



**INSTITUTO DE HIGIENE E  
MEDICINA TROPICAL**  
DESDE 1902



**UNIVERSIDADE  
NOVA  
DE LISBOA**

**Universidade Nova de Lisboa**  
**Instituto de Higiene e Medicina Tropical**

**Diversidade genética e resistência aos anti-retrovirais no vírus da  
imunodeficiência humana tipo 1 em circulação  
em Benguela, Angola**

**Ana Raquel Dias Quitéria**

**DISSERTAÇÃO PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS BIOMÉDICAS  
ESPECIALIDADE EM BIOLOGIA MOLECULAR EM SAÚDE TROPICAL E  
INTERNACIONAL**

**Julho 2019**



INSTITUTO DE HIGIENE E  
MEDICINA TROPICAL  
DESDE 1902



UNIVERSIDADE  
**NOVA**  
DE LISBOA

**Universidade Nova de Lisboa**

**Instituto de Higiene e Medicina Tropical**

Diversidade genética e resistência aos anti-retrovirais  
no vírus da imunodeficiência humana tipo 1 em  
circulação em Benguela, Angola

**Autor:** Ana Raquel Dias Quitéria

**Orientador:** Prof. Doutor João Piedade

**Co-orientadora:** Prof.<sup>a</sup> Doutora Filomena Martins Pereira

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do grau de  
Mestre em Ciências Biomédicas especialidade em Biologia Molecular em Saúde Tropical e  
Internacional

A presente dissertação foi redigida segundo o acordo ortográfico de 1945.

## Elementos bibliográficos

---

- ❖ Quitéria R., Yefimenko L., Martins Pereira F., Piedade J. Diversidade genética e resistência aos anti-retrovirais no vírus da imunodeficiência humana tipo 1 em circulação na cidade de Benguela, Angola. *Special GHTM Session – Second GHTM Antimicrobial Resistance Awareness Day*. 13 de Novembro de 2018. Lisboa, Portugal  
(Comunicação oral)
  
- ❖ Quitéria R., Yefimenko L., Martins Pereira F., Piedade J. Genetic diversity and drug resistance-associated mutations in human immunodeficiency virus type 1 circulating in Benguela, western Angola. *IV NOVAhealth Genetics Workshop*. 21 de Março de 2019. Lisboa, Portugal  
(Comunicação em póster; Anexo 1)
  
- ❖ Quitéria R., Yefimenko L., Martins Pereira F., Piedade J. Genetic diversity and drug resistance-associated mutations in human immunodeficiency virus type 1 (HIV-1) circulating in Benguela, western Angola. *7<sup>th</sup> European Congress of Virology*. 28 de Abril – 1 de Maio de 2019. Roterdão, Holanda  
(Comunicação em póster; Resumo n.º: 340; Código póster: PE54; Anexo 2)

## **Agradecimentos**

---

Se não fosse o apoio e incentivo de muito boa gente, esta dissertação não teria visto a luz do dia nem tinha o mesmo valor. Por isso, não posso deixar de expressar o meu agradecimento àqueles que tornaram possível a sua concretização.

Ao Professor Doutor João Piedade, orientador desta tese, pela dedicação, disponibilidade, apoio e paciência. Todas as críticas e sugestões, a partilha de conhecimentos e o trabalho em conjunto foram essenciais para que este projecto chegasse a bom porto depois de todas as tempestades.

À Dr.<sup>a</sup> Lesya Yefimenko e à Professora Doutora Filomena Martins Pereira por fornecerem as amostras biológicas e respectivos dados epidemiológicos. Obrigada pela disponibilidade e simpatia demonstrada sempre que necessitei de esclarecer alguma dúvida.

À Técnica Superior Ângela Lopes pela preciosa ajuda laboratorial, simpatia, e disponibilidade durante a realização deste projecto.

Aos meus amigos pelo apoio e motivação constantes, foram todos muito importantes, mas não posso deixar de agradecer especialmente à Patrícia Borges, Mónica Ferreira e Paloma Gonçalves. As nossas horas de almoço, pausas e fofocas foram essenciais para a manutenção da minha sanidade mental durante esta etapa complexa, vou levar-vos comigo para sempre!

À minha família, em especial à minha mãe e à minha irmã, cujo apoio incondicional, motivação, carinho e compreensão foram e são muito importantes para mim. Obrigada por tudo, mesmo!!! Sem vocês não estaria aqui hoje.

E ao Ricardo, que tal como a minha mãe e a minha irmã, foi um porto de abrigo durante todo o mestrado e me ajudou a manter o foco mesmo quando tudo parecia correr mal. Obrigada por seres o meu chato!

Muito obrigada a todos!

## Resumo

---

Em Angola, assiste-se a um constante aumento da complexidade da epidemia do vírus da imunodeficiência humana tipo 1 (HIV-1), caracterizada pela circulação da maioria dos subtipos do grupo M e crescente frequência de mutações associadas a resistência aos anti-retrovirais (ARVs), especialmente aos inibidores da transcriptase reversa (RT). A introdução do acesso universal à terapia anti-retroviral trouxe melhorias na qualidade de vida das pessoas a viver com o HIV, mas o aparente aumento da transmissão primária de vírus resistentes traz dificuldades ao tratamento e controlo da infecção. Por outro lado, os dados disponíveis sobre a infecção são ainda escassos e maioritariamente provenientes de Luanda. Assim, os objectivos deste estudo foram caracterizar as estirpes de HIV-1 circulantes na província de Benguela (Angola ocidental) e estimar o respectivo grau de resistência primária aos inibidores da protease (PR) e da RT actualmente certificados.

Foi incluído um total de 111 indivíduos diagnosticados com infecção pelo HIV-1, entre Agosto de 2016 e Janeiro de 2017, no Hospital Geral de Benguela, sendo que, na sua totalidade, estes nunca tinham sido submetidos a terapia anti-retroviral. As amostras de sangue periférico colhidas foram armazenadas em cartões *Whatman Indicating FTA® Elute Micro Cards* (GE Healthcare, Reino Unido) e o DNA proviral presente nas células infectadas utilizado como matriz para a amplificação por dupla *nested* PCR das regiões codificantes da PR (460 pb) e RT (650 pb). A classificação genotípica das 103 sequências obtidas foi realizada com base em análise filogenética e de *bootscanning*, tendo-se recorrido ao programa *HIVdb – Genotypic Resistance Interpretation Algorithm* da *Stanford University HIV Drug Resistance Database* para a pesquisa de mutações associadas a resistência aos inibidores da PR e da RT, e outros polimorfismos genéticos, e determinação dos respectivos perfis fenotípicos putativos de resistência.

Na globalidade, na amostra populacional em estudo, verificou-se o predomínio dos subtipos C (40,4%) e F1 (19,1%), seguidos da CRF02\_AG (12,8%) e subtipos G (8,5%), A (4,3%), B, D e J (2,1%, cada), e ainda uma pequena percentagem de formas recombinantes únicas (URFs) (6,4%) e sequências não genotipáveis (U) (2,1%). Encontrou-se um conjunto diversificado de mutações associadas a resistência aos diferentes inibidores da PR e da RT considerados na análise, incluindo mutações principais e acessórias, das quais uma parte considerável tem relevância para a vigilância epidemiológica da resistência transmitida aos ARVs no HIV-1. A grande maioria das sequências da região codificante da PR encontra-se associada a perfis fenotípicos de susceptibilidade (92,0%), mas uma percentagem considerável das sequências da região codificante da RT apresentava algum tipo de resistência aos respectivos inibidores (34,0%).

Resumindo, foi encontrada uma elevada diversidade genética nas estirpes do HIV-1 circulantes em Benguela, destacando-se o considerável grau de resistência primária aos ARVs observado, em particular, relativamente aos inibidores da RT. Deste modo, os resultados obtidos neste estudo reforçam a importância da monitorização da emergência de mutações associadas a resistência aos diferentes ARVs e sua transmissão em Angola, considerando o seu impacto directo na efectividade da terapêutica e controlo da infecção pelo HIV-1.

**Palavras-chave:** HIV-1, Angola, diversidade genética, terapia anti-retroviral, resistência primária aos anti-retrovirais

## Abstract

---

In Angola, there has been an increase in the complexity of the human immunodeficiency virus type 1 (HIV-1) epidemic, characterized by the circulation of the majority of group M subtypes and the increasing frequency of drug resistance mutations (DRMs), especially to reverse transcriptase (RT) inhibitors. The introduction of universal access to antiretroviral therapy has brought improvements in the quality of life of people living with HIV, but the apparent increase in the primary transmission of resistant viruses brings difficulties to treatment and control of the infection. On the other hand, available data on the infection are still scarce and mainly from Luanda. So, the aims of this study were to genetically characterize the circulating HIV-1 strains in the province of Benguela (western Angola) and to assess the corresponding degree of primary drug resistance to currently licensed protease (PR) and RT inhibitors.

A total of 111 individuals diagnosed with HIV-1 infection between August 2016 and January 2017 at the General Hospital of Benguela were included, all of whom were treatment-naïve. Peripheral blood samples were collected and stored on Whatman Indicating FTA<sup>®</sup> Elute Micro Cards (GE Healthcare, UK) and the proviral DNA used as matrix for the double nested PCR amplification of PR (460 bp) and RT (650 bp) coding regions. The genotypic classification of the 103 obtained nucleotide sequences was based on their phylogenetic and bootscanning analyses. The HIVdb program – Genotypic Resistance Interpretation Algorithm (Stanford University, USA) was used for characterization of DRMs to PR and RT inhibitors, and other genetic polymorphisms, and determination of their putative phenotypic resistance profiles.

Overall, in this population sample, there was a predominance of subtypes C (40.4%) and F1 (19.1%), followed by CRF02\_AG (12.8%), and subtypes G (8.5%), A (4.3%), B, D and J (2.1%, each), with a small percentage of unique recombinant forms (URFs) (6.4%) and unclassified sequences (U) (2.1%). A diverse set of DRMs to the different PR and RT inhibitors considered in the analysis was found, including major and accessory mutations, of which a considerable proportion is relevant to the epidemiological surveillance of transmitted HIV-1 drug resistance. The great majority of PR coding sequences was associated to phenotypic profiles of susceptibility to the different inhibitors (92.0%), but a considerable proportion of RT coding sequences showed some level of resistance to RT inhibitors (34.0%).

In summary, a high genetic diversity was found in the HIV-1 strains circulating in Benguela, associated to a considerable degree of primary drug resistance, especially to RT inhibitors. The results obtained in this study reinforce the importance of monitoring the emergence of DRMs and their transmission in Angola, due to their obvious direct impact on therapy efficacy and control of the HIV-1 infection.

**Keywords:** HIV-1, Angola, genetic diversity, antiretroviral therapy, primary drug resistance

## Índice

---

|                                                                              |           |
|------------------------------------------------------------------------------|-----------|
| Índice de figuras.....                                                       | x         |
| Índice de tabelas.....                                                       | xiii      |
| Lista de abreviaturas, acrónimos e siglas .....                              | xiv       |
| <b>1 Introdução .....</b>                                                    | <b>1</b>  |
| <b>1.1 Vírus da imunodeficiência humana e sida .....</b>                     | <b>1</b>  |
| <b>1.2 Origem biológica do HIV.....</b>                                      | <b>2</b>  |
| <b>1.3 Dados epidemiológicos globais .....</b>                               | <b>4</b>  |
| <b>1.4 Taxonomia e características gerais do HIV .....</b>                   | <b>5</b>  |
| <b>1.5 Estrutura e organização genómica do HIV-1 .....</b>                   | <b>6</b>  |
| <b>1.6 Ciclo replicativo do HIV-1.....</b>                                   | <b>8</b>  |
| <b>1.7 Diversidade genética .....</b>                                        | <b>11</b> |
| <b>1.7.1 Classificação do HIV-1.....</b>                                     | <b>11</b> |
| <b>1.7.2 Distribuição mundial.....</b>                                       | <b>12</b> |
| <b>1.7.3 Mecanismos responsáveis e impacto da diversidade genética .....</b> | <b>13</b> |
| <b>1.8 Gene <i>pol</i>.....</b>                                              | <b>14</b> |
| <b>1.8.1 Protease .....</b>                                                  | <b>15</b> |
| <b>1.8.2 Transcriptase reversa.....</b>                                      | <b>16</b> |
| <b>1.9 Terapia anti-retroviral.....</b>                                      | <b>17</b> |
| <b>1.10 Resistência associada à terapia anti-retroviral.....</b>             | <b>18</b> |
| <b>1.11 Monitorização de resistência aos anti-retrovirais .....</b>          | <b>20</b> |
| <b>1.12 Inibidores da protease e da transcriptase reversa .....</b>          | <b>21</b> |
| <b>1.12.1 Mutações de resistência.....</b>                                   | <b>22</b> |
| <b>1.13 HIV-1 em Angola.....</b>                                             | <b>26</b> |
| <b>1.13.1 Dados epidemiológicos gerais .....</b>                             | <b>26</b> |

|                                                                                                                   |    |
|-------------------------------------------------------------------------------------------------------------------|----|
| 1.13.2 Epidemiologia molecular .....                                                                              | 27 |
| 1.13.2.1 Genótipos circulantes .....                                                                              | 27 |
| 1.13.2.2 Mutações de resistência aos inibidores da protease e da transcriptase reversa .....                      | 29 |
| 1.14 Problema de investigação .....                                                                               | 30 |
| 1.14.1 Objectivos .....                                                                                           | 31 |
| 2 Material e métodos.....                                                                                         | 33 |
| 2.1 Caracterização geral da amostra populacional em estudo.....                                                   | 33 |
| 2.1.1 Colheita das amostras e rastreio da infecção pelo HIV-1 .....                                               | 33 |
| 2.2 Preparação de discos dos cartões FTA® .....                                                                   | 34 |
| 2.3 Amplificação por PCR das regiões codificantes da protease e da transcriptase reversa no gene <i>pol</i> ..... | 35 |
| 2.3.1 Caracterização dos <i>primers</i> utilizados .....                                                          | 35 |
| 2.3.2 Planificação do protocolo de amplificação .....                                                             | 36 |
| 2.3.2.1 Reacções de PCR.....                                                                                      | 39 |
| 2.4 Visualização dos produtos de PCR.....                                                                         | 41 |
| 2.5 Purificação dos produtos de PCR.....                                                                          | 41 |
| 2.6 Sequenciação nucleotídica dos produtos de PCR.....                                                            | 42 |
| 2.7 Análise bioinformática das sequências nucleotídicas.....                                                      | 43 |
| 2.7.1 Edição das sequências e confirmação da sua identidade .....                                                 | 43 |
| 2.7.2 Genotipagem automatizada das sequências de HIV-1 .....                                                      | 43 |
| 2.7.3 Análise filogenética .....                                                                                  | 44 |
| 2.7.3.1 Análise de sequências divergentes .....                                                                   | 44 |
| 2.8 Análise bioinformática das sequências aminoacídicas.....                                                      | 45 |
| 2.8.1 Pesquisa de mutações associadas a resistência a anti-retrovirais .....                                      | 45 |
| 2.9 Análise estatística .....                                                                                     | 46 |

|                                                                                                                          |            |
|--------------------------------------------------------------------------------------------------------------------------|------------|
| <b>3 Resultados.....</b>                                                                                                 | <b>47</b>  |
| <b>3.1 Caracterização da amostra populacional em estudo .....</b>                                                        | <b>47</b>  |
| <b>3.2 Amplificação por PCR das regiões codificantes da protease e da transcriptase reversa no gene <i>pol</i> .....</b> | <b>48</b>  |
| <b>3.3 Análise bioinformática das sequências nucleotídicas.....</b>                                                      | <b>50</b>  |
| 3.3.1 Edição das sequências e confirmação da sua identidade .....                                                        | 50         |
| 3.3.2 Genotipagem automatizada das sequências de HIV-1 .....                                                             | 51         |
| 3.3.3 Análise filogenética .....                                                                                         | 52         |
| <b>3.4 Análise bioinformática das sequências aminoacídicas.....</b>                                                      | <b>63</b>  |
| 3.4.1 Pesquisa de mutações associadas a resistência a anti-retrovirais .....                                             | 63         |
| 3.4.1.1 Região codificante da protease .....                                                                             | 64         |
| 3.4.1.2 Região codificante da transcriptase reversa.....                                                                 | 66         |
| <b>4 Discussão e conclusão .....</b>                                                                                     | <b>70</b>  |
| <b>5 Referências bibliográficas .....</b>                                                                                | <b>94</b>  |
| <b>6 Anexos .....</b>                                                                                                    | <b>123</b> |
| <b>Anexo 1: Póster apresentado no IV NOVAhealth Genetics Workshop.....</b>                                               | <b>123</b> |
| <b>Anexo 2: Póster apresentado no 7<sup>th</sup> European Congress of Virology .....</b>                                 | <b>124</b> |
| <b>Anexo 3: Dados demográficos e clínicos da população em estudo.....</b>                                                | <b>125</b> |
| <b>Anexo 4: Composição dos tampões utilizados .....</b>                                                                  | <b>131</b> |
| <b>Anexo 5: Protocolo <i>Whatman</i><sup>TM</sup> <i>FTA</i><sup>®</sup> <i>Elute</i><sup>TM</sup> .....</b>             | <b>132</b> |
| <b>Anexo 6: Nomenclatura dos nucleótidos e aminoácidos.....</b>                                                          | <b>133</b> |
| <b>Anexo 7: Mapas genómicos de formas recombinantes circulantes seleccionadas .....</b>                                  | <b>135</b> |
| <b>Anexo 8: Resultados da genotipagem rápida automatizada das sequências de referência CRF49_cpx utilizadas .....</b>    | <b>137</b> |

|                                                                                                                                                                |            |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Anexo 9: Resultados finais comparados da genotipagem das sequências em estudo para as regiões codificantes da protease e da transcriptase reversa .....</b> | <b>139</b> |
| <b>Anexo 10: Resultados da pesquisa de mutações de resistência aos ARVs .....</b>                                                                              | <b>141</b> |
| <b>Anexo 11: Perfis fenotípicos putativos de resistência aos ARVs.....</b>                                                                                     | <b>145</b> |
| <b>Anexo 12: Resultados da pesquisa de assinaturas genéticas relativos às mutações de resistência aos ARVs .....</b>                                           | <b>149</b> |

## Índice de figuras

---

|                                                                                                                                                                                                                                                                      |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figura 1:</b> Relações filogenéticas entre HIV-1 e SIVcpzPtt/SIVgor, e HIV-2 e SIVsmm. ....                                                                                                                                                                       | 2  |
| <b>Figura 2:</b> Coincidência geográfica entre o habitat natural dos primatas hospedeiros de SIVcpzPtt/SIVgor e SIVsmm e a região de onde são provenientes os primeiros casos de infecção com HIV-1 e a região endêmica do HIV-2, respectivamente. ....              | 3  |
| <b>Figura 3:</b> Distribuição global das pessoas a viver com HIV em 2017. ....                                                                                                                                                                                       | 4  |
| <b>Figura 4:</b> Diagrama representativo de uma partícula viral (imatura e madura) do HIV-1. ....                                                                                                                                                                    | 6  |
| <b>Figura 5:</b> Mapa genómico do HIV-1, na sua forma proviral. ....                                                                                                                                                                                                 | 7  |
| <b>Figura 6:</b> Esquema representativo do ciclo replicativo do HIV-1. ....                                                                                                                                                                                          | 9  |
| <b>Figura 7:</b> Representação da estrutura da protease do HIV-1. ....                                                                                                                                                                                               | 15 |
| <b>Figura 8:</b> Representação da estrutura da transcriptase reversa do HIV-1. ....                                                                                                                                                                                  | 16 |
| <b>Figura 9:</b> Conjunto das mutações associadas à resistência aos PIs (listadas pela IAS-USA), estando destacadas a negrito as mutações com maior relevância clínica. ....                                                                                         | 23 |
| <b>Figura 10:</b> Conjunto das mutações associadas a resistência aos NRTIs (listadas pela IAS-USA), estando destacadas a negrito as mutações com maior relevância clínica. ....                                                                                      | 24 |
| <b>Figura 11:</b> Conjunto das mutações associadas a resistência aos NNRTIs (listadas pela IAS-USA), estando destacadas a negrito as mutações com maior relevância clínica. ...                                                                                      | 25 |
| <b>Figura 12:</b> Mapa de Angola e respectivas províncias, ficando a região Norte acima da linha vermelha e a região Sul abaixo da linha azul. ....                                                                                                                  | 29 |
| <b>Figura 13:</b> Esquema representativo da amplificação das regiões codificantes da protease e da transcriptase reversa do gene <i>pol</i> do HIV-1, por realização de três PCRs consecutivas, para a obtenção de fragmentos de 460 e 650 pb, respectivamente. .... | 38 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figura 14:</b> Representação gráfica, em percentagem, dos dados demográficos da amostra populacional em estudo, relativamente a género, idade, estado civil e nível de escolaridade. ....                                                                                                                                                                                                                                                                   | 47 |
| <b>Figura 15:</b> Representação gráfica, em percentagem, dos dados demográficos e clínicos da amostra populacional em estudo, relativamente a local de residência, área de residência e outras infecções diagnosticadas. ....                                                                                                                                                                                                                                  | 48 |
| <b>Figura 16:</b> Representação gráfica dos resultados obtidos (em percentagem) após amplificação por dupla <i>nested</i> PCR das regiões codificantes da protease e da transcriptase reversa. ....                                                                                                                                                                                                                                                            | 49 |
| <b>Figura 17:</b> Visualização dos padrões de migração dos produtos de amplificação por dupla <i>nested</i> PCR das regiões codificantes da protease (à esquerda de MM) e transcriptase reversa (à direita de MM), após electroforese em gel de agarose a 1,2% (p/v), em tampão TAE 0,5 × e corado com brometo de etídeo (0,5 µg/mL).. ....                                                                                                                    | 50 |
| <b>Figura 18:</b> A) Relações filogenéticas entre 113 sequências nucleotídicas da região codificante da protease, baseadas num alinhamento múltiplo de 307 nucleótidos. B) Gráficos seleccionados da análise de <i>bootscanning</i> , das sequências indicadas em A) com uma seta roxa, utilizando os mesmos modelos aplicados à árvore filogenética, com recurso ao <i>software SimPlot</i> (versão 3.5.1) com janela 100 pb e passo 20 pb. ....              | 55 |
| <b>Figura 19:</b> A) Relações filogenéticas entre 117 sequências nucleotídicas da região codificante da transcriptase reversa, baseadas num alinhamento múltiplo de 532 nucleótidos. B) Gráficos seleccionados da análise de <i>bootscanning</i> , das sequências indicadas em A) com uma seta roxa, utilizando os mesmos modelos aplicados à árvore filogenética, com recurso ao <i>software SimPlot</i> (versão 3.5.1) com janela 100 pb e passo 20 pb. .... | 56 |
| <b>Figura 20:</b> Representação gráfica da proporção relativa dos genótipos do HIV-1 obtidos a partir da análise filogenética das sequências nucleotídicas de um total de 47 amostras (PR+RT).....                                                                                                                                                                                                                                                             | 63 |

|                                                                                                                                                                                                                                                                                   |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figura 21:</b> Esquema representativo de todas as mutações associadas a resistência encontradas nas sequências em estudo da região codificante da protease, incluindo principais acessórias e outras mutações.....                                                             | 65 |
| <b>Figura 22:</b> Representação gráfica dos níveis relativos de resistência (em percentagem) para os diferentes inibidores da protease, referentes às sequências em estudo da respectiva região codificante.....                                                                  | 66 |
| <b>Figura 23:</b> Esquema representativo de todas as mutações associadas a resistência encontradas nas sequências em estudo da região codificante da transcriptase reversa, incluindo mutações associadas a resistência apenas na presença de outras mutações. ..                 | 67 |
| <b>Figura 24:</b> Representação gráfica dos níveis relativos de resistência (em percentagem) para os diferentes inibidores da transcriptase reversa, em percentagem, nucleosídicos e não nucleosídicos, referentes às sequências em estudo da respectiva região codificante. .... | 69 |

## Índice de tabelas

---

|                                                                                                                                                                                                                                                                                        |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Tabela 1:</b> Resumo das funções das proteínas codificadas pelo genoma do HIV-1. ....                                                                                                                                                                                               | 7  |
| <b>Tabela 2:</b> <i>Primers</i> utilizados na 1. <sup>a</sup> e 2. <sup>a</sup> <i>nested</i> PCRs.....                                                                                                                                                                                | 36 |
| <b>Tabela 3:</b> Condições de amplificação aplicadas na primeira PCR. ....                                                                                                                                                                                                             | 40 |
| <b>Tabela 4:</b> Condições de amplificação aplicadas na segunda PCR.....                                                                                                                                                                                                               | 40 |
| <b>Tabela 5:</b> Condições de amplificação aplicadas na terceira PCR.....                                                                                                                                                                                                              | 40 |
| <b>Tabela 6:</b> Resultados da genotipagem automatizada para as regiões codificantes da protease e da transcriptase reversa utilizando o <i>software REGA HIV-1 Subtyping Tool</i> (versão 3.0). ....                                                                                  | 52 |
| <b>Tabela 7:</b> Número total de sequências e tamanho do alinhamento múltiplo (em nucleótidos) utilizados na construção das árvores filogenéticas PR e RT.....                                                                                                                         | 53 |
| <b>Tabela 8:</b> Resultados da genotipagem para as regiões codificantes da protease e da transcriptase reversa após construção de árvores filogenéticas. ....                                                                                                                          | 57 |
| <b>Tabela 9:</b> Resultados finais da genotipagem para as regiões codificantes da protease e da transcriptase reversa após conclusão da respectiva análise filogenética e da pesquisa de eventos de recombinação genética e/ou de hipermutação por ocorrência de hiperadenilação. .... | 61 |

## **Lista de abreviaturas, acrónimos e siglas**

---

|                 |                                                                                        |
|-----------------|----------------------------------------------------------------------------------------|
| <b>APOBEC</b>   | <i>Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like</i>                |
| <b>ARVs</b>     | <i>Anti-retrovirais</i>                                                                |
| <b>CA</b>       | <i>Capsid protein</i>                                                                  |
| <b>CD4</b>      | <i>Cluster of differentiation type 4</i>                                               |
| <b>CDC</b>      | <i>Centers for Disease Control and Prevention</i>                                      |
| <b>CDK9</b>     | <i>Cyclin-dependent kinase 9</i>                                                       |
| <b>CRF</b>      | <i>Circulating recombinant form</i>                                                    |
| <b>DC-SIGN</b>  | <i>Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin</i> |
| <b>DNA</b>      | <i>Ácido desoxirribonucleico</i>                                                       |
| <b>dNTP</b>     | <i>Desoxirribonucleotídeo trifosfatado</i>                                             |
| <b>EIA</b>      | <i>Ensaio imunoenzimático</i>                                                          |
| <b>EUA</b>      | <i>Estados Unidos da América</i>                                                       |
| <b>FDA</b>      | <i>U.S. Food and Drug Administration</i>                                               |
| <b>gp</b>       | <i>Glicoproteína</i>                                                                   |
| <b>HAART</b>    | <i>Highly active antiretroviral therapy</i>                                            |
| <b>HGB</b>      | <i>Hospital Geral de Benguela</i>                                                      |
| <b>HIV</b>      | <i>Human immunodeficiency virus</i>                                                    |
| <b>HMGI(Y)</b>  | <i>High-mobility group DNA-binding protein</i>                                         |
| <b>HTLV-III</b> | <i>Human T-lymphotropic virus type III</i>                                             |
| <b>IAS-USA</b>  | <i>International Antiviral Society-USA</i>                                             |
| <b>IHMT</b>     | <i>Instituto de Higiene e Medicina Tropical</i>                                        |
| <b>IN</b>       | <i>Integrase</i>                                                                       |
| <b>INSTIs</b>   | <i>Integrase strand transfer inhibitors</i>                                            |
| <b>IUPAC</b>    | <i>International Union of Pure and Applied Chemistry</i>                               |
| <b>LAV</b>      | <i>Lymphadenopathy-associated virus</i>                                                |
| <b>LTR</b>      | <i>Long terminal repeat</i>                                                            |
| <b>MA</b>       | <i>Matrix protein</i>                                                                  |
| <b>MAPK</b>     | <i>Mitogen-activated protein kinase</i>                                                |
| <b>mRNA</b>     | <i>RNA mensageiro</i>                                                                  |

|                               |                                                                                                                                        |
|-------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| <b>NC</b>                     | <i>Nucleocapsid protein</i>                                                                                                            |
| <b>NES</b>                    | <i>Nuclear export signal</i>                                                                                                           |
| <b>NF-κB</b>                  | <i>Nuclear factor κ-light-chain-enhancer of activated B cells</i>                                                                      |
| <b>NLS</b>                    | <i>Nuclear localization signal</i>                                                                                                     |
| <b>NNRTIs</b>                 | <i>Nonnucleoside reverse transcriptase inhibitors</i>                                                                                  |
| <b>NRTIs</b>                  | <i>Nucleoside reverse transcriptase inhibitors</i>                                                                                     |
| <b>OMS</b>                    | Organização Mundial da Saúde                                                                                                           |
| <b>pb</b>                     | Par de bases                                                                                                                           |
| <b>PBS</b>                    | <i>Primer binding site</i>                                                                                                             |
| <b>PCR</b>                    | <i>Polymerase chain reaction</i>                                                                                                       |
| <b>PIs</b>                    | <i>Protease inhibitors</i>                                                                                                             |
| <b>PIC</b>                    | <i>Pre-integration complex</i>                                                                                                         |
| <b>PPT</b>                    | <i>Polypurine tract</i>                                                                                                                |
| <b>PR</b>                     | Protease                                                                                                                               |
| <b>Pr55<sup>Gag</sup></b>     | Proteína precursora das proteínas estruturais de HIV-1 codificadas pelo gene <i>gag</i>                                                |
| <b>Pr160<sup>GagPol</sup></b> | Proteína precursora das proteínas estruturais e enzimáticas de HIV-1 codificadas, respectivamente, pelos genes <i>gag</i> e <i>pol</i> |
| <b>RDC</b>                    | República Democrática do Congo                                                                                                         |
| <b>RNA</b>                    | Ácido ribonucleico                                                                                                                     |
| <b>RNase</b>                  | Ribonuclease                                                                                                                           |
| <b>RRE</b>                    | <i>Rev responsive element</i>                                                                                                          |
| <b>RT</b>                     | Transcriptase reversa                                                                                                                  |
| <b>RTC</b>                    | <i>Reverse transcription complex</i>                                                                                                   |
| <b>RTIs</b>                   | <i>Reverse transcriptase inhibitors</i>                                                                                                |
| <b>SIV</b>                    | <i>Simian immunodeficiency virus</i>                                                                                                   |
| <b>SIVcpz<sup>Ptt</sup></b>   | SIV do chimpanzé-central (subespécie <i>Pan troglodytes troglodytes</i> )                                                              |
| <b>SIVgor</b>                 | SIV do gorila-ocidental-das-terras-baixas (subespécie <i>Gorilla gorilla gorilla</i> )                                                 |
| <b>SIVsmm</b>                 | SIV do mangabei-fuliginoso (espécie <i>Cercocebus atys</i> )                                                                           |
| <b>SU</b>                     | <i>Surface protein</i>                                                                                                                 |

|               |                                                   |
|---------------|---------------------------------------------------|
| <b>TA</b>     | Temperatura ambiente                              |
| <b>TAE</b>    | Tris-acetato-EDTA                                 |
| <b>TAM</b>    | <i>Thymidine analogue-associated mutation</i>     |
| <b>TM</b>     | <i>Transmembrane protein</i>                      |
| <b>tRNA</b>   | RNA de transferência                              |
| <b>U</b>      | <i>Unclassified</i>                               |
| <b>UNAIDS</b> | <i>Joint United Nations Programme on HIV/AIDS</i> |
| <b>URF</b>    | <i>Unique recombinant form</i>                    |

## 1 Introdução

---

### 1.1 Vírus da imunodeficiência humana e sida

A infecção crónica com o vírus da imunodeficiência humana (HIV, do inglês *human immunodeficiency virus*) está na origem da síndrome da imunodeficiência adquirida (sida) que, em consequência das dimensões sem precedentes atingidas, tem um forte impacto na saúde pública, tratando-se por isso de uma patologia de elevada importância ao nível da saúde global (1). A síndrome caracteriza-se por uma depleção gradual do número de linfócitos T auxiliares CD4<sup>+</sup> (do inglês, *cluster of differentiation type 4*) em circulação, resultante, em parte, da rápida replicação do vírus nas mesmas, levando a uma disfunção progressiva do sistema imunitário do indivíduo infectado associada ao aparecimento de infecções com agentes oportunistas e, frequentemente, alguns tipos de neoplasias (2–4).

Os primeiros casos de infecção com HIV foram descritos nos Estados Unidos da América (EUA) em 1981, após relatos de vários casos de patologias raras, como infecções com *Pneumocystis jirovecii*, sarcoma de Kaposi e linfadenopatias persistentes generalizadas, em grupos de jovens homossexuais previamente saudáveis que apresentavam um quadro clínico de grave deficiência imunitária de causa desconhecida (5–10). O aumento da incidência de novos casos, inclusive noutros grupos populacionais (e.g. consumidores de drogas por via intravenosa), levou ao reconhecimento da síndrome e à sua definição pelo CDC (do inglês, *Centers for Disease Control and Prevention*) como “uma doença, pelo menos moderadamente preditiva de um defeito na imunidade mediada por células, ocorrendo numa pessoa sem causa conhecida de menor resistência a essa doença” (11), utilizando-se pela primeira vez o acrónimo SIDA em 1982.

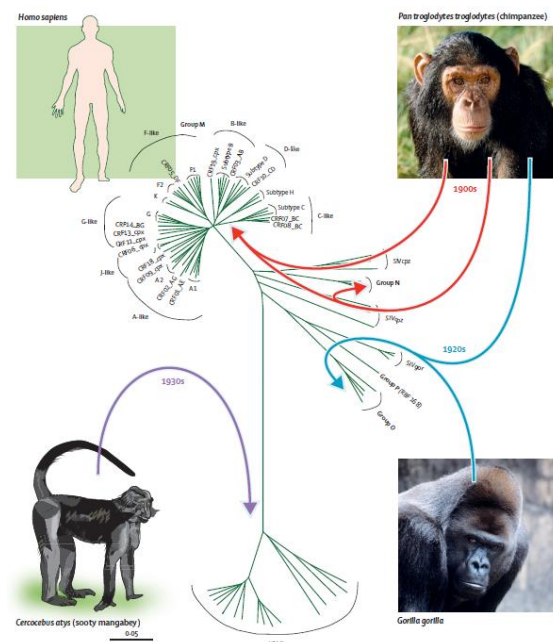
No ano seguinte, Barré-Sinoussi e colaboradores isolam pela primeira vez o agente etiológico da sida, então designado LAV (do inglês, *lymphadenopathy-associated virus*), a partir de um gânglio linfático de um indivíduo homossexual com linfadenopatia generalizada (12,13). Em 1984, uma outra equipa de investigadores reporta igualmente o isolamento do agente etiológico da sida designando-o de HTLV-III (do inglês, *human T-lymphotropic virus type III*) (14–17), verificando-se mais tarde que era idêntico a LAV. Surge em 1986 a proposta do Comité Internacional de Taxonomia de Vírus para a

designação oficial do vírus causador da sida como vírus da imunodeficiência humana (HIV) (18,19). No mesmo ano, é isolado um outro vírus em indivíduos com sida oriundos da África Ocidental, semelhante ao então conhecido HIV mas geneticamente diferente do mesmo, apresentando características antigénicas e relações filogenéticas distintas (20–23). O HIV passou assim a ser classificado em dois tipos distintos: HIV-1 e HIV-2 – o primeiro é responsável pela pandemia, enquanto o segundo é endémico e permanece confinado a países da África Ocidental, possuindo apenas alguns focos noutras regiões com ligações aos países endémicos (1,24–26).

## 1.2 Origem biológica do HIV

O HIV teve origem em múltiplas transmissões de carácter zoonótico de vírus da imunodeficiência símia (SIV, do inglês *simian immunodeficiency virus*) de primatas da África Central-Occidental para humanos, sendo a exposição cutânea, ou das membranas mucosas, a sangue ou secreções de animais infectados – em resultado de actividades como a caça e manuseamento de carne dos primatas para consumo alimentar ou a captura e troca dos animais que são também mantidos como animais de estimação – apontada como a causa mais plausível de transmissão (1,27,28). Diversos estudos permitiram verificar uma elevada semelhança filogenética do HIV-1 com o SIV do chimpanzé-central (subespécie *Pan troglodytes troglodytes*) (SIVcpz<sub>Ptt</sub>) e o SIV do gorila-occidental-das-terras-baixas (subespécie *Gorilla gorilla gorilla*) (SIVgor), e do HIV-2 com o SIV do mangabeí-fuliginoso (espécie *Cercocebus atys*) (SIVsmm) (1,23,29–31).

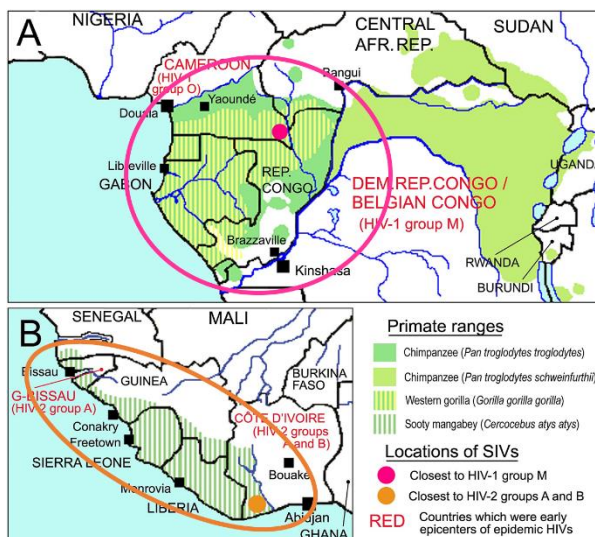
Análises filogenéticas revelaram que os dois tipos de HIV formam linhagens evolutivas monofiléticas com os respectivos SIVs (Figura 1) (23,24,30), verificando-se ainda uma coincidência geográfica entre o habitat natural dos primatas hospedeiros naturais dos SIVs referidos (32) e as áreas de maior



**Figura 1: Relações filogenéticas entre HIV-1 e SIVcpzPtt/SIVgor, e HIV-2 e SIVsmm – retirado de (24).**

diversidade genética dos HIVs onde terá ocorrido a transmissão cruzada dos vírus para os humanos. As espécies portadoras de SIVcpzPtt e SIVgor são encontradas na região da África Central, entre os Camarões e a República Democrática do Congo (RDC), de onde são provenientes os primeiros casos de infecção com HIV-1 (24,27,30), e a região endémica do HIV-2, na costa ocidental africana, entre o Senegal e a Costa do Marfim, coincide com o habitat natural da espécie portadora de SIVsmm (28,33,34) (Figura 2).

A elevada prevalência de SIVcpz<sub>Ptt</sub>, SIVgor e SIVsmm nos seus hospedeiros naturais (23,27) é indicativa de um longo período de circulação destes vírus nas respectivas espécies hospedeiras, o que reforça a sua ancestralidade em relação aos respectivos HIVs. Pelos motivos apresentados, foi proposto que o HIV-1 terá tido origem em transmissões cruzadas do SIVcpz<sub>Ptt</sub> e/ou do SIVgor em pelo menos quatro ocasiões



**Figura 2: Coincidência geográfica entre o habitat natural dos primatas hospedeiros de SIVcpzPtt/SIVgor e SIVsmm e a região de onde são provenientes os primeiros casos de infecção com HIV-1 e a região endêmica do HIV-2, respectivamente – adaptado de (32).**

diferentes, e que o HIV-2 terá resultado de transmissões cruzadas do SIVsmm em pelo menos nove eventos independentes (23). Estudos recentes colocam Kinshasa, capital da RDC, no epicentro da epidemia, indicando que terá sido a partir desta cidade que a infecção com HIV-1 se terá propagado para outras regiões da África Central, inicialmente a um ritmo lento, desde 1920, aproximadamente (35,36). Após 1960, verificou-se um aumento na velocidade de propagação do vírus que se expandiu um pouco por todo o

mundo, tendo sido apontadas como possíveis causas o aumento dos fenómenos migratórios entre diferentes países, as alterações nos comportamentos sexuais (verificando-se um aumento da promiscuidade e da prevalência de infecções sexualmente transmissíveis) e a implementação em África de planos de vacinação e programas terapêuticos em grande escala, onde a utilização repetitiva de seringas não esterilizadas era uma prática comum (35–37).

### 1.3 Dados epidemiológicos globais

Desde o início da epidemia nos anos 80 do século passado, 77,3 milhões [59,9–100 milhões] de pessoas foram infectadas com HIV, tendo morrido até à data cerca de 35,4 milhões [28,9–41,5 milhões] de pessoas com patologias relacionadas com sida (38). Estimativas recentes da UNAIDS (do inglês, *Joint United Nations Programme on HIV/AIDS*) indicam que em 2017 existiam 36,9 milhões [31,1–43,9 milhões] de pessoas infectadas com HIV, das quais cerca de 75% [55-92%] tinham conhecimento da sua condição (38,39). A maioria da população infectada encontra-se concentrada no continente africano (Figura 3), em particular na África Oriental e Meridional, onde 19,6 milhões [17,5–22,0 milhões] de pessoas vivem com HIV (correspondendo a 53% da totalidade das mesmas em todo o mundo) e onde foi contabilizado cerca de 45% do total de novas infecções a nível global (38,39). Para esse mesmo ano, estimou-se que o número de novas infecções era de 1,8 milhões [1,4–2,4 milhões], correspondendo a uma diminuição de 47% desde o seu pico em 1996, e que o número de mortes por causas relacionadas com sida era de 940 000 [670 000–1,3 milhões], indicando uma diminuição superior a 51% desde o seu pico em 2004 (38,39). Segundo dados da Organização Mundial da Saúde (OMS), a prevalência global da infecção por HIV era de 0,8% [0,6–0,9%] em 2017, sendo superior em África, onde a prevalência é de 4,1% [3,4–4,8%].

A nível global, das pessoas que em 2017 viviam com sida e tinham conhecimento da sua condição, 79% estavam em tratamento e, dessas, 81% tinham cargas virais indetectáveis, verificando-se que mais de metade das pessoas que viviam com HIV (59% [44–73%]) tinham acesso à terapia anti-retroviral (38). Este último valor corresponde a um total estimado de 21,7 milhões [19,1–22,6 milhões] de pessoas, observando-se um aumento em 2,3 milhões desde o final de 2016 (38,39).

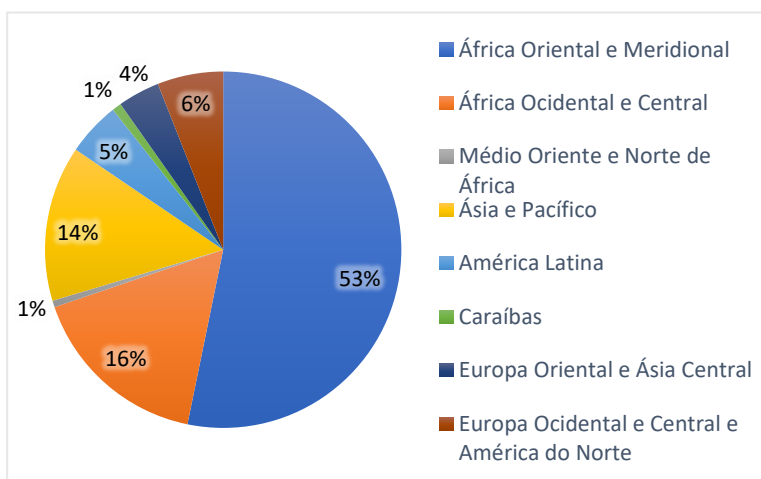


Figura 3: Distribuição global das pessoas a viver com HIV em 2017 (38).

Os dados disponíveis sugerem ainda que 47% das novas infecções, a nível global em 2017, ocorreram nas populações mais vulneráveis – homens que fazem sexo com homens, consumidores de drogas injectáveis, profissionais do sexo, prisioneiros e indivíduos transgénero – e nos seus parceiros sexuais, verificando-se também que 1,8 milhões [1,3–2,4 milhões] de crianças e jovens com menos de 15 anos de idade ficaram infectados e que apenas metade dos mesmos se encontravam em tratamento (38,39).

### 1.4 Taxonomia e características gerais do HIV

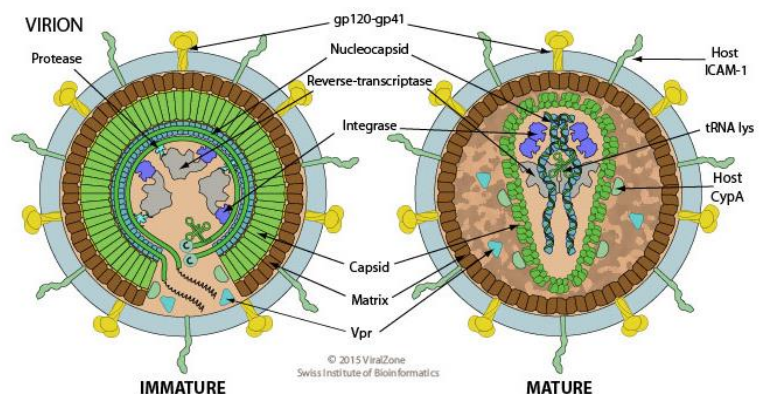
Taxonomicamente, o HIV-1 e o HIV-2 pertencem à ordem *Ortervirales*, família *Retroviridae*, subfamília *Orthoretrovirinae*, género *Lentivirus* (40,41), integrando o grupo dos lentivírus dos primatas. Ambos possuem como vias de transmissão o contacto sexual sem protecção, a exposição parentérica – que inclui a inoculação percutânea e transfusões com sangue contaminado – e a transmissão vertical de mãe para filho, tanto durante a gestação ou no momento do parto como através da amamentação (2,3,42). O diagnóstico da infecção pode ser realizado por pesquisa de anticorpos anti-HIV no soro, plasma, sangue total ou fluido oral, com recurso, mais comum, a ensaios imunoenzimáticos (EIA, do inglês *enzyme immunoassay*) ou imunocromatográficos, ou por pesquisa directa do vírus. Esta pode ser realizada através da detecção do genoma viral – tanto de RNA (ácido ribonucleico) viral, como de DNA (ácido desoxirribonucleico) proviral, com recurso a técnicas de PCR (do inglês, *polymerase chain reaction*) – ou do antígeno p24 (proteína da cápside viral, de expressão precoce), ou, muito raramente, através do isolamento do vírus *in vitro*. A história clínica do indivíduo tem importância na selecção dos testes de diagnóstico mais adequados (43–45).

O HIV infecta principalmente linfócitos T auxiliares e macrófagos via o receptor CD4 e os co-receptores CCR5 e CXCR4, entre outros, apresentando tropismo diferencial para estes últimos. Na maioria dos casos, para infectar estas células, o vírus interage primeiramente com células dendríticas da submucosa através do receptor DC-SIGN (do inglês, *dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin*), as quais são importantes células apresentadoras de antígenos e também podem ser infectadas, uma vez que expressam igualmente o receptor CD4 e o co-receptor CCR5. Da infecção das células dendríticas resulta a apresentação antigénica de epitopos virais às células-alvo, que ocorre após maturação e migração destas células para os gânglios

linfáticos mais próximos do local primário da infecção (46–48). Ambos os vírus levam a falência imunológica, conduzindo o indivíduo infectado ao estado de sida. No entanto, o HIV-2 é menos virulento e transmissível que o HIV-1, caracterizando-se a sua infecção por um período assintomático mais longo e cargas virais mais baixas, o que resulta numa taxa de progressão de doença mais lenta, justificando, em parte, a sua distribuição geográfica limitada e menor prevalência (49–53).

## 1.5 Estrutura e organização genómica do HIV-1

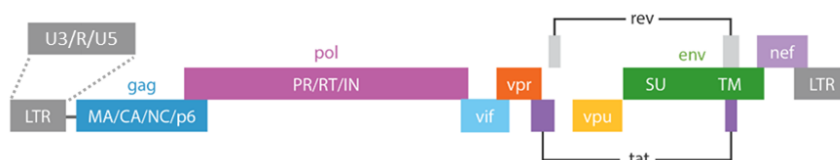
A estrutura dos viriões do HIV-1 (Figura 4) é aproximadamente esférica, com cerca de 100 nm de diâmetro, apresentando uma simetria complexa e um invólucro lipídico, proveniente da membrana citoplasmática das células infectadas. O invólucro contém as glicoproteínas de superfície gp120 (ou SU, do inglês *surface protein*) e transmembranar gp41 (ou TM, do inglês *transmembrane protein*) unidas por ligações não covalentes sob a forma de homotrímeros de heterodímeros. Estas glicoproteínas virais derivam de um único precursor poliproteico (gp160) e estão envolvidas na entrada do vírus nas células-alvo. Envolvida pelo invólucro lipídico encontra-se a matriz viral, formada pela proteína p17 (ou MA, do inglês *matrix protein*), subjacente à qual se encontra a cápside viral, composta pela proteína p24 (ou CA, do inglês *capsid protein*), que possui a forma de um cone truncado. No interior desta última encontra-se o genoma viral, ao qual está associada a proteína da nucleocápside (p7 ou NC, do inglês *nucleocapsid protein*), as enzimas virais, nomeadamente, a protease (PR), a transcriptase reversa (RT) e a integrase (IN), a proteína p6 e as proteínas virais acessórias Vpu, Vif, Vpr e Nef (3).



**Figura 4: Diagrama representativo de uma partícula viral (imatura e madura) do HIV-1 – disponível em:** [http://viralzone.expasy.org/resources/HIV\\_virion.jpg](http://viralzone.expasy.org/resources/HIV_virion.jpg).

do inglês *nucleocapsid protein*), as enzimas virais, nomeadamente, a protease (PR), a transcriptase reversa (RT) e a integrase (IN), a proteína p6 e as proteínas virais acessórias Vpu, Vif, Vpr e Nef (3).

O genoma viral é dimérico, composto por duas moléculas idênticas de RNA de cadeia simples e polaridade positiva com cerca de 9,5 kb cada, possuindo um total de nove genes codificando para diferentes proteínas e enzimas virais (2,3), distinguindo-se os genomas de HIV-1 e HIV-2 pela presença de um gene não homólogo – *vpu* para o HIV-1 e *vpx* para o HIV-2 (54,55). O genoma possui estruturas *cap* e cauda *polyA* nas



**Figura 5: Mapa genômico do HIV-1, na sua forma proviral – adaptado de [https://www.abmgood.com/marketing/knowledge\\_base/img/Lentivirus\\_System/lenti\\_virus\\_genome\\_map.png](https://www.abmgood.com/marketing/knowledge_base/img/Lentivirus_System/lenti_virus_genome_map.png).**

extremidades 5' e 3' respectivamente, e a forma dimérica é devida à sua associação às proteínas da NC e ao emparelhamento de sequências invertidas. Na forma de DNA proviral (Figura 5) – molécula de cadeia dupla resultante da retrotranscrição do RNA viral – possui, em ambas as extremidades, regiões reguladoras da expressão gênica designadas LTR (do inglês, *long terminal repeats*) – implicadas não só na transcrição viral, mas anteriormente também na retrotranscrição e integração. As LTR têm cerca de 600 pb e estrutura U3-R-U5, apresentando na extremidade 5' a região PBS (do inglês, *primer binding site*) e na extremidade 3' a região PPT (do inglês, *polypurine tract*). A região codificante do genoma possui uma localização central e nove genes: *gag*, codificante para as proteínas estruturais do virião – MA, CA e NC – e ainda a proteína p6; *pol*, que codifica as enzimas virais PR, RT e IN; *env*, codificante para as glicoproteínas do invólucro SU e TM; *tat* e *rev*, que codificam as proteínas reguladoras Tat e Rev, respectivamente; e *vif*, *vpr*, *vpu* e *nef*, codificantes para as proteínas acessórias Vif, Vpr, Vpu e Nef respectivamente. As principais funções das proteínas codificadas por cada um dos genes encontram-se resumidas na Tabela 1 (3,56–59).

**Tabela 1: Resumo das funções das proteínas codificadas pelo genoma do HIV-1.**

|                   |            |                                                                                                                                                                                                                                                           |
|-------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Genes estruturais | <i>gag</i> | <b>MA (p17):</b> Proteína estrutural da matriz; envolvida no transporte do complexo de pré-integração para o núcleo; encaminha as poliproteínas Gag e Gag-Pol para a membrana citoplasmática e promove a sua incorporação na partícula viral em formação. |
|                   |            | <b>CA (p24):</b> Proteína estrutural da cápside; envolvida na montagem do virião.                                                                                                                                                                         |

|                          |            |                                                                                                                                                                                            |
|--------------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                          |            | <b>NC (p7):</b> Proteína estrutural da nucleocápside; envolvida no empacotamento do RNA genómico e na montagem do virião; chaperone de RNA.                                                |
|                          |            | <b>p6:</b> Envolvida na montagem e libertação do virião da célula; promove a incorporação da proteína Vpr no virião.                                                                       |
|                          | <i>pol</i> | <b>PR (p10):</b> Responsável pelo processamento proteolítico das poliproteínas Gag e Gag-Pol, durante a maturação do virião e aquisição de infecciosidade.                                 |
|                          |            | <b>RT (p66/51):</b> Responsável pela retrotranscrição do genoma viral.                                                                                                                     |
|                          |            | <b>IN (p32):</b> Responsável pela inserção do DNA complementar no genoma da célula (formação do provírus).                                                                                 |
|                          | <i>env</i> | <b>SU (gp120):</b> Glicoproteína de superfície; intervém na ligação do vírus à célula-alvo – ligação a receptores (CD4) e co-receptores (CCR5/CXCR4) celulares.                            |
| <b>Genes reguladores</b> |            | <b>TM (gp41):</b> Glicoproteína transmembranar; intervém na entrada do vírus na célula-alvo – fusão do invólucro lipídico com a membrana citoplasmática.                                   |
|                          | <i>tat</i> | <b>Tat (p14):</b> Activador transcricional – promove a activação da RNA polimerase II celular para a transcrição do DNA proviral.                                                          |
|                          | <i>rev</i> | <b>Rev (p19):</b> Envolvida na exportação nuclear de transcritos virais não processados, ou parcialmente processados, para o citoplasma.                                                   |
| <b>Genes acessórios</b>  | <i>vif</i> | <b>Vif (p23):</b> Promove a infecciosidade viral (envolvida na degradação da proteína APOBEC3G) e a paragem do ciclo celular na fase G2.                                                   |
|                          | <i>vpr</i> | <b>Vpr (p15):</b> Envolvida no transporte do complexo de pré-integração para o núcleo; promove paragem do ciclo celular na fase G2/M; induz apoptose.                                      |
|                          | <i>vpu</i> | <b>Vpu (p16):</b> Promove a degradação de CD4, facilitando a libertação do virião da célula infectada.                                                                                     |
|                          | <i>nef</i> | <b>Nef (p27):</b> Modulação negativa de CD4 e MHC I e II, promovendo o escape imunitário; promove a infecciosidade viral; inibe a apoptose; determinante multifactorial de patogenicidade. |

## 1.6 Ciclo replicativo do HIV-1

O ciclo replicativo do HIV-1 (Figura 6) consiste numa série de eventos sequenciais que ocorrem, na sua maioria, dentro da célula infectada, com vista à produção de novos viriões capazes de infectar outras células-alvo, podendo esses eventos ser divididos em duas fases, precoce e tardia, consoante ocorram, respectivamente, antes ou depois do início da replicação do genoma viral (60).

O ciclo replicativo viral inicia-se com a ligação (ou adsorção) dos viriões às células-alvo através da interacção da gp120 com o receptor celular CD4. A consequente

alteração conformacional da gp120 permite a sua interação com o co-receptor CCR5 ou CXCR4 (dependo do tropismo viral). A ligação da gp120 ao co-receptor induz um rearranjo conformacional na gp41, expondo o péptido de fusão que se insere na membrana citoplasmática e leva à fusão desta última com invólucro lipídico viral, resultando na libertação da cápside viral no citoplasma da célula (56,58,60).

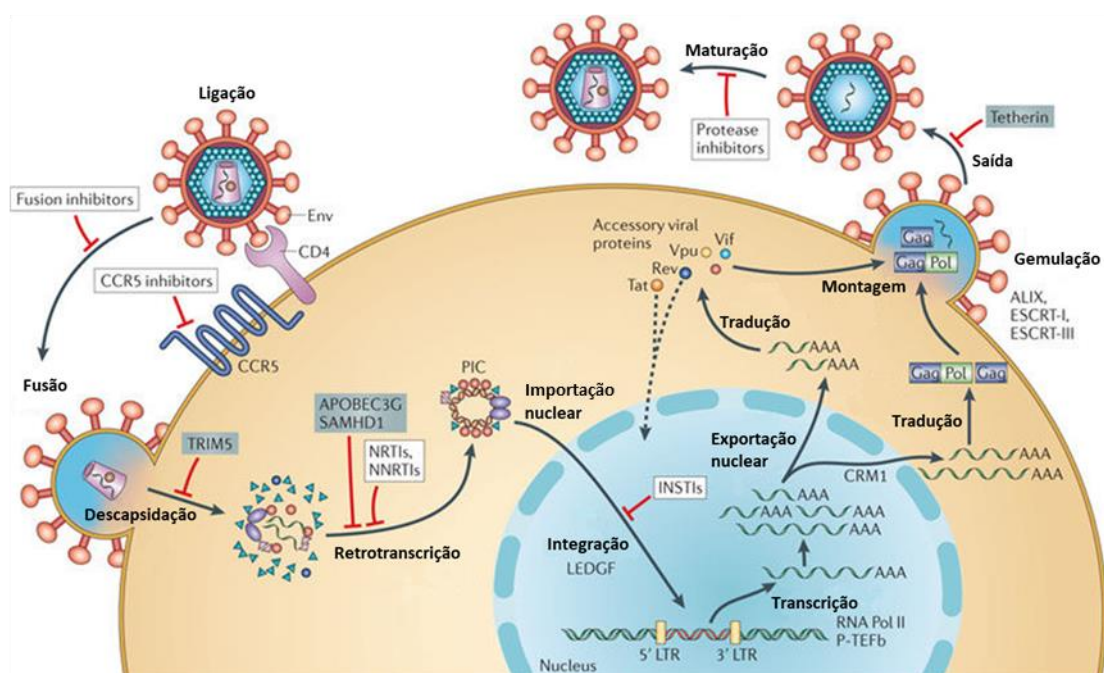


Figura 6: Esquema representativo do ciclo replicativo do HIV-1 – adaptado de <http://www.nature.com/nrmicro/journal/v10/n4/images/nrmicro2747-f1.jpg>.

Após a entrada, segue-se a descapsidação, na qual ocorre uma degradação parcial da cápside viral – envolvendo os factores celulares MAPK (do inglês, *mitogen-activated protein kinase*) e ciclofilina A e a proteína viral MA (46,56) – que resulta na formação do complexo de transcrição reversa (RTC, do inglês *reverse transcription complex*), composto pelas duas moléculas de RNA genómico revestidas pela NC, o *primer* de origem celular tRNA<sub>Lys3</sub>, a RT, a IN e as proteínas virais MA e Vpr (56,61). O RTC fica depois ancorado aos microfilamentos de actina do citoesqueleto e a RT inicia a retrotranscrição do genoma viral para a síntese do DNA complementar, o qual é depois transportado para o núcleo da célula infectada (utilizando a rede de microtúbulos) após a formação do complexo de pré-integração (PIC, do inglês *pre-integration complex*) – composto pelo DNA complementar, IN, RT, MA, Vpr e a proteína celular HMGI(Y) (do inglês, *high-mobility group DNA-binding protein*), entre outras (56,58,60,61). Depois da

entrada do PIC no núcleo através de um poro nuclear intacto, o DNA complementar é integrado no genoma da célula infectada por acção da IN – tendo sido identificados diferentes factores celulares importantes nesta etapa do ciclo replicativo (56,58,62,63) – ou circulariza, formando círculos 1-LTR ou 2-LTR, transcricionalmente activos, a partir dos quais parecem ser produzidas, mas de modo basal, as proteínas virais Nef, Tat e Rev (3,56,64).

Uma vez integrado no genoma celular, o DNA proviral é transcrito pela RNA polimerase II celular, após activação das LTR por factores de transcrição celulares, como, entre outros, o NF- $\kappa$ B (do inglês, *nuclear factor  $\kappa$ -light-chain-enhancer of activated B cells*) (3), produzindo-se, inicialmente, quantidades basais das proteínas virais Nef, Tat e Rev, cujos mRNA (RNA mensageiro) sofrem processamento de *splicing* múltiplo. Depois de produzida em quantidades suficientes, a proteína Tat, em associação com a proteína celular ciclina T1, liga-se ao elemento TAR (do inglês, *transactivation response region*), existente na extremidade 5' do DNA proviral, e recruta a cinase CDK9 (do inglês, *cyclin-dependent kinase 9*) para a fosforilação da extremidade C-terminal da RNA polimerase II, o que resulta no grande aumento da sua processividade e permite o alongamento dos transcritos (56,60) – que podem ou não sofrer *splicing*, originando mRNAs não processados e mRNAs parcial ou totalmente processados (60,65) – e a replicação do genoma viral. O transporte para o citoplasma dos mRNAs não processados (genoma viral) e parcialmente processados, que, ao contrário dos mRNAs totalmente processados, ficam retidos no núcleo, é assegurado pela proteína Rev – detentora dos domínios NLS e NES (do inglês, *nuclear localization signal* e *nuclear export signal*) – que reconhece e se liga à estrutura secundária de RNA RRE (do inglês, *Rev responsive element*) localizada no gene *env* e presente nesses transcritos (56,60,66).

Os diferentes mRNAs são depois traduzidos e produzidas as diferentes proteínas virais, tanto no citoplasma, onde é produzida a maioria das proteínas virais, como no retículo endoplasmático rugoso, onde é produzido e processado o precursor poliproteico gp160 que dará origem às proteínas gp120 e gp41. As proteínas estruturais precursoras são transportadas para junto da membrana citoplasmática para a posterior montagem dos viriões, a qual é coordenada maioritariamente por Pr55<sup>Gag</sup> e Pr160<sup>Gag-Pol</sup> e outros factores celulares, sendo ainda influenciada pelas glicoproteínas do invólucro viral e envolvendo o recrutamento de duas moléculas de RNA viral (para a constituição do genoma viral) e

da proteína Vpr (3,56,60). A saída das partículas virais imaturas, não infecciosas, da célula infectada ocorre por gemulação e envolve as proteínas p6 e Vpu, terminando o ciclo replicativo do HIV-1 com a sua maturação (que pode ou não coincidir com a saída da célula), durante a qual ocorre um rearranjo estrutural das proteínas Pr55<sup>Gag</sup> e Pr160<sup>Gag-Pol</sup> mediado pela PR (que é activada aquando da gemulação) produzindo-se novos viriões (3,56,60).

### 1.7 Diversidade genética

#### 1.7.1 Classificação do HIV-1

Com base na análise filogenética de numerosas sequências genómicas de várias estirpes virais com diferentes origens geográficas, foi possível classificar o HIV-1 em diferentes grupos, subtipos, sub-subtipos e ainda diversas formas recombinantes (67). Desse modo, identificam-se quatro linhagens filogenéticas distintas do HIV-1 resultantes da transmissão dos respectivos SIVs para humanos em quatro ocasiões diferentes: os grupos M, N, O e P (M e N têm origem no SIVcpz<sub>Ptt</sub> e O e P no SIVgor) (1,24,30). Em relação aos grupos N e P, foi reportado um número reduzido de infecções somente nos Camarões, num total de vinte indivíduos infectados com vírus do grupo N e dois indivíduos infectados com vírus do grupo P (68–71). O grupo O encontra-se na sua maioria restrito à África Central-Occidental, em particular aos Camarões e países vizinhos onde representa cerca de 1 a 2% de todas as infecções com HIV-1, sendo a sua prevalência relativamente baixa e estimando-se que tenha infectado um total de 100 000 indivíduos (72,73).

O grupo M é responsável pela pandemia, contabilizando mais de 90% das infecções a nível global, e apresenta a maior diversidade genética dos quatro grupos de HIV-1, incluindo nove subtipos designados A-D, F-H, J e K (1,24,74). Adicionalmente, os subtipos A e F incluem sub-subtipos, A1-A4, F1 e F2, sendo importante referir que os subtipos B e D deveriam, na verdade, corresponder também a dois sub-subtipos do mesmo subtipo genético, uma vez que possuem relações filogenéticas mais próximas entre si do que com os restantes subtipos – no entanto, foi conservada a nomenclatura original de modo a manter a concordância com os artigos previamente publicados (67,75,76). Existem também várias formas recombinantes únicas (URFs, do inglês *unique recombinant forms*) e circulantes (CRFs, do inglês *circulating recombinant forms*),

resultantes da recombinação genética entre os diferentes subtipos e sub-subtipos, em consequência de episódios de co-infecção (simultânea ou sequencial – superinfecção) num mesmo indivíduo, favorecidos pela sua co-circulação em áreas geográficas restritas (2,77–79). As CRFs correspondem a vírus com genomas mosaico, idênticos entre si, que foram identificados em pelo menos três indivíduos não relacionados epidemiologicamente, sendo, por esse motivo, reconhecida como epidemiologicamente relevante a sua contribuição para a epidemia (67,80,81), estando descritas até à data cerca de 98 CRFs (82). Por outro lado, as URFs correspondem a vírus com genomas mosaico que foram identificados num único indivíduo ou em conjuntos de indivíduos epidemiologicamente relacionados (80,81). Em relação à nomenclatura, as CRFs são designadas com um número de identificação (números consecutivos por ordem de descoberta) seguido dos subtipos, ou CRFs, se recombinantes de segunda geração, envolvidos (*e.g.*, CRF02\_AG).

### 1.7.2 Distribuição mundial

A distribuição geográfica das diferentes formas genéticas do HIV-1, em particular do grupo M, apresenta um padrão específico, ainda que heterogéneo, que reflecte a história da pandemia, em consequência, entre outros, de uma via de transmissão mais comum ou de efeitos fundadores resultantes de uma migração viral accidental (78). A nível global, os subtipos de HIV-1 do grupo M com maior prevalência são A, B e C, sendo este último o mais frequente – representando cerca de metade das infecções no mundo – apesar de ser o subtipo B aquele que está mais representado nas bases de dados (78,83–85). O subtipo C encontra-se distribuído um pouco por todo o mundo, mas sendo particularmente predominante na África Meridional e Oriental e na Índia, onde é responsável pela esmagadora maioria das infecções (78,83,86). O subtipo A é predominante nos países da Europa de Leste que constituíam a antiga União Soviética e ainda nalguns países da África Central e Oriental (78,84,87). Por outro lado, o subtipo B – que se pensa derivar de uma introdução pontual de um vírus relacionado com subtipo D no Haiti e EUA (88) – é predominante na Europa Central e Ocidental, nas Américas, especialmente do Norte e Central, e na Austrália, sendo igualmente comum em diferentes países do norte de África, Sudeste Asiático e Médio Oriente, e em determinadas populações da Rússia (78,84,87). As CRFs com maior relevância na pandemia são a CRF01\_AE, predominante

no Sudeste Asiático, e a CRF02\_AG, com predomínio na África Oriental e Centro-Oriental, verificando-se ainda uma contribuição crescente da CRF07\_BC também encontrada em circulação no Sudeste Asiático e na China (74,78,83–86).

Em África, verifica-se uma distribuição regional dos diferentes subtipos e formas recombinantes relativamente estável – predominando na África Oriental, os subtipos A e D; na África Meridional, o subtipo C; na África Ocidental, os subtipos A e G e as formas recombinantes CRF02\_AG e CRF06\_cpx; e no norte de África, o subtipo B e a forma recombinante CRF02\_AG –, sendo os subtipos A, C e D e a forma recombinante CRF02\_AG responsáveis por cerca de 85% do total de novas infecções (85,89).

### 1.7.3 Mecanismos responsáveis e impacto da diversidade genética

Uma das características mais relevantes do HIV-1 é a sua elevada diversidade genética, que resulta na existência no indivíduo infectado de um conjunto de estirpes virais distintas, mas geneticamente relacionadas, com propriedades antigénicas e fenotípicas ligeiramente diferentes, designado de *quasispecies*, sobre o qual actua a selecção natural favorecendo os vírus com melhor *fitness* (90,91). Diversos factores têm contribuído para o aumento da diversidade genética do HIV-1, que possui como principais causas:

- ❖ A rápida replicação viral, estimando-se que sejam produzidos cerca de  $10^{10}$  viriões por dia num indivíduo infectado, com tempos médios de semi-vida muito curtos (92);
- ❖ Elevada taxa de erros, principalmente durante a retrotranscrição, devido à ausência de actividade correctora 3'-exonuclease (*proofreading*) da RT e como resultado da actividade de factores mutagénicos do hospedeiro como a APOBEC3G, estimando-se que sejam introduzidas, em média,  $10^{-5}$ - $10^{-4}$  mutações por nucleótido incorporado, por ciclo replicativo (90,93);
- ❖ Eventos de recombinação genética promovidos pela RT em consequência de dupla infecção e de fenómenos de *template switching*, devido à capacidade da enzima “saltar” entre as duas moléculas de RNA genómico durante a retrotranscrição (facilitada pela dimerização do genoma), o que permite a utilização de ambas as cadeias como matriz na síntese do DNA proviral (90,93),

estimando-se que ocorram aproximadamente dois a três eventos de recombinação, por genoma, por ciclo replicativo (94);

- ❖ Elevadas pressões selectivas ambientais por parte do sistema imunitário do indivíduo infectado e da terapia anti-retroviral, que resultam na selecção e fixação das estirpes virais com mutações de escape imunitário e de resistência apresentando a melhor *fitness* (90,93,95).

A elevada diversidade genética verificada neste vírus possui um impacto considerável em vários aspectos, incluindo a terapêutica e a progressão da doença. Resulta, muito frequentemente, no surgimento de mutações de resistência e consequentes diferenças na resposta à terapêutica, podendo até levar à falência da mesma, tendo-se inclusive verificado diferenças na resposta à terapêutica entre subtipos. Outras consequências possíveis incluem diferenças na velocidade de progressão para sida e nas taxas de transmissão do vírus, dificuldades acrescidas no desenvolvimento de uma vacina eficaz, tendo já sido utilizadas diversas abordagens com pouco ou nenhum sucesso, e ainda uma sensibilidade diferencial das diferentes formas genéticas nos testes de diagnóstico e de carga viral, cuja performance é afectada, podendo resultar em resultados falsos negativos e medições erróneas da carga viral que, por sua vez, condicionam a decisão terapêutica (81,86,90,91,95,96).

### 1.8 Gene *pol*

O gene *pol* é uma das regiões mais bem conservadas do genoma do HIV-1, possuindo um total de 3012 nucleótidos, dos quais 2844 codificam as enzimas virais PR, RT e IN (97). Este gene não possui um codão de iniciação e encontra-se parcialmente sobreposto ao gene *gag* numa grelha de leitura diferente, motivo pelo qual é transcrito juntamente com este último, designando-se de Gag-Pol o mRNA obtido (98). Devido à existência de um codão *stop* no final da sequência codificante do gene *gag*, é necessária uma mudança de grelha de leitura durante a tradução de Gag-Pol para que os produtos do gene *pol* sejam expressos, o que resulta na síntese de uma poliproteína de fusão de 160 kDa designada Pr160<sup>Gag-Pol</sup> (98,99). A necessidade de uma mudança de grelha de leitura leva a que a síntese de Pr55<sup>Gag</sup> (poliproteína precursora das proteínas codificadas pelo gene *gag*) ocorra em maior percentagem do que a síntese de Pr160<sup>Gag-Pol</sup> (95% vs. 5%)

(100), o que é, no entanto, favorável para a eficiência do ciclo replicativo viral, uma vez que é necessária uma grande quantidade de proteínas estruturais para a montagem dos viriões.

## 1.8.1 Protease

A PR do HIV-1 é uma aspartil protease, constituindo uma das enzimas proteolíticas mais extensivamente estudadas nas últimas três décadas (101). Na sua forma activa, a PR apresenta-se sob a forma de um homodímero simétrico (Figura 7) formado por dois monómeros idênticos de 99 aminoácidos, cuja actividade catalítica depende da manutenção desta estrutura dimérica – estabilizada pela interacção entre os terminais N e C de cada monómero (resíduos aminoácidos 1-4 e 96-99) (101–103). O centro activo da enzima localiza-se na interface entre os dois monómeros e possui a tríade catalítica D25-T26-G27 (sequência altamente conservada nas proteases aspárticas), contribuindo cada monómero com um resíduo catalítico (D25 e D25') (101,102,104). Outra característica da PR é a presença, acima do centro activo, de estruturas flexíveis em ambos os monómeros, designadas *flap*, que interagem entre si mudando a sua orientação (abrindo e fechando) de modo a permitir a ligação e estabilização dos substratos (ou inibidores) relativamente ao centro activo da enzima (102–105).

A PR desempenha um papel fundamental no ciclo replicativo viral, nomeadamente na maturação e aquisição de infecciosidade dos viriões, sendo responsável pela clivagem das poliproteínas Pr55<sup>Gag</sup> e Pr160<sup>Gag-Pol</sup> (3,57,106). Consequentemente, a sua inibição resulta na produção de partículas virais imaturas e não infecciosas, razão pela qual se trata de um excelente alvo terapêutico (105). A descoberta de resíduos de aminoácidos bem conservados críticos para a ligação do substrato ao centro activo da enzima (D25, G27, A28 e D29, localizados no centro activo, e G49, localizado na região flexível *flap*), permitiu a respectiva classificação como

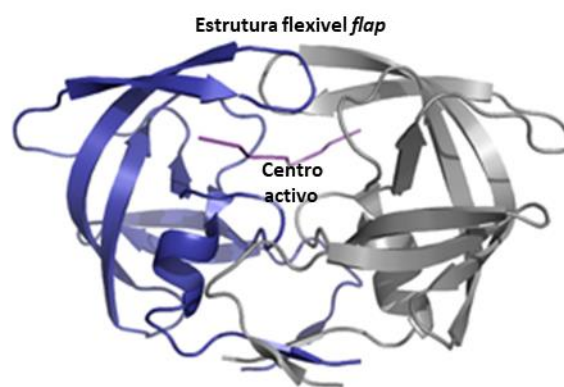


Figura 7: Representação da estrutura da protease do HIV-1 – adaptado de [https://www.umassmed.edu/contentassets/cdb4c5f734e64944bb8c017128262f71/hiv1pr\\_general.png](https://www.umassmed.edu/contentassets/cdb4c5f734e64944bb8c017128262f71/hiv1pr_general.png).

potenciais alvos terapêuticos de inibidores da PR, dada a sua importância para o correcto funcionamento da enzima (104).

### 1.8.2 Transcriptase reversa

A RT do HIV-1 existe na forma de um heterodímero assimétrico constituído por duas subunidades, p66 e p51 (Figura 8), possuindo duas actividades enzimáticas distintas que em conjunto permitem a síntese do DNA proviral a partir do genoma de RNA: DNA polimerase, com a capacidade de utilizar como matriz, tanto moléculas de RNA, como moléculas de DNA; e RNase (ribonuclease) H, que degrada especificamente as cadeias de RNA dos híbridos intermediários de DNA/RNA formados durante o processo da retrotranscrição (107–109). As subunidades p66 e p51 partilham a extremidade N-terminal e são constituídas por 560 e 440 aminoácidos respectivamente, sendo que a subunidade p51 deriva da clivagem proteolítica, mediada pela PR, do domínio RNase H C-terminal da subunidade p66 (108,109). Os centros activos de ambas as actividades enzimáticas encontram-se na subunidade p66, sendo que a subunidade p51 possui um papel estrutural importante para a estabilidade e actividade do heterodímero (107,110).

O domínio da DNA polimerase é constituído por quatro subdomínios: *fingers*, *palm* e *thumb*, responsáveis pela polimerização, e *connection*, que permite a ligação ao domínio RNase H (107,110). A configuração do local de ligação aos ácidos nucleicos permite que as moléculas de RNA viral estejam em contacto com ambos os domínios catalíticos da enzima, possuindo o subdomínio *thumb* um papel importante no correcto

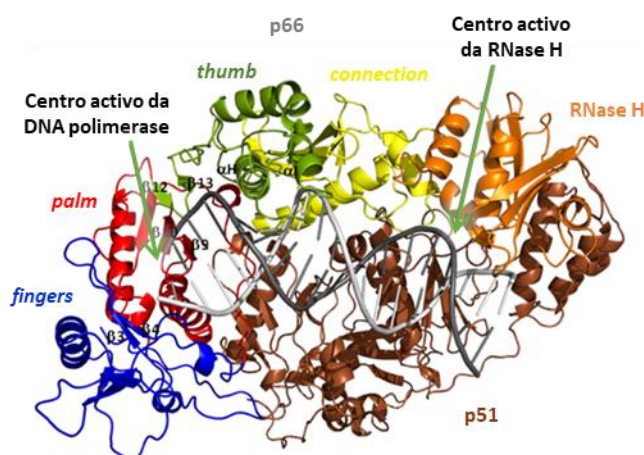


Figura 8: Representação da estrutura da transcriptase reversa do HIV-1 – adaptado de [http://www.mdpi.com/viruses/viruses-02-00606/article\\_deploy/html/images/viruses-02-00606-g001.png](http://www.mdpi.com/viruses/viruses-02-00606/article_deploy/html/images/viruses-02-00606-g001.png).

posicionamento da molécula durante a polimerização (107). O centro activo do domínio da DNA polimerase é composto por uma tríade catalítica (D110, D185 e D186) no subdomínio *palm* que interage com dois iões divalentes ( $Mg^{2+}$ , utilizado *in vivo*, e  $Mn^{2+}$ ) necessários à actividade catalítica. Os resíduos D185 e D186 fazem parte do motivo

YXDD (onde, no caso do HIV-1, X é M), altamente conservado nas RTs retrovirais e fundamental para a manutenção da respectiva actividade (107). Existem outros domínios conservados, como é o caso do DNA *primer grip* que permite o correcto posicionamento da extremidade 3'-OH do *primer* tRNA<sub>Lys3</sub> no centro activo da polimerase para o ataque nucleofílico ao dNTP (desoxirribonucleotídeo trifosfatado) a incorporar na cadeia nascente (107,111–113). Mutações nestes e noutros domínios conservados resultam em alterações na conformação do centro activo da polimerase, conduzindo à sua inactividade, ou a uma diminuição da sua actividade, comprometendo o ciclo replicativo viral (111–114).

O domínio da RNase H encontra-se na extremidade C-terminal da subunidade p66 e possui duas estruturas principais bem conservadas: o centro activo, composto por quatro aminoácidos (D442, E478, D498 e D549), e o RNase H *primer grip* (composto pela rede de aminoácidos T473, I505, K476, Q475, Y501, Q448, N474 e R500) que interage com as cadeias de DNA e RNA próximas do centro activo (115). Estudos *in vivo* e *in vitro* demonstraram que mutações nestes resíduos podem influenciar a actividade e especificidade da RNase H, assim como da DNA polimerase, resultando na perda de infecciosidade dos viriões (115).

Diversos estudos revelam a importância de determinadas sequências e resíduos de aminoácidos na estabilidade do heterodímero p51/p66 e na interacção entre as duas subunidades da RT e ainda que mutações nestes locais afectam a sua função, sugerindo que a actividade da enzima depende da manutenção da sua forma heterodimérica, o que faz desses locais potenciais alvos terapêuticos dos inibidores da RT (108,114,116–119).

### 1.9 Terapia anti-retroviral

O tratamento da infecção pelo HIV-1 é realizado com recurso à utilização de anti-retrovirais (ARVs), nomeadamente, inibidores de entrada (antagonista de co-receptor e inibidor de fusão), da RT [NRTIs (do inglês, *nucleoside reverse transcriptase inhibitors*) e NNRTIs (do inglês, *nonnucleoside reverse transcriptase inhibitors*)], da IN (INSTIs, do inglês *integrase strand transfer inhibitors*) e da PR (PIs, do inglês *protease inhibitors*), que, de um modo geral, actuam por forma a inibir a replicação viral impedindo a infecção de novas células e a progressão da doença (2,45,120). Uma vez que não é possível eliminar ou curar a infecção, devido, em parte, à ocorrência de latência do vírus em

células infectadas, a terapia anti-retroviral tem como objectivo principal a diminuição da carga viral para níveis indetectáveis (<50 cópias/mL de plasma – embora testes mais recentes permitam detectar até 20 cópias/mL de plasma, pelo que os valores de referência poderão alterar-se mediante a sua utilização na monitorização da doença), mas também a manutenção da contagem de linfócitos T CD4<sup>+</sup> em níveis normais (acima das 600 células/mm<sup>3</sup> de sangue) (45,120,121).

Inicialmente, os ARVs eram administrados em regimes de monoterapia, tendo-se introduzido em meados dos anos 90 do século passado o conceito de terapia combinada, após o desenvolvimento de inibidores da PR e da RT, revolucionando o tratamento da infecção (45,120). A utilização da terapia anti-retroviral de alta eficácia (HAART, do inglês *highly active antiretroviral therapy*), que consiste na administração de uma combinação de, pelo menos, três ARVs com, pelo menos, dois alvos terapêuticos distintos, de forma a evitar o desenvolvimento de resistências, permitiu a redução da morbilidade e mortalidade associadas à infecção com HIV e à sida, transformando esta síndrome numa patologia crónica tratável e passível de ser controlada, em vez de permanecer uma doença progressiva inevitavelmente fatal por falta de um tratamento eficaz (2,45,52,120). Uma das questões mais relevantes e amplamente debatidas em relação à terapia anti-retroviral refere-se ao momento do seu início. Actualmente, considera-se que a terapêutica deverá ser iniciada o mais cedo possível, uma vez que se verificam benefícios clínicos para os infectados e vantagens na prevenção da infecção, sendo recomendado pela OMS a administração da terapia anti-retroviral a todos os indivíduos com um diagnóstico de HIV confirmado, independentemente das contagens de linfócitos T CD4<sup>+</sup> e estado clínico, dando prioridade aos indivíduos com doença avançada (122,123). Algumas agências internacionais e associações governamentais, como é o caso do *U.S. Department of Health and Human Services*, consideram ainda que a terapia anti-retroviral deve ser iniciada logo após o diagnóstico da infecção e reforçam também a importância do seu diagnóstico precoce (124).

### 1.10 Resistência associada à terapia anti-retroviral

A emergência de resistências associadas ao uso de ARVs no tratamento da infecção com HIV-1 é uma forte limitação ao sucesso da HAART. Tal como já foi referido, a elevada diversidade genética do HIV-1 é um dos factores potenciadores da

aquisição de mutações de resistência (e consequentes diferenças na resposta terapêutica) devido à emergência de estirpes resistentes no conjunto da *quasispecies* viral. Do mesmo modo, a fraca adesão terapêutica, relacionada com a toxicidade dos ARVs, é outro factor que, a par das interacções medicamentosas e factores intrínsecos do hospedeiro, relacionados, entre outros aspectos, com o metabolismo dos ARVs, pode resultar no desenvolvimento de estirpes resistentes, em consequência de uma supressão viral incompleta, decorrente de alterações dos tempos de semi-vida dos ARVs e das suas concentrações no organismo do hospedeiro (125,126). Além disso, a selecção de estirpes virais resistentes aos ARVs está ainda dependente de aspectos como a barreira genética para aquisição de resistência (definida como o número mínimo de mutações necessário para superar a pressão selectiva do fármaco), característica de cada classe de ARVs, e a *fitness* viral das estirpes mutadas (definida, genericamente, como a capacidade de um vírus se adaptar e replicar sob pressão selectiva (127)). Somente estirpes com mutações de resistência que permitam a replicação sob administração de um determinado regime terapêutico poderão tornar-se dominantes no conjunto da *quasispecies* viral, o que poderá resultar em falência terapêutica, com perda do controlo da infecção.

Um dos mais graves problemas de saúde pública relativos ao HIV-1 é a transmissão directa de vírus resistentes, pois esta limita grandemente as opções terapêuticas aplicáveis, uma vez que, apesar dos indivíduos infectados com esses vírus poderem nunca ter sido submetidos a terapia anti-retroviral, são portadores de estirpes virais com mutações de resistência aos ARVs desde o momento da infecção (128–130). Na ausência de mutações de resistência, os ARVs existentes parecem ser igualmente eficazes no tratamento da infecção com os diferentes subtipos do HIV-1. No entanto, diversos estudos relatam a associação de determinadas mutações com subtipos específicos do vírus, o que resulta em diferenças nos padrões de resistência entre subtipos (126,128,130). Outros estudos reportam a importância clínica de variantes minoritárias resistentes, principalmente no caso de ARVs com uma baixa barreira genética (como os NNRTIs), pois uma única mutação pode determinar o sucesso, ou insucesso, da terapêutica aplicada (129–131). Todos estes factores, incluindo ainda a decisão sobre o momento em que é iniciada a terapêutica e o próprio regime terapêutico que é aplicado (130), dificultam o combate contra a infecção, sendo a elevada diversidade genética do vírus um ponto fulcral nesta questão.

### 1.11 Monitorização de resistência aos anti-retrovirais

A monitorização de resistências aos ARVs pode ser realizada com recurso a métodos fenotípicos ou genotípicos, ambos consistindo em testes de susceptibilidade aos ARVs. Os ensaios fenotípicos permitem medir directamente o grau de susceptibilidade aos ARVs dos vírus em estudo com recurso a culturas celulares, sendo, em algumas técnicas, utilizados vírus recombinantes. Estes, obtidos através da clonagem molecular de segmentos (amplificados por técnicas de PCR) de um gene do vírus selvagem em estudo (PR, RT, IN ou *env*) em vectores de clonagem baseados no HIV-1, defectivos para o gene utilizado, são introduzidos em células permissivas para que ocorra replicação viral *in vitro*, sujeitando essas células a diferentes concentrações dos fármacos de interesse. O grau de susceptibilidade dos vírus recombinantes é então estimado através do cálculo da concentração de fármaco necessária para inibir a replicação viral em 50 (IC50) ou 90% (IC90), medindo-se, desse modo, o efeito líquido da combinação da grande maioria das mutações para o gene em estudo (126,128,132,133).

Os ensaios genotípicos permitem determinar perfis genéticos de resistência dos vírus analisados, baseando-se na sequenciação nucleotídica dos genes virais (ou segmentos destes) cujos produtos constituem alvos terapêuticos dos fármacos de interesse. Estes métodos permitem a detecção directa de mutações de resistência e de outros polimorfismos genéticos, identificados de forma semi-automatizada recorrendo a ferramentas de bioinformática que permitem a comparação do perfil genético obtido, após sequenciação, com o perfil genético da estirpe selvagem de referência. Este tipo de ensaios é utilizado mais frequentemente, em detrimento dos testes fenotípicos, por serem menos dispendiosos, morosos e laboriosos (128,132,134). No entanto, a interpretação dos resultados pode constituir uma tarefa complexa devido à existência de inúmeras mutações de resistência aos ARVs que originam diversos perfis genotípicos (apresentando diferentes níveis de resistência). Por outro lado, ao possuírem uma baixa sensibilidade na detecção de mutações nas estirpes minoritárias, levando a que uma parte das mutações não seja detectada, incluindo as que se encontram “arquivadas” no genoma das células infectadas ou as que ainda não foram caracterizadas, encontram-se dependentes da informação presente nas bases de dados existentes que, por sua vez, necessita de ser constantemente actualizada por equipas de especialistas (126,128,132).

Actualmente, existem disponíveis para consulta e utilização diversas bases de dados internacionais (*e.g. Stanford University HIV Drug Resistance Database* e *Los Alamos HIV Sequence Database*) e ferramentas de bioinformática para interpretação de dados genotípicos especificamente direccionadas para o HIV-1 (*e.g. HIVdb Program – Genotypic Resistance Interpretation Algorithm*) e ainda sistemas baseados na opinião de especialistas que permitem uma abordagem fenotípica virtual (*e.g. EuResist Prediction Engine* da *EuResist Network*). Contudo, os resultados nem sempre são totalmente coincidentes quando aplicados às mesmas mutações, uma vez que os diferentes sistemas de interpretação não são necessariamente baseados nas mesmas premissas, sendo utilizados algoritmos distintos (135).

### 1.12 Inibidores da protease e da transcriptase reversa

Os PIs são compostos que mimetizam o domínio de ligação dos substratos da PR (poliproteínas virais) ao centro activo da enzima, ao qual se ligam (não sendo, no entanto, clivados pela PR devido à presença de uma ligação hidroxietileno), actuando como inibidores competitivos das reacções proteolíticas necessárias para a maturação dos viriões, o que resulta na produção de partículas virais imaturas e não infecciosas (136–138). Actualmente, existem nove PIs aprovados pela FDA (do inglês, *U.S. Food and Drug Administration*) em utilização para o tratamento da infecção pelo HIV (139), um dos quais, o ritonavir, é utilizado como potenciador farmacocinético (por se tratar de um potente inibidor do citocromo P450 permite aumentar substancialmente as concentrações de outros PIs no plasma, quando co-administrado, potenciando o seu efeito terapêutico) (138,140,141). Propostas para novos PIs, que interferem com a dimerização da enzima, ou com a flexibilidade da região *flap*, têm surgido, em consequência da constante procura de novos alvos terapêuticos e do desenvolvimento de novos fármacos (105,142).

Os NRTIs interagem com o centro activo do domínio da polimerase da RT, actuando como terminadores de cadeia na síntese do DNA proviral, quando incorporados na cadeia nascente, devido à ausência de um resíduo 3'-OH na sua estrutura molecular, o que impede a formação da ligação fosfodiéster necessária para a continuação da reacção de polimerização (107,137). Constituem inibidores competitivos e são produzidos sob a forma de pró-fármacos, sendo necessária a sua fosforilação (adição de três grupos fosfato) por enzimas celulares para a sua conversão na forma farmacologicamente activa

(107,136,137). Presentemente, existem sete NRTIs aprovados pela FDA em utilização para o tratamento da infecção pelo HIV (139), constituindo um deles, o tenofovir, uma excepção, uma vez que se trata de um análogo de nucleótido e não de nucleósido, estando, no entanto, inserido nesta classe de ARVs pelo facto de a sua estrutura molecular e mecanismo de acção ser semelhante ao dos NRTIs – possui um grupo fosfato que não pode ser removido por hidrolases, necessitando somente da adição de dois grupos fosfato para a sua conversão na forma activa (107,136,137).

Os NNRTIs são moléculas de natureza hidrofóbica com diferentes estruturas moleculares que, ao contrário dos NRTIs, não necessitam de ser metabolizados a nível celular para actuarem, tratando-se de inibidores não competitivos específicos da RT do HIV-1, não afectando a RT do HIV-2 (109,137). Actuam através da sua ligação a um centro alostérico próximo do centro activo do domínio da polimerase da RT (associado ao subdomínio *palm* da subunidade p66), o que resulta na alteração da sua conformação afectando o funcionamento da enzima (107,109,136,137). Actualmente, existem seis NNRTIs aprovados pela FDA para o tratamento do HIV, um dos quais, o doravirine, foi aprovado em Agosto de 2018 e outro, o delavirdine, já não é recomendado para o tratamento da infecção pelo HIV-1, segundo as directrizes do *U.S. Department of Health and Human Services*, estando apenas cinco em utilização (124,139,143,144).

### 1.12.1 Mutações de resistência

As mutações de resistência aos diferentes ARVs podem ser classificadas como principais (*major*) ou acessórias (*minor*), sendo que as primeiras são capazes de reduzir por si só a susceptibilidade aos ARVs, enquanto as últimas apenas reduzem a susceptibilidade aos ARVs (ou aumentam a *fitness* replicativa) quando em combinação com uma mutação principal (134). No entanto, uma mutação considerada principal para um determinado fármaco pode ser considerada acessória para um outro, pelo que estas classificações não estão estritamente definidas para cada mutação e são, inclusive, influenciadas por outros factores, como a frequência da mutação em indivíduos com falência virológica ou do polimorfismo em indivíduos não tratados e a sua localização na estrutura da proteína alvo (134,145).

A elevada resistência aos PIs resulta de uma acumulação gradual de mutações principais e acessórias cuja combinação permite à PR exercer a sua função enzimática na

presença do inibidor, contrariando a sua actividade e aumentando a resistência ao mesmo, sendo esta a razão pela qual são designados de inibidores de elevada barreira genética (132,134,146). As primeiras mutações *major* a ser adquiridas são, na sua maioria, substituições na própria região codificante da PR, em particular no local de ligação ao substrato (reduzindo a afinidade entre o inibidor e a enzima mutante). Sendo, contudo, normalmente insuficientes para inibir totalmente a actuação do fármaco, são depois seleccionadas mutações acessórias que, potenciando o grau de resistência, permitem reparar a função enzimática e aumentar a *fitness* replicativa (*e.g.* alterações na estrutura do centro activo, na região flexível *flap* e da estabilidade do dímero) e mutações que introduzem modificações nos locais de clivagem principais das poliproteínas precursoras virais com vista a restaurar o processamento proteico e aumentar a produção da própria PR, tratando-se, deste modo, a aquisição de resistência aos PIs de um processo “passo-a-passo” (120,132,134,146–148). O facto de a estrutura química da maioria dos PIs ser muito semelhante leva a que algumas mutações principais originem perfis fenotípicos de múltipla resistência, o que constitui um dos problemas mais relevantes da terapêutica com PIs (132,134). Na Figura 9 encontram-se listadas as mutações na região codificante da

|                              |                             |              |              |              |              |              |              |              |              |              |              |              |              |              |              |              |              |              |              |              |              |              |              |              |
|------------------------------|-----------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Atazanavir<br>+/- ritonavir* | L<br>10<br>F<br>V<br>C      | G<br>16<br>E | K<br>20<br>R | L<br>24<br>I | V<br>32<br>I | L<br>33<br>I | E<br>34<br>Q | M<br>36<br>I | M<br>46<br>I | G<br>48<br>V | I<br>50<br>L | F<br>53<br>L | I<br>54<br>Y | D<br>60<br>E | I<br>62<br>V | I<br>64<br>M | A<br>71<br>V | G<br>73<br>C | V<br>82<br>A | I<br>84<br>V | I<br>85<br>V | N<br>88<br>S | L<br>90<br>M | I<br>93<br>L |
| Darunavir/<br>ritonavir*     | V<br>11<br>I                |              |              |              | V<br>32<br>I | L<br>33<br>F |              |              | I<br>47<br>V |              | I<br>50<br>V | I<br>54<br>M |              |              |              |              | T<br>74<br>P | L<br>76<br>V |              | I<br>84<br>V |              |              | L<br>89<br>V |              |
| Fosamprenavir/<br>ritonavir  | L<br>10<br>F<br>I<br>R<br>V |              |              |              | V<br>32<br>I |              |              |              | M<br>46<br>I | I<br>47<br>V | I<br>50<br>V | I<br>54<br>L |              |              |              |              | G<br>73<br>S | L<br>76<br>V | V<br>82<br>A | I<br>84<br>V |              |              | L<br>90<br>M |              |
| Indinavir/<br>ritonavir**    | L<br>10<br>I<br>R<br>V      | K<br>20<br>M | L<br>24<br>I | V<br>32<br>I |              |              |              | M<br>36<br>I | M<br>46<br>I |              |              | I<br>54<br>V |              |              |              |              | A<br>71<br>V | G<br>73<br>S | L<br>76<br>V | V<br>77<br>I | I<br>82<br>A | I<br>84<br>V | L<br>90<br>M |              |
| Lopinavir/<br>ritonavir*     | L<br>10<br>F<br>I<br>R<br>V | K<br>20<br>M | L<br>24<br>I | V<br>32<br>I | L<br>33<br>F |              |              | M<br>36<br>I | I<br>46<br>V | I<br>47<br>L | I<br>50<br>V | F<br>53<br>L | I<br>54<br>V |              |              | L<br>63<br>P | A<br>71<br>V | G<br>73<br>S | L<br>76<br>V | V<br>82<br>A | I<br>84<br>V |              | L<br>90<br>M |              |
| Nelfinavir***                | L<br>10<br>F<br>I           |              |              |              | D<br>30<br>N |              |              | M<br>36<br>I | M<br>46<br>I |              |              |              |              |              |              |              | A<br>71<br>V |              | V<br>77<br>I | V<br>82<br>A | I<br>84<br>V | N<br>88<br>D | L<br>90<br>M |              |
| Saquinavir/<br>ritonavir*    | L<br>10<br>I<br>R<br>V      |              | L<br>24<br>I |              |              |              |              |              | G<br>48<br>V |              |              | I<br>54<br>V |              | I<br>62<br>V |              |              | A<br>71<br>V | G<br>73<br>S | V<br>77<br>I | V<br>82<br>A | I<br>84<br>V |              | L<br>90<br>M |              |
| Tipranavir/<br>ritonavir     | L<br>10<br>V                |              |              |              | L<br>33<br>F |              |              | M<br>36<br>I | K<br>43<br>T | M<br>46<br>L | I<br>47<br>V | I<br>54<br>A | Q<br>58<br>E |              |              | H<br>69<br>K |              | T<br>74<br>P |              | V<br>82<br>T | N<br>83<br>D | I<br>84<br>V | L<br>89<br>I |              |

Figura 9: Conjunto das mutações associadas à resistência aos PIs (listadas pela IAS-USA), estando destacadas a negrito as mutações com maior relevância clínica (149). Legenda: para cada fármaco, ou combinação, em cima é indicado o aminoácido presente no vírus selvagem (susceptível), o número indica a respectiva posição na sequência e em baixo estão indicados o(s) aminoácido(s) encontrado(s) nos vírus mutados associado(s) a resistência.

PR associadas à resistência aos PIs, segundo a IAS-USA (do inglês, *International Antiviral Society-USA*) (149).

A resistência aos inibidores da RT (RTIs, do inglês *reverse transcriptase inhibitors*) está correlacionada com mutações que ocorrem na região codificante da própria RT, existindo, no caso dos NRTIs, dois mecanismos associados à resistência (150). Um deles consiste na aquisição de mutações, ditas discriminatórias, que permitem à RT distinguir entre os dNTPs celulares e os seus análogos, aumentando a sua selectividade e impedindo a incorporação dos segundos na cadeia nascente do DNA (128,132,150). O outro consiste na aquisição de mutações que afectam a actividade fosforolítica da RT e conferem à enzima a capacidade de excisar os NRTIs incorporados na cadeia nascente de DNA proviral a partir da extremidade 3', revertendo a terminação da cadeia e permitindo a continuação da sua polimerização num processo designado de *primer unblocking*. Estas designam-se mutações associadas aos análogos de timidina (TAM, do inglês *thymidine analogue-associated mutations*) por estarem associadas à utilização de inibidores análogos a esse dNTP (120,128,132). Fenómenos de resistência

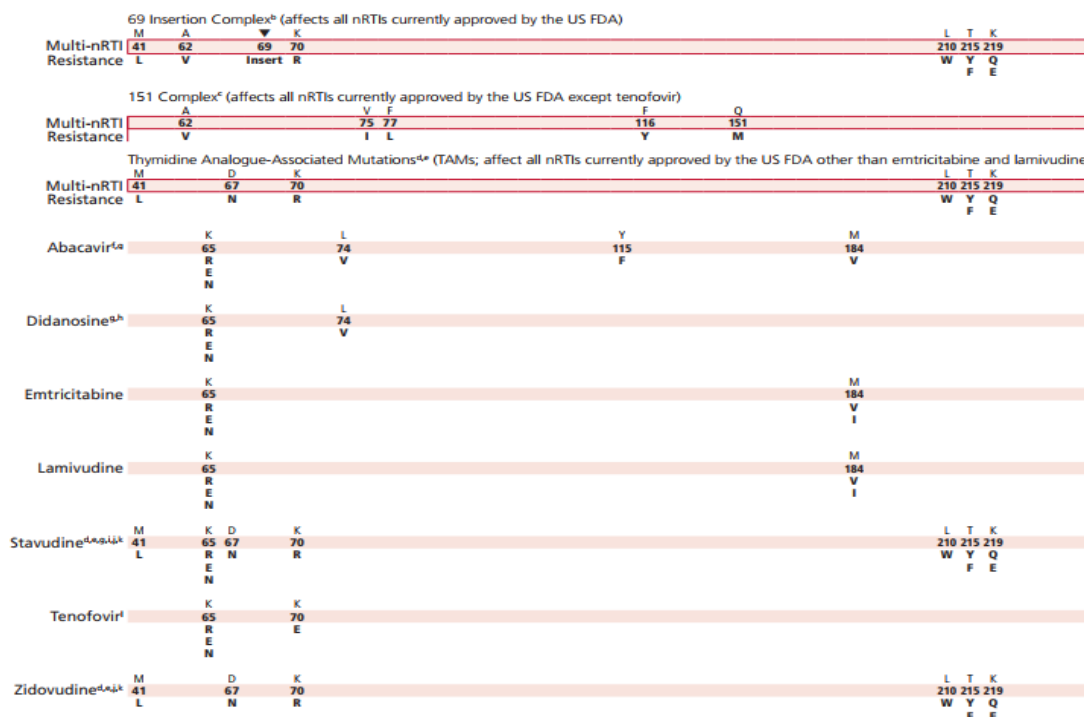
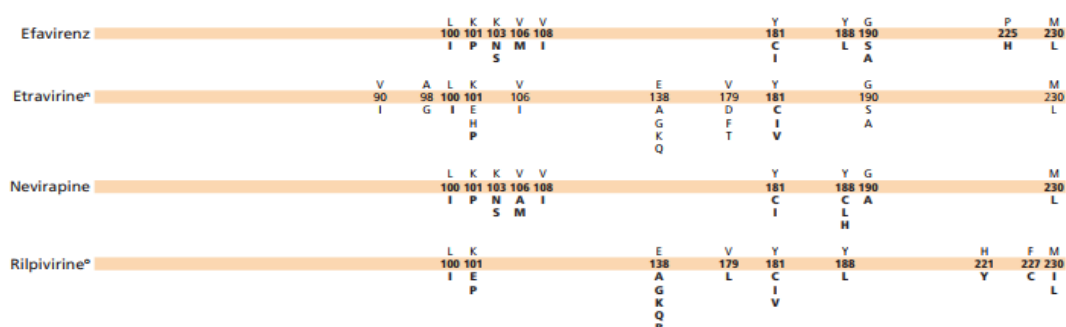


Figura 10: Conjunto das mutações associadas a resistência aos NRTIs (listadas pela IAS-USA), estando destacadas a negrito as mutações com maior relevância clínica (149). Legenda: para cada fármaco, em cima é indicado o aminoácido presente no vírus selvagem (susceptível), o número indica a respectiva posição na sequência, e em baixo estão indicados o(s) aminoácido(s) encontrado(s) nos vírus mutados associado(s) a resistência.

cruzada aos NRTIs (aos quais estão também associadas as TAM) resultam em dificuldades na aplicação de uma terapêutica eficaz, particularmente em indivíduos que já fizeram tratamento com este tipo de ARVs (132,150). Na Figura 10 encontram-se listadas as mutações na região codificante da RT associadas à resistência aos NRTIs, segundo a IAS-USA (149).

No caso dos NNRTIs, a resistência resulta da aquisição de mutações no local de ligação do fármaco (ou resíduos adjacentes) que interferem com as interações entre este e o seu local de ligação na RT, reduzindo a afinidade da ligação e interferindo, desse modo, com a capacidade de interacção com os resíduos do centro alostérico (147,150). Estes RTIs possuem uma baixa barreira genética, pelo que apenas uma ou duas mutações podem resultar em elevados níveis de resistência, com um impacto mínimo na *fitness* replicativa do vírus (120,128,147). Na classe dos NNRTIs verificam-se níveis elevados de resistência cruzada (150) que resultam do facto de a maioria das mutações reduzirem a susceptibilidade a múltiplos NNRTIs e de a baixa barreira genética permitir a emergência de linhagens independentes multi-resistentes, mesmo que não seja possível detectar todas as mutações por sequenciação nucleotídica de base populacional (128,147). Por outro lado, foi verificada a associação entre algumas mutações de resistência aos NRTIs (em particular TAMs) e hipersusceptibilidade aos NNRTIs, sendo que indivíduos com este tipo de mutações apresentam uma melhor resposta à terapêutica com estes últimos (126,132). Na Figura 11 encontram-se listadas as mutações na região codificante da RT associadas à resistência aos NNRTIs, segundo a IAS-USA (149).



**Figura 11:** Conjunto das mutações associadas a resistência aos NNRTIs (listadas pela IAS-USA), estando destacadas a negrito as mutações com maior relevância clínica (149). Legenda: para cada fármaco, em cima é indicado o aminoácido presente no vírus selvagem (susceptível), o número indica a respectiva posição na sequência, e em baixo estão indicados o(s) aminoácido(s) encontrado(s) nos vírus mutados associado(s) a resistência.

### 1.13 HIV-1 em Angola

Angola é um país da costa ocidental africana que faz fronteira com a República do Congo, a RDC, a Zâmbia e a Namíbia, possuindo em 2017 uma população total estimada de 29,8 milhões de pessoas (151). Estudos recentes (152) relevam que Angola terá desempenhado um papel importante na disseminação inicial da pandemia de HIV-1 juntamente com a RDC, onde a mesma teve origem no início do séc. XX. O elevado número de imigrantes angolanos a viver em Kinshasa e os movimentos migratórios dos mesmos (e de outros indivíduos), entre Angola e a RDC (entre outros países), durante os anos 40 do século passado, levaram à introdução do vírus no país, a partir da sua origem, ainda numa fase inicial da pandemia, justificando a elevada diversidade genética encontrada em Angola, por sua vez comparável àquela que é verificada na RDC e outros países vizinhos (152). Por outro lado, a Guerra Colonial nos anos 60 e 70, e a guerra civil que se seguiu (1975-2002), terão também contribuído para o aumento da diversidade genética, em consequência da deslocação temporária de parte da população angolana para os países vizinhos durante a guerra, e ainda para o padrão epidemiológico da epidemia em Portugal, em consequência da introdução do vírus no país resultante dos movimentos populacionais entre os dois países antes, durante e após a guerra (153,154). Em 2004, foi introduzido em Angola o acesso universal à terapia anti-retroviral, facto com impacto considerável na epidemia local, em particular na melhoria da qualidade de vida das pessoas a viver com HIV (155).

#### 1.13.1 Dados epidemiológicos gerais

Segundo dados recentes da UNAIDS (39,156), em 2017 foram registadas em Angola 27 000 [20 000–37 000] novas infecções com HIV e 13 000 [9500–17 000] mortes relacionadas com sida, verificando-se para indivíduos com idades entre 15 e 49 anos uma taxa de prevalência de 1,9% [1,6–2,3%] e existindo cerca de 310 000 [260 000–360 000] pessoas a viver com HIV, na sua maioria mulheres (um total de 190 000 [160 000–220 000], apresentando uma taxa de prevalência de 2,6% [2,1–3,0%] vs. 1,2% [1,0–1,5%] verificada nos homens da mesma faixa etária). Do total de pessoas que viviam com HIV em Angola em 2017 apenas cerca de 26% [22–30%] estavam em tratamento e, em 2016, do total de 280 000 [240 000–330 000] pessoas a viver com HIV no país, apenas cerca de 40% [33–47%] tinham conhecimento da sua condição e a percentagem de indivíduos

virologicamente suprimidos era de 16% [13–19%], não estando disponíveis dados referentes a 2017 sobre estes dois aspectos. Estima-se também que o número de novas infecções em crianças em 2017 foi de 5500 [3700–7100], sendo que 14% [10–17%] destas tinham acesso à terapia anti-retroviral, assim como 34% [24–42%] das mulheres grávidas a viver com HIV, estimando-se que foram evitadas 1200 [<1000–1500] novas infecções graças à prevenção da transmissão vertical. A percentagem de indivíduos com idades entre 15 e 49 anos com conhecimento acerca da prevenção da infecção em 2015 era de somente cerca de 30%. No entanto, o uso de preservativo em relações sexuais de alto risco em 2016 era de aproximadamente 63,3% e 32,1% em homens e mulheres da mesma faixa etária, respectivamente. Por outro lado, a informação sobre as populações mais vulneráveis é bastante escassa, conhecendo-se apenas alguns dados sobre o conhecimento da sua condição, prevalência da infecção, cobertura da terapia anti-retroviral e uso de preservativo de três populações (profissionais do sexo, homens que fazem sexo com homens e indivíduos transgénero). Já o investimento monetário realizado em 2017 relacionado com o HIV em Angola foi de cerca de 20 milhões de dólares americanos, consistindo na sua maioria em doações internacionais. Apesar disso, o número de novas infecções com HIV e mortes relacionadas com sida aumentou em 1 e 29% respectivamente, desde 2010 e, em 2016, 31% dos indivíduos com idades entre 15 e 49 anos reportaram atitudes discriminatórias para com pessoas a viver com HIV.

### **1.13.2 Epidemiologia molecular**

#### **1.13.2.1 Genótipos circulantes**

Diversos estudos permitiram ajudar a determinar o perfil epidemiológico molecular da epidemia de HIV-1 em Angola, que se caracteriza pela circulação da maioria dos subtipos do grupo M, verificando-se uma elevada diversidade genética, à semelhança do que foi observado em países vizinhos, sugerindo que o vírus circula no país há um longo período de tempo (154,155,157–162). No entanto, os seus resultados poderão não representar, com exactidão, a situação epidemiológica do país, considerando a cobertura geográfica altamente fragmentada, havendo até províncias não representadas, e a origem conveniente da maioria das amostras caracterizadas.

De um modo geral, os diferentes estudos revelam um predomínio dos subtipos não-B, sendo que os mais recentes (161,162), cujas amostras eram provenientes de várias regiões de Angola, revelam que os subtipos C e F1 são os mais frequentes (16–36,9% e 14–20%, respectivamente), mas também a existência de uma elevada proporção (superior a 40%) de formas recombinantes. Observou-se uma mudança em relação ao que foi reportado em estudos anteriores, nos quais o subtipo A (e respectivos sub-subtipos) era apontado como o mais frequente (de 13,8% para 5–10,7%), tendo Bártoleto *et al.* (162) avançado com a hipótese de que esta mudança estará relacionada com o aumento da prevalência do subtipo C no sul da RDC e na Zâmbia (161,162).

A CRF mais frequentemente descrita é a CRF02\_AG (correspondendo a 14–16,9% das formas recombinantes encontradas), mas a maioria das formas recombinantes anteriormente reportadas são URFs (42–83,1%). A sua frequência aumentou significativamente entre 2001 e 2009, resultante de eventos de recombinação genética entre diferentes subtipos, CRFs e sequências não classificadas (designadas por “U”, do inglês *unclassified*), sendo que estas últimas foram detectadas em cerca de metade das formas recombinantes e constituem uma pequena porção do total das sequências obtidas em ambos os estudos (161,162). As URFs mais frequentes (10,8–21%) continham genomas mosaico com sequências do subtipo H e sequências U, designando-se de U/H. Tendo sido encontradas em indivíduos não relacionados epidemiologicamente e formando, na análise filogenética, grupos monofiléticos suportados por valores de *bootstrap* significativos, aparentam ter origem em Luanda e estar já estabelecidas no país, podendo eventualmente representar uma nova CRF.

Em termos da distribuição das variantes encontradas, verifica-se um predomínio do subtipo B (12) e do subtipo C na região Sul (45,5%), enquanto na região Centro existe um predomínio das diferentes CRFs e URFs (42%), em particular das formas recombinantes U/H (16%), reflectindo o movimento das populações dentro e fora do país ao longo do tempo. A análise das redes de transmissão da infecção revelou um elevado número de grupos de transmissão, contendo 35,7% das sequências analisadas, sendo que a maioria continha somente sequências provenientes de Luanda, embora alguns incluíssem também sequências com origem em diferentes regiões do país (154,158,160,161,163), reflectindo a complexidade da epidemia.



**Figura 12: Mapa de Angola e respectivas províncias, ficando a região Norte acima da linha vermelha e a região Sul abaixo da linha azul – adaptado de <http://www.africa-turismo.com/imagens/mapa-angola.jpg>.**

### 1.13.2.2 Mutações de resistência aos inibidores da protease e da transcriptase reversa

No que se refere a mutações de resistência, nos diversos estudos realizados (153,155,160,162) não foram descritas mutações principais associadas à resistência aos PIs, tendo, no entanto, sido detectadas diversas mutações acessórias e polimorfismos genéticos em indivíduos não tratados, sugerindo que algumas estirpes de origem angolana possuem uma baixa barreira genética para aquisição de resistência a alguns PIs, sendo os vírus do subtipo F os mais polimórficos (162). Por exemplo, as mutações L10I e L10V, associadas a resistência à maioria dos PIs na presença de outras mutações, foram encontradas em cerca de 15,7 a 17,7% das estirpes, verificando-se um aumento da sua frequência em relação ao que foi descrito em estudos anteriores (entre 5 e 10%) (160,162).

Por outro lado, a mutação V11I, associada à resistência ao darunavir, foi detectada em estirpes do subtipo A, F e CRF02\_AG em diferentes percentagens (160,162), com valores superiores àqueles que estavam descritos na *Stanford University HIV Drug Resistance Database* à data de publicação dos diferentes estudos. Por fim, o polimorfismo A71V foi encontrado em dois indivíduos infectados com vírus do subtipo D, sendo descrito pela primeira vez neste subtipo viral (162).

Em relação a mutações associadas à resistência aos RTIs, no estudo realizado por Afonso *et al.* (161) foram detectadas mutações de resistência aos NRTIs em 10,5% dos indivíduos não tratados (com as mutações M184V e T215F entre as mais frequentes) e mutações de resistência aos NNRTIs em 14% dos indivíduos não tratados (sendo as mutações K103N e Y181C as mais comuns), verificando-se, na população em estudo, uma frequência de transmissão de mutações de resistência a NRTIs ou NNRTIs elevada (16,3%) e uma frequência de transmissão de mutações de resistência a ambas as classes de RTIs intermédia (8%). No estudo realizado por Bártolo *et al.* (162) foi igualmente detectada a mutação K103N (que confere resistência elevada a nevirapine e efavirenz), entre outras mutações e polimorfismos genéticos, mas apenas num indivíduo não tratado infectado com um vírus do subtipo G, representando uma diminuição da prevalência desta mutação em relação a estudos anteriores (160).

### 1.14 Problema de investigação

Os dados disponíveis indicam que a epidemia do HIV-1 em Angola tem vindo, aparentemente, a aumentar de complexidade, em particular, na diversidade genética das estirpes circulantes, sugerindo um aumento dos fenómenos de co-infecção e/ou superinfecção, e na transmissão de estirpes com mutações de resistência a indivíduos não infectados (resistência primária), tendo o acesso à terapia anti-retroviral um impacto igualmente considerável nessa complexidade. O padrão epidemiológico actualmente observado em Angola, ainda que incompleto, parece claramente influenciado pela rápida propagação de estirpes resistentes, desde a implementação do acesso universal à terapia anti-retroviral, devida, possivelmente, ao uso intermitente dos ARVs. Outros factores determinantes serão a entrada no país de indivíduos infectados oriundos de países onde a HAART está disponível há mais tempo, o uso não monitorizado de ARVs adquiridos no

mercado negro, o uso não reportado de ARVs por questões sociais e dificuldades no acesso ao tratamento em certas regiões do país (155,160–162).

Embora os diversos estudos realizados até hoje sobre a caracterização epidemiológica molecular das estirpes de HIV-1 circulantes em Angola tenham contribuído com informação relevante para a compreensão da epidemia no país, os seus resultados estarão ainda longe de representar com exactidão a verdadeira situação epidemiológica, uma vez que se baseiam num número limitado de indivíduos infectados e em coberturas geográficas restritas, incluindo apenas (ou em elevada percentagem) amostras colhidas na cidade de Luanda ou localizações próximas. A informação sobre a epidemia de HIV-1 noutras regiões do país é, por esse motivo, quase inexistente, o que leva a um potencial enviesamento da informação. Por exemplo, até à data, apenas cinco estudos reportam a inclusão de amostras com origem na província de Benguela nas suas análises, as quais constituem percentagens muito reduzidas do total das amostras analisadas (153,154,160–162). Daqui, surge a necessidade da realização de estudos adicionais com vista a melhor retratar a situação verificada noutras regiões de Angola, contexto em que surge este trabalho e o estudo transversal (*“Sífilis e infecção pelos vírus da imunodeficiência humana e da hepatite B em Benguela, Angola”*, coordenação L. Yefimenko e F. Martins Pereira) no qual o mesmo se insere.

### 1.14.1 Objectivos

Na tentativa de responder ao problema de investigação identificado anteriormente, foram definidos para este projecto dois objectivos principais:

- ❖ Caracterizar filogeneticamente as estirpes do HIV-1 em circulação na província de Benguela, Angola;
- ❖ Estimar o grau de resistência primária aos ARVs inibidores da PR e da RT nas estirpes do HIV-1 em circulação na província de Benguela, Angola.

Estes objectivos principais vão de encontro ao objectivo geral do estudo no qual este trabalho se insere e que consiste na caracterização de infecções sexualmente transmitidas (sífilis, infecção pelo HIV e hepatite B) em indivíduos internados e utentes de consultas externas ou do banco de urgência do Hospital Geral de Benguela (HGB). De

modo concretizar os objectivos principais definidos, foram igualmente estabelecidos os seguintes objectivos específicos:

- ❖ Amplificar por dupla *nested* PCR as regiões codificantes da PR e da RT, no gene *pol*;
- ❖ Obter as sequências nucleotídicas de ambas as cadeias dos amplicões de PR e RT obtidos por métodos de sequenciação de base populacional;
- ❖ Classificar genotipicamente, por métodos de análise filogenética manual, as sequências nucleotídicas codificantes para a PR e RT obtidas;
- ❖ Pesquisar mutações de resistência aos ARVs e outros polimorfismos genéticos associados a resistência por comparação das sequências putativas de aminoácidos obtidas com sequências consenso do subtipo B;
- ❖ Avaliar a associação de polimorfismos genéticos específicos com os subtipos virais encontrados.

Tendo em conta os objectivos aqui definidos, o presente estudo é classificado, segundo Fortin (164), como um estudo descritivo simples.

## 2 Material e métodos

---

### 2.1 Caracterização geral da amostra populacional em estudo

O presente estudo incluiu amostras de sangue de um total de 111 indivíduos seropositivos para o HIV-1 que se encontravam internados ou eram utentes de consultas externas ou do banco de urgência do HGB, na cidade de Benguela, unidade sanitária de referência na região Centro-Oeste de Angola. Para a obtenção de informação sobre as características demográficas e clínicas foi aplicado um questionário estruturado, pré-codificado, de preenchimento voluntário, a todos os indivíduos que assinaram um termo de consentimento livre e esclarecido, estando resumida no Anexo 3 a informação recolhida. Os indivíduos incluídos neste estudo tinham idades compreendidas entre 0 e 78 anos, sendo 42% do género masculino e 58% do género feminino. A totalidade dos indivíduos incluídos nunca foi submetida a terapia anti-retroviral e o diagnóstico da infecção pelo HIV-1 foi realizado à data da colheita das amostras, entre Agosto de 2016 e Janeiro de 2017. É importante referir que inicialmente estavam incluídas neste estudo 202 amostras, não tendo, no entanto, sido possível analisá-las na totalidade devido a atrasos causados pelas dificuldades sentidas na optimização dos protocolos experimentais (descritas em maior detalhe nas secções seguintes). O estudo transversal, mencionado anteriormente, no qual este trabalho se insere obteve um parecer favorável para a sua realização da Comissão Científica e de Ética do HGB.

#### 2.1.1 Colheita das amostras e rastreio da infecção pelo HIV-1

As amostras de sangue foram colhidas, após a obtenção do consentimento informado, por punção venosa, sendo de imediato testadas para a presença de anticorpos anti-HIV com recurso a dois testes imunocromatográficos rápidos. O primeiro teste utilizado foi o *VIH 1/2* (Healgen Scientific, EUA), um teste de terceira geração que permite identificar, em soro, plasma ou sangue total, anticorpos anti-HIV-1 e -2, não distinguindo entre os dois tipos virais (165). O segundo teste foi o *Anti-HIV 1/2 Test* (Türklab, Turquia), que permite a detecção de anticorpos anti-HIV-1 e -2, incluindo HIV-1 do grupo O, produzidos contra os antígenos virais gp120, gp41, gp36, e p24, mas não a distinção entre os tipos virais (166). Para as amostras positivas para, pelo menos, um dos testes realizados, foram armazenadas algumas gotas de sangue total em cartões

*Whatman Indicating FTA<sup>®</sup> Elute Micro Cards* (GE Healthcare, Reino Unido), seguindo as especificações do fabricante. Este tipo de cartões permite manter as amostras à temperatura ambiente (TA) sem que haja degradação significativa, uma vez que após lise das células quando estas entram em contacto com os cartões, as proteínas desnaturam e os ácidