduos (n=78; 47%) estavam infectados por membros do complexo *M. tuberculosis*.

Atendendo a que as amostras incluídas no presente estudo foram recolhidas em suporte de papel de filtro foi necessário testar um método de extracção de ADN mais adequado. Para isso, recorreu-se a uma amostra de conveniência de indivíduos infectados por VIH-1 (n=40), onde foram comparados os resultados de amplificação de fragmentos virais a partir dos extraídos pelo método de extracção de ADN tradicional (coluna, QIAGEN) e pelo método alternativo (método de *Chelex*). Após realização dos procedimentos em paralelo, verificou-se um maior sucesso de amplificação da região LTR do VIH-1 com o método alternativo relativamente ao procedimento tradicional (respectivamente, 75% vs 52%). De salientar, que este grupo de amostras testadas correspondia a indivíduos sob influência de anti-retrovíricos (terapia HAART) possuindo níveis de carga vírica indetectáveis (ARN VIH-1 <LDL), sendo à partida expectável a ocorrência de resultados falsos negativos.

Assim, após a avaliação destes resultados foi aplicado o método de *Chelex* a amostras seleccionadas da população em estudo da Guiné-Bissau. A detecção da infecção VIH-1 e VIH-2 foi feita através da técnica *Nested* PCR com amplificação de fragmentos específicos da região LTR, em iguais condições às anteriormente usadas para o grupo de amostras de conveniência. No total da população em estudo (n=165) foram diagnosticados 26 indivíduos com infecção VIH-1 e 7 indivíduos com infecção VIH-2, correspondendo a uma prevalência no grupo de 15,8% para VIH-1 e 4,2% para VIH-2. Verificou-se que 63 (47,7%) casos estavam infectados com tuberculose e 15 (9,1%) casos co-infectados com VIH e membros do complexo *M. tuberculosis*. Não foram observadas diferenças estatisticamente significativas entre a infecção com CMTB e o risco de aquisição de infecção pelo VIH ( $\chi^2=0,055$ ;  $p=0,848$ ).

Nos casos de infecção VIH-1 confirmada (n=26), procedeu-se à caracterização molecular dos vírus com a amplificação por PCR da região V3-V5 de *env* e gene *nef*. O sucesso da

amplificação da região *env* foi de 88% e do gene *nef* foi de 69%. As razões que podem justificar o insucesso da amplificação obtido podem estar relacionadas com uma reduzida concentração de ADN extraído (valores da absorvância no espectrofotómetro baixos); um reduzido título de ADN proviral nas amostras; alterações nas sequências virais nos locais de hibridação dos *primers* por recombinação genética do vírus; e também, mais provavelmente, uma degradação do ADN resultante das condições do material de colheita do sangue (papel de filtro) e do seu armazenamento e transporte, já que as amostras sofreram uma grande amplitude térmica até ao seu processamento no Laboratório.

Foram construídas 21 sequências consenso para a região *env* e 11 sequências consenso para o gene *nef*, cada uma das respectivas sequências associadas a cada um dos indivíduos em estudo. Com o intuito de classificar as sequências de VIH-1 obtidas foram construídas duas árvores filogenéticas utilizando também sequências de referência seleccionadas da base de dados internacional. A selecção das sequências atendeu à sua origem geográfica e às principais formas virais envolvidas na pandemia da infecção VIH/SIDA, especialmente no continente africano. A construção das árvores filogenéticas baseou-se no método de *Neighbour-Joining*, que é eficiente, rápido e o adequado para grandes conjuntos de múltiplos dados<sup>82</sup>. As distâncias genéticas das sequências virais foram calculadas de acordo com o modelo da substituição de nucleótidos *Kimura* a dois parâmetros, que assume taxas diferentes entre transições (A-G e C-T) e transversões (A-C, A-T, G-C e G-T)<sup>83</sup>.

A análise da árvore filogenética correspondente à região *env* permitiu classificar 19 (90%) sequências como agrupando com o subtipo A e 2 (10%) sequências como agrupando com o subtipo G, respectivamente com valores de *bootstrap* de 70% e 90%. Surgiu a hipótese de ocorrência de eventos de recombinação intragénica ao nível desta região, uma vez que, nenhuma das sequências estudadas (agrupamento A) foi observada em clusters contendo as sequências referência, para o subtipo A (A1 a A4), e apenas algumas formaram clusters isolados entre si, suportados por elevados valores de *bootstrap* (G7E, G188E, G10E, G13E e G14E; 100% *bootstrap*). O facto da maioria das sequências analisadas agruparem sem valores de *bootstrap* significativos e de forma dispersa no agrupamento, embora nos indique que uma porção do genoma analisado possa ter essa origem (subtipo A), também levanta indícios de que possíveis eventos de recombinação tenham ocorrido com outras e diferentes formas genéticas virais resultando em sequências menos semelhantes entre si. A análise por *bootscanning* a algumas sequências permitiu confirmar esta suspeita, revelando o envolvimento de outros subtipos e CRF, como por exemplo o subtipo B e a CRF02\_AG. Na verdade, por exemplo, para a sequência G10E, um primeiro segmento do genoma revelou ser mais semelhante ao subtipo B seguido de um segundo segmento mais semelhante à CRF02\_AG. No entanto, na sequência G2E, embora os resultados sugerissem o envolvimento das mesmas formas virais, B e CRF02\_AG, observou-se uma alteração da sua ordem nos eventos de recombinação. A localização destas duas sequências na árvore filogenética está em acordo com o descrito. Por outro lado, as sequências que agrupam com valor de *bootstrap* de 100% com a sequência G10E, possuem o mesmo padrão de recombinação (B/CRF02AG) (dados não apresentados). Curiosamente, um

estudo de amostras colhidas entre 1994 e 1997 na Guiné-Bissau para caracterização molecular de V3 *env* e RT VIH-1 indicou o domínio completo da forma IbNG-*like* subtipo A/G (subtipo A em V3 e subtipo G em RT). Apenas um caso entre 27 estudados foi classificado de subtipo B<sup>43</sup>.

No presente estudo, a análise de *nef* permitiu classificar 7 (64%) sequências como subtipo G, mais concretamente agrupando com referências CRF02\_AG, e também 4 (36%) sequências classificadas como subtipo J, respectivamente com valores de *bootstrap* de 97%, 96% e 86%. De salientar, que as 7 sequências que foram classificadas de recombinantes AG, localizam-se fora de um cluster com valor de *bootstrap* de 100% que contém as sequências de referência do subtipo G e CRF14\_BG. Adicionalmente, uma sequência (CRF02\_AG IbNG) derivada do estudo molecular VIH-1 realizado em 1999 na Guiné-Bissau<sup>43</sup>, não agrupou de forma evidente com as sequências em análise, sugerindo uma constituição genómica diferente. No entanto, 10 anos de evolução molecular na epidemia VIH/SIDA decorreram na Guiné-Bissau. Na verdade, e para esta região genómica, para além da presença do subtipo J que ainda não tinha sido descrita no país, a análise por *bootscanning* mostrou semelhança de partes das sequências em estudo com o subtipo C (G194N, AG/C) e com o subtipo H (G2N, AG/H/AG). De realçar, também, que desde o final da década de 90, a infecção VIH-1 se encontra em franca expansão no país.

A análise global de ambas as regiões genómicas (V3-V5 de *env* e *nef*), confirmou a inexistência de vírus de subtipos puros e todos os casos foram classificados como formas recombinantes, algumas das quais sofreram vários eventos de recombinação resultando genomas virais em mosaico. Detectou-se a presença de sequências com composição genética muito semelhante entre si, evidenciando regiões de homologia com os subtipos A e CRF02\_AG mas aparentemente diferentes das CRF que se encontram descritas na actualidade ou das que foram descritas na população da Guiné-Bissau há cerca de 10 anos atrás (IbNG-*like* subtipo A/G)<sup>43</sup>. Devido ao facto de ainda persistirem condições políticas e económicas instáveis na Guiné-Bissau, as migrações populacionais são muito comuns no país. Desta forma, a intensa circulação de indivíduos infectados pode favorecer a selecção e a circulação de algumas formas em detrimento de outras, explicando a evolução da epidemiologia molecular da infecção e a elevada complexidade das formas virais em circulação. Contudo, estes resultados enfatizam a necessidade de que se analise a maior proporção do genoma possível de modo a diminuir as limitações na classificação dos vírus e torná-la mais correcta e próxima da realidade.

Para calcular o grau de divergência entre as sequências nucleotídicas analisadas, as médias das respectivas distâncias genéticas foram calculadas, tendo por base as diferentes regiões analisadas e/ou os subtipos observados. Comparativamente a *nef*, os valores intra- e inter-subtipo mais elevados obtidos para *env* podem traduzir a enorme variabilidade que caracteriza esta região, em resultado da maior pressão selectiva, quer exercida pelo sistema imunológico, quer imposta pelos anti-retrovíricos no tratamento dos doentes. Estas diferenças são corroboradas pelo estudo realizado por van Harmelen *et al.* em 2001<sup>90</sup> que sugere que as distâncias genéticas em *env*,

independentemente dos subtipos, mostram ser superiores comparativamente a outros genes virais. De notar também, que os resultados mostram que as distâncias inter-subtipo, independentemente da região genómica analisada, são sempre superiores das distâncias intra-subtipo.

A análise das sequências aminoacídicas das regiões V3-V5 de Env e de Nef foi centrada em observar o grau de conservação ou disrupção de domínios estruturais e funcionais descritos na literatura como biologicamente importantes na infecção VIH-1.

Relativamente à região V3-V5 de Env, o motivo estrutural composto por duas cisteínas, C<sub>1</sub> e C<sub>35</sub>, localizadas nos limites da região<sup>87</sup>, apresentou-se totalmente conservado prevendo a sua normal conformação. De uma forma global, verificou-se a existência de alguns polimorfismos. No entanto, os indícios sugerem que estes podem eventualmente estar associados a diferentes formas genéticas virais (subtipos não-B), e não propriamente, a uma alteração do motivo que impeça a realização de uma determinada função. A existência de um resíduo de glutamina (Q) na posição 18 na maioria dos casos envolvidos no estudo, leva a crer que poderá existir uma associação entre a alteração R→Q para os subtipos A e G. Um estudo revelou que a presença do resíduo R é comum no subtipo B, sendo encontrado em 67% das sequências referência<sup>92</sup>. A arginina (R) é um resíduo com carga positiva e carácter básico, não pertencendo ao mesmo grupo bioquímico que a glutamina (Q). A alteração R→Q foi também descrita para o subtipo C<sup>93</sup>.

Relativamente aos 3 potenciais locais de N-glicosilação identificados na região analisada (I, II, III), estes apresentaram-se bastante conservados em praticamente todos os casos estudados. Apenas na sequência G2E foi identificado um número reduzido de locais de N-glicosilação. Pensa-se que pressões selectivas actuam no sentido de manter um número mínimo de locais de N-glicosilação<sup>39</sup>, e uma vez que o padrão de glicosilação poderá ser importante na replicação e transmissibilidade do vírus, as regiões altamente glicosiladas constituem deste modo uma estratégia importante para escape do vírus ao sistema imunitário do hospedeiro<sup>65</sup>.

A análise dos motivos estruturais e funcionais da proteína Nef, revelou uma proteína completa e potencialmente funcional, contendo a maioria dos motivos conservados. Os motivos envolvidos na conformação de Nef, nomeadamente o sinal de miristilação (MGxxxS<sub>1</sub>) e o motivo ligação à molécula CD4 e local de clivagem da proteína Nef (WL<sub>57</sub>)<sup>69</sup>, encontraram-se 100% conservados. A conservação do motivo de miristilação é fundamental na associação de Nef com a membrana citoplasmática e consequente desempenho das funções da proteína<sup>68</sup>.

O resíduo de metionina (M<sub>20</sub>)<sup>94</sup>, envolvido na regulação negativa de MHC I, apresentou-se totalmente conservado, o mesmo não acontecendo com o motivo da hélice α (RRAE<sub>21</sub>)<sup>88</sup>, implicado na mesma função. Esta apresentou exclusivamente a forma RQNP, caracterizada essencialmente pela presença de resíduos com carga neutra (Q e N) e por uma prolina (aminoácido apolar), ou seja, uma alteração por resíduos com propriedades bioquímicas diferentes que poderia sugerir disrupção do motivo com comprometimento da função associada. No entanto, consultando a base de dados de sequências VIH-1, verificamos que para este motivo, existem polimorfismos descritos essencialmente em vírus não-B, e concretamente para as CRF02\_AG é descrita a forma RQTP, muito semelhante à observada nas sequências em análise.

No que diz respeito ao motivo ácido EEEE<sub>62</sub>, envolvido na regulação negativa de MHC I, através da ligação a PACS-1<sup>73</sup>, verificou-se que no total das sequências em estudo, e contrariamente, ao observado em sequências do subtipo B, a existência do motivo na forma EGEE (inserção do resíduo G com carga neutra), já observada e descrita noutros estudos<sup>41,95</sup>.

Os motivos implicados nas vias de sinalização intracelulares, entre os quais, os motivos de tetraprolina (PxxPxxPxRP<sub>69</sub>) e de prolina (PxxP<sub>147</sub>), identificados como locais de ligação a domínios SH3 de diversas cinases (Hck, Src e Vav)<sup>74</sup> apresentaram-se relativamente aos resíduos de prolina totalmente conservados. No entanto, no motivo RR<sub>105</sub>, implicado na ligação a PAK<sup>96</sup>, verificou-se a alteração do primeiro resíduo, R/K<sub>105</sub>→Q, que poderá ter impacto tanto no processo de enrolamento como na consequente estabilidade de Nef.

A consulta à base de dados de sequências VIH-1 confirmou que alterações dos tipos R→K ou K→R são frequentemente observadas (aminoácidos de propriedades bioquímicas semelhantes) e o resíduo Q tem uma posição conservada, independentemente dos subtipos virais em causa, logo imediatamente a seguir ao motivo. Contudo, com excepção da sequência G7N, quer o segundo resíduo do motivo (R ou K) quer os restantes resíduos associados à ligação a PAK descritos como sendo P<sub>69</sub>, L<sub>76</sub> e F<sub>121</sub><sup>97</sup>, mantiveram-se conservados, sugerindo que a activação desta proteína não será comprometida.

O local de fosforilação pela cinase PKC (RPMTYK<sub>77</sub>), manteve-se totalmente conservado, revelando ser potencialmente possível a fosforilação de Nef<sup>98</sup>. Por seu lado, também os motivos envolvidos na ligação a moléculas intervenientes no tráfego celular, nomeadamente, à tioesterase humana (FPD<sub>121</sub>)<sup>71</sup>, a β-COP (EE<sub>154</sub>)<sup>99</sup>, a complexos adaptadores (E/DxxxLL<sub>160</sub>)<sup>100</sup> e a V1H (DD<sub>174</sub>)<sup>72</sup> observaram-se de uma forma geral bem conservados ou com alterações que não sugerem comprometer as funções de regulação negativa de CD4 e/ou MHC I e II. Contudo, no motivo EE<sub>154</sub>, verificou-se a alteração E<sub>155</sub>→K em todas as sequências, que implica uma alteração de carga no resíduo dessa posição. No entanto, a relativa frequência com que estas alterações ocorrem naturalmente levanta dúvidas sobre a importância funcional deste motivo. No motivo DDPxxE<sub>174</sub>, envolvido na ligação à cinase c-Raf1<sup>89</sup> foram detectadas alterações como, D<sub>174</sub>→E e P<sub>176</sub>→E. A primeira alteração não parece comprometer a função do motivo, uma vez o aspartato (D) e o glutamato (E) são ambos ácidos com características bioquímicas semelhantes. No entanto, a alteração P<sub>176</sub>→E já poderá eventualmente interferir na função de Nef, pois a prolina é um aminoácido apolar com estrutura cíclica em que o azoto se encontra ligado a dois átomos de carbono e o aspartato é um aminoácido com carácter ácido. Contudo, atendendo à elevada frequência desta alteração, observada em sequências da base de dados, uma vez mais parece tratar-se de um polimorfismo de sequências não B.

## VI. CONCLUSÃO

A Guiné-Bissau durante os anos 80 e 90 foi identificada como sendo o país da África Ocidental com uma das maiores prevalências da infecção VIH-2, comparativamente à infecção por VIH-1, que na altura era quase inexistente. No entanto, a partir do final da década de 90, a situação inverteu-se, com um declínio da prevalência de VIH-2 e um aumento acentuado da infecção VIH-1. Os resultados do presente estudo estão de acordo com a realidade deste país, uma vez que se verificou uma prevalência de 15,8% para VIH-1 e de 4,2% para VIH-2 no grupo estudado. Foi obtida uma proporção de 47% de infecção com membros do complexo *Mycobacterium tuberculosis*, confirmando que a tuberculose também é um grave problema de Saúde Pública no país. A percentagem de co-infecção VIH e Tuberculose encontrada foi de 9,1%. No entanto, não foram detectadas diferenças significativas entre indivíduos infectados ou não com membros do CMTB e o risco de aquisição da infecção por VIH.

Em resultado da caracterização molecular *env* e *nef* de VIH-1, não foram detectadas sequências virais de subtipos puros e a maioria foi classificada como sendo formas genéticas A/G, embora em algumas delas se confirmasse também o envolvimento de outros subtipos virais, tais como o J, H, C e B, sugerindo ser genomas complexos, resultantes da recombinação eventualmente de recombinantes A/G, descritos no passado (*IbNG-like* subtipo A/G) na Guiné-Bissau. Apenas a análise do genoma total destes vírus poderia permitir uma classificação correcta e pormenorizada destas formas virais.

A análise das sequências aminoacídicas das regiões V3-V5 de Env e da proteína Nef revelou uma conservação dos motivos funcionais descritos na literatura, reflectindo a sua importância na infecciosidade e replicação viral. No entanto, foram observados alguns polimorfismos que podem estar associados a formas virais não B, especialmente observados no subtipo G e nas CRF01\_AE e CRF02\_AG (consulta de sequências da base de dados VIH). Contudo, a existência destes polimorfismos “naturais”, poderá influenciar a funcionalidade das respectivas proteínas nas diversas formas virais. Estudos *in vivo* e também *in vitro* com expressão e avaliação de funções das proteínas virais que abrangem formas genéticas diferentes do subtipo B podem de algum modo contribuir para o conhecimento das repercussões destas assinaturas no estabelecimento e na progressão da infecção e mostrar novos caminhos de investigação para o combate a esta infecção, nomeadamente através do desenho de uma vacina eficaz.

**VII. REFERÊNCIAS BIBLIOGRÁFICAS**

1. UNAIDS. 2009. Report on the Global Aids Epidemic.
2. Flint SJ, Enquist LW, Krug RM, *et al.* 2000. *Principles of Virology- Molecular Biology, Pathogenesis and Control of Animal Diseases*. 2ª Ed. ASM Press, Washington DC.
3. Barré-Sinoussi F, Chermann JC, Rey F, *et al.* 1983. Isolation of a T- Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune-Deficiency Syndrome (AIDS). *Science*, 220(4599):868-71.
4. Clavel F, Guetard D, Brunvezinet F, *et al.* 1986. Isolation of a New Human Retrovirus from West-African Patients with AIDS. *Science*, 233(4761):343-6.
5. Knipe DM, Howley DE, Griffin RA, *et al.* 2001. *Fields Virology*. 4ª Ed. Lippincott Williams & Wilkins, Philadelphia, USA.
6. McCutchan FE. 2006. Global Epidemiology of HIV. *J. Med. Virol.*, 78(S1):S7-12.
7. Hanh BH, Shaw GM, De Cock KM, *et al.* 2000. AIDS – AIDS as a Zoonosis: Scientific and Public Health Implications. *Science*, 287(5453):607-14.
8. Gao F, Bailes E, Robertson DL, *et al.* 1999. Origin of HIV-1 in the Chimpanzee *Pan troglodytes troglodytes*. *Nature*, 397(6718):436-41.
9. Hirsch VM, Olmsted RA, Murpheykorb M, *et al.* 1989. An African Primate Lentivirus (SIV<sub>sm</sub>) Closely Related to HIV-2. *Nature*, 339(6223):389-92.
10. Simon V, Ho DD, Karim QA. 2006. HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. *The Lancet*, 368: 489-504.
11. Barré-Sinoussi F. 1996. HIV as the Cause of AIDS. *The Lancet*, 348:31-35.
12. Sougrat R, Bartesaghi A, Lifson JD, *et al.* 2007. Electron Tomography of the Contact between T Cells and SIV/HIV-1: Implications of Viral Entry. *PLoS Pathog.*, 3:e63.
13. Levy JA. 2007. *HIV and the Pathogenesis of AIDS*, 3ª ed, ASM Press, Washington, DC.
14. Freed EO. 2001. HIV Replication. *Somatic Cell and Molecular Genetics* 26:13-33.
15. Hutchinson JF. 2001. Biology and Evolution of HIV. *Annu. Rev. Anthropol.*, 30: 85-108.
16. Azevedo- Pereira JM. Ciclo Biológico de VIH In: Antunes F (ed). 2008. *Manual sobre SIDA*. Permanyer Portugal, Lisboa, pp.9-26.
17. Stevenson M. 2003. Dissecting HIV through RNA Interference. *Nat. Rev. Immunol.*, 3 (11):851-8.
18. Jaffe HW, Bregman DJ, Selik RM. 1983. Acquired immune deficiency syndrome in the United States: the first 1,000 cases. *J. Infect. Dis.*, 148:339-45.
19. Sousa AE, Victorino R. Imunopatogénese e Resposta Imunológica In: Antunes F (ed). 2008. *Manual sobre SIDA*. Permanyer Portugal, Lisboa, pp.53.75.
20. Berger EA, Murphy PM, Farber JM. 1999. Chemokine Receptors as HIV-1 Coreceptors: Roles in Viral Entry, Tropism, and Disease. *Annu. Rev. Immunol.*, 17:657-700.
21. In [schools-wikipedia.org/wp/h/HIV.htm](http://schools-wikipedia.org/wp/h/HIV.htm).
22. Roberts J, Bebenek B, Loeb L. 1988. The accuracy of reverse transcriptase from HIV-1. *Science*, 242:1171-3.
23. Robertson D, Sharp P, McCutchan F, *et al.* 1995. Recombination in HIV-1. *Nature*, 374:124-6.

## VII.REFERÊNCIAS BIBLIOGRÁFICAS

24. Ramirez BC, Simon-Loriere E, Galetto R, *et al.* 2008. Implications of Recombination for HIV Diversity. *Virus Research*, 134:64-73.
25. Gupta RM, Sahni AK, Jena J, *et al.* 2004. Genomic Diversity of Human Immunodeficiency Viruses, *Contemporary Issue*, 61: 267-270.
26. Pinto ME, Struchiner CJ. 2006. A diversidade do HIV-1: uma ferramenta para o estudo da pandemia, *Rev. Cad. Saúde Pública*, Rio de Janeiro, 22 (3): 473-484.
27. Gao F, Trask SA, Hui H, *et al.* 2001. Molecular Characterization of a Highly Divergent HIV Type 1 Isolate Obtained Early in the AIDS Epidemic from the Democratic Republic of Congo. *AIDS Research and Human Retroviruses*, 17(22): 1217-22.
28. Plantier JC, *et al.* 2009. A New Human Immunodeficiency Virus derived from Gorillas. *Nat Med*.
29. Taylor BS, Sobieszczyk ME, McCutchan FE, *et al.* 2008. The Challenge of HIV-1 Subtype Diversity. *N. Engl. J. Med.*, 358(15):1590-602.
30. *HIV Sequence Compendium* 2009. Carla Kuiken, Thomas Leitner, Brian Foley *et al.* editors. 2009. Publisher: Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico. LA-UR 09-03280.
31. Jacobs GB, Loxton AG, Laten A, *et al.* 2009. Emergence and Diversity of Different HIV-1 Subtypes in South Africa, 2000-2001. *Journal of Medical Virology*, 81: 1852-1859.
32. Buonaguro M, Tornesello ML, Buonaguro FM. 2007. Human Immunodeficiency Virus Type 1 Subtype Distribution in the Worldwide Epidemic: Pathogenic and Therapeutic Implications. *J. Virol*, 81(19):10209-19.
33. Osmanov S, Pattou C, Walker N, Schwardlander B, *et al.* 2002. Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. *J Acquir Immune Defic Syndr.*, 29:184-90.
34. Tabeke Y, Kusagawa S, Motomura K. 2004. Molecular epidemiology of HIV: Tracking AIDS pandemic. *Pediatr. Int.* 46: 236-44.
35. Peeters M, Toure-Kane C, Nkengansong JN. 2003. Genetic Diversity of HIV in Africa: impact on diagnosis, treatment, vaccine development and trials. *AIDS* 17:2547-60.
36. Kandathil AJ, Ramalingam S, Kannangai R, *et al.* 2005. Molecular epidemiology of HIV. *Indian J. Med. Res.* 121:333-44.
37. Lau KA, Wang B, Saksena NK. 2007. Emerging trends of HIV epidemiology in Asia. *AIDS Reviews*.
38. Perrin L, Kaiser L, Yerly S. 2003. Travel and the spread of HIV-1 genetic variants, *Rev. The Lancet Infectious Diseases, Vol 3*: 22-27.
39. Esteves A, Parreira R, Venenno T, *et al.* 2002. Molecular Epidemiology of HIV Type 1 Infection in Portugal: High Prevalence of Non-B Subtypes, *AIDS Res. Hum. Retroviruses*, 18 (5): 313-325.
40. Palma AC, Araújo F, Duque V, *et al.* 2007. Molecular epidemiology and prevalence of drug resistance-associated mutations in newly diagnosed HIV-1 patients in Portugal. *Infection, Genetics and Evolution*, 7: 391-98.
41. Parreira R, Pádua E, Piedade J, *et al.* 2005. Genetic analysis of human immunodeficiency virus type 1 *nef* in Portugal: Subtyping, identification of mosaic genes, and amino acid sequence variability. *J Med Virol*, 77:8-16.
42. Pádua E, Parreira R, Tendeiro R, *et al.* 2009. Potential impact of viral load and genetic makeup of HIV type 1 on mother-to-child transmission: characterization of *env-C2V3C3* and *nef* sequences. *AIDS Res Hum Retroviruses*, 25:1171-7.
43. Andersson S, Norrgren H, Dias F, *et al.* 1999. Molecular Characterization of Human Immunodeficiency Virus (HIV)-1 and -2 in Individuals from Guinea-Bissau with Single or Dual Infections: Predominance of a Distinct HIV-1 Subtype A/G Recombinant in West Africa. *Virology*, 262: 312-20.

## VII.REFERÊNCIAS BIBLIOGRÁFICAS

44. Damond F, Descamps D, Farfara I, *et al.* 2001. Quantification of Proviral Load of Human Immunodeficiency Virus Type 2 Subtypes A and B using Real-time PCR. *J Clinical Microbiology*, 39(12): 4264-68.
45. In [www.aidglobal.org](http://www.aidglobal.org).
46. Laeyendecker O, Li X, Arroyo M, *et al.* 2006. The Effect of HIV Subtype on Rapid Disease Progression in Rakai, Uganda. *13<sup>th</sup> Conference on retroviruses and Opportunistic Infections (abstract no 44LB)*.
47. Baeten D, *et al.* 2007. HIV-1 subtype D infection associated with faster disease progression than subtype A, in spite of similar HIV-1 plasma viral loads. *14<sup>th</sup> Conference on retroviruses and Opportunistic Infections (abstract no 68)*.
48. Nelson KE, *et al.* 2007. Survival of blood donors and their spouses with HIV-1 subtype E (CRF01\_AE) infection in northern Thailand, 1992-2007. *AIDS Volume 21 Supplement 6*.
49. Pope M, Frankel SS, Mascola JR, *et al.* 1997. Human immunodeficiency virus type 1 strains of subtypes B and E replicate in cutaneous dendritic cell-T-cell mixtures without displaying subtype-specific tropism. *J Virol*, 71 (10): 8001-07.
50. Dittmar MT, Simmons G, Hibbitts S, *et al.* 1997. Langerhans cell tropism of human immunodeficiency virus type 1 subtype A through F isolates derived from different transmission groups. *J Virol*, 71 (10): 8008-13.
51. Murray, Embree, *et al.* 2000. Effect of human immunodeficiency virus (HIV) type 1 viral genotype on mother-to-child transmission of HIV-1. *J Infect Dis.*, 2(181).
52. Lal RB, Chakrabarti S, Yang C. 2005. Impact of genetic diversity of HIV-1 on diagnosis, antiretroviral therapy & vaccine development. *Indian J. Med. Res.*, 121:287-314.
53. Buttler IF, Pandrea I, Marx PA, *et al.* 2007. HIV Genetic diversity: Biological and public health consequences. *Curr. HIV Res.*, 5:23-45.
54. Cane PA. 2009. New developments in HIV drug resistance. *J. Antimicrob. Chemother.*, 64:37-40.
55. Moyle G, Gatell J, Perno F, *et al.* 2008. Potential for new antiretroviral to address unmet needs in the management of HIV-1 infection. *AIDS Patient Care STDS*. 22:459-71.
56. CDC HIV/AIDS Facts. January 2008.
57. Gerald Friedland, MD. 2007. Tuberculosis, Drug Resistance, and HIV/AIDS: A Triple Threat. *Current Infectious Disease Reports*, 9: 252-261.
58. Koivula T, Cristea-Fernstrom M, Chryssanthou E, *et al.* 2004. Genetic Diversity in clinical isolates of *Mycobacterium avium* complex from Guinea-Bissau, West Africa. *Microbes Infect*, 6(14):1320-5.
59. Kwong PD, Wyatt R, Robinson J. 1998. Structure of an HIV gp120 Envelope Glycoprotein in Complex with the CD4 Receptor and a Neutralizing Human Antibody. *Nature*, 393(6686):648-59.
60. Milich L, Margolin B, Swanstrom R. 1993. V3 Loop of the Human-Immunodeficiency Virus Type 1 Env Protein: Interpreting Sequence Variability. *J. Virol.*, 67(9):5623-34.
61. Jensen MA, Li F-S, van't Wout AB, *et al.* 2003. Improved Coreceptor Usage Prediction and Genotypic Monitoring of R5-to-X4 Transition by Motif Analysis of Human Immunodeficiency Virus Type 1 env V3 Loop Sequences. *J. Virol.*, 77(24):13376-88.
62. Pollakis G, Kang S, Kliphuis A, *et al.* 2001. N-Linked Glycosylation of the HIV Type-1 gp120 Envelope Glycoprotein as a Major Determinant of CCR5 and CXCR4 Coreceptor Utilization. *J. Biol. Chem.* 276(16):13433-41.
63. Xiao L, Owen SM, Goldman I, *et al.* 1998. CCR5 Coreceptor Usage of Non-Syncytium-Inducing Primary HIV-1 Is Independent of Phylogenetically Distinct Global HIV-1 Isolates: Delineation of Consensus Motif in the V3 Domain that Predicts CCR-5 Usage. *Virology*, 240(1):83-92.
64. Williamson S. 2003. Adaptation in the Env Gene of HIV-1 and Evolutionary Theories of Disease Progression. *Mol. Biol. Evol*, 20(8):1318-25.

## VII.REFERÊNCIAS BIBLIOGRÁFICAS

65. Wyatt R, Sodroski J. 1998. The HIV-1 Envelope Glycoproteins: Fusogens, Antigens, and Immunogens. *Science*, 280(5371):1884-8.
66. Kirchhoff F, Schindler M, Specht A, Arhel N, Munch J. 2008. Role of Nef in primate lentiviral immunopathogenesis. *Cell Mol. Life Sci.* 65(17):2621-36.
67. Klotman ME, Kim SY, Buchbinder A, *et al.* 1992. Kinetics of Expression of Multiply Spliced RNA in Early Human Immunodeficiency Virus Type 1 Infection of Lymphocytes and Monocytes. *Proc. Natl. Acad. Sci.*, 89(3):1148.
68. Geyer M, Fackler OT, Peterlin BM. 2001. Structure-function Relationships in HIV-1 Nef. *EMBO Rep.*, 2(7):580-5.
69. Freund J, Kellner R, Konvalinka J, *et al.* 1994. A Possible Regulation of Negative Factor (Nef) Activity of Human Immunodeficiency Virus Type 1 by the Viral Protease. *Eur. J. Biochem.*, 223:589-93.
70. Le Gall S, Erdtmann L, Schwarz O, *et al.* 1998. Nef interacts with the  $\mu$  subunit of clathrin adaptor complexes and reveals a cryptic sorting signal in MHC-I molecules. *Immunity*, 8(4):483-95.
71. Liu LX, Heveker N, Fackler OT, *et al.* 2000. Mutation of a Conserved Residue (D123) Required for Oligomerization of Human Immunodeficiency Virus Type 1 Nef Protein Abolishes Interaction with Human Thioesterase and Results in Impairment of Nef Biological Functions. *J Virol.*, 74(11):5310-9.
72. Lu X, Yu H, Liu SH, *et al.* 1998. Interactions Between HIV1 Nef and Vacuolar ATPase Facilitate the Internalization of CD4. *Immunity*, 8(5):647-56.
73. Piguet V, Demareux N, Trono D, *et al.* 2000. HIV-1 Nef Protein Binds to the Cellular Protein PACS-1 to Downregulate Class I Major Histocompatibility Complexes. *Nat. Cell Biol.* 2(3):163-7.
74. Saksela K, Cheng G, Baltimore D. 1995. Proline-rich (PxxP) Motifs in HIV-1 Nef Bind to SH3 Domains of a Subset of Src Kinases and are required for the Enhanced Growth of Nef Viruses but not for Downregulation of CD4. *EMBO J.*, 14(3):484-91.
75. Piguet V, Trono D. 1999. A Structure-function Analysis of the Nef Protein of Primate Lentiviruses. In CL Kuiken CL, Foley B, Hahn B, *et al.* 1999. *Human Retroviruses and AIDS*. Los Alamos National Laboratory, NM. pp. 448-59.
76. Pamer E, Cresswell P. 1998. Mechanisms of MHC Class I - Restricted Antigen Processing. *Annu Rev Immunol*, 16:323-58.
77. Semple M, Loveday C, Weller I, *et al.* 1991. Direct measurement of viraemia in patients infected with HIV-1 and its relationship to disease progression and zidovudine therapy. *J Med Virol*, 35:38-45.
78. Berry N, Ariyoshi K, Jobe O, *et al.* 1994. HIV type 2 proviral load measured by quantitative polymerase chain reaction correlates with CD4 lymphopaenia in HIV type-2-infected individuals. *AIDS Res Hum Retroviruses*, 10:1031-37.
79. Delwart EL, Shpaer EG, Louwagie J, *et al.* 1993. Genetic relationships determined by a DNA heteroduplex mobility assay: Analysis of HIV-1 *env* genes. *Science*, 262:1257-61.
80. Larkin MA, Blackshields G, Brown NP, *et al.* 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23:2947-8.
81. Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.*, 41:95-8.
82. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, 4(4):406-25.
83. Kimura M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J Molec. Evol*, 16:111-20.
84. Tamura K, Dudley J, Nei M, *et al.* 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol. Biol. Evol.*, 24(8):1596-9.

## VII.REFERÊNCIAS BIBLIOGRÁFICAS

85. Lole KS, Bollinger RC, Paranjape RS, *et al.* 1999. Full-length Human Immunodeficiency Virus Type 1 Genomes from Subtype C-infected Seroconverters in India, with Evidence of Intersubtype recombination. *J. Virol.*, 73(1):152-60.
86. Zhang M, *et al.* 2004. Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology*. 14(12):1229-46.
87. Hwang SS, Boyle TJ, Lyerly HK, *et al.* 1992. Identification of envelope V3 loop as the major determinant of CD4 neutralization sensitivity of HIV-1. *Science* 257:535-7.
88. Mangasarian A, Piguet V, Wang JK, *et al.* 1999. Nef-induced CD4 and major histocompatibility complex class I (MHC-I) down-regulation are governed by distinct determinants: N-terminal alpha helix and proline repeat of Nef selectively regulate MHC-I trafficking. *J Virol*, 73:1964-73.
89. Hodge DR, Dunn KJ, Pei GK, *et al.* 1998. Binding of c-Raft kinase to a conserved acidic sequence within the carboxylterminal region of the HIV-1 Nef protein. *J Biol Chem.*, 273(25):15727-33.
90. van Harmelen J, Williamson C, Kim B, *et al.* 2001. Characterization of full-length HIV type 1 subtype C sequences from South Africa. *AIDS Res. Hum. Retroviruses*, 17 (16):1527-31.
91. Novitsky VA, Montano MA, McLane MF, *et al.* 1999. Molecular Cloning and Phylogenetic Analysis of Human Immunodeficiency Virus Type 1 Subtype C: a Set of 23 Full-Length Clones from Botswana. *J. Virol.*, 73(5): 4427-32.
92. Ataman-Onal Y, Coiffier C, Giraud A, *et al.* 1999. Comparison of complete env Gene Sequences from Individuals with Symptomatic Primary HIV Type 1 infection. *AIDS Res. Hum. Retroviruses*, 15 (11):1035-39.
93. Engelbrecht S, Villiers T, Sampson CC, *et al.* 2001. Genetic Analysis of the Complete gag and env Genes of HIV Type 1 Subtype C Primary Isolates from South Africa. *AIDS Res. Hum. Retroviruses*, 17(16):1533-47.
94. Blagoveshchenskaya AD, Thomas L, Feliciangeli SF, Hung CH, *et al.* 2002. HIV-1 nef downregulates MHC-1 by a PACS-1 and P13K-regulated Arf6 endocytic pathway. *Cell*, 111:853-866.
95. Jubier-Maurin V, Saragosti S, Perret JL, *et al.* 1999. Genetic Characterization of the nef Gene from Human Immunodeficiency Virus Type 1 Group M Strains Representing Genetic Subtypes A, B, C, E, F, G, and H. *AIDS Res. Hum. Retroviruses*, 15(1):23-32.
96. Renkema GH, Manninen A, Mann DA, *et al.* 1999. Identification of the Nef-associated kinase as p21-activated kinase 2. *Curr. Biol.*, 9(23):1407-10.
97. Manninen A, Hiipakka M, Vihinen M, *et al.* 1998. SH3-domain binding function of HIV-1 Nef is required for association with a PAK-related kinase. *Virology*, 250(2):273-82.
98. Kennelly PJ, Krebs EG. 1991. Consensus sequences as substrate specificity determinants for protein kinases and protein phosphatases. *J Biol Chem*, 266:15555-8.
99. Piguet V, Gu F, Foti M, *et al.* 1999. Nef-induced CD4 Degradation: a Diacidic-based Motif in Nef functions as a Lysosomal Targeting Signal Through the Binding of  $\beta$ -COP in Endosomes. *Cell* 97(1):63-73.
100. Bresnahan PA, Yonemoto W, Ferrell S, *et al.* 1998. A Dileucine Motif in HIV-1 Nef Acts as an Internalization Signal for CD4 Downregulation and Binds the AP-1 Clathrin Adaptor. *Curr. Biol.*, 8(22):1235-8.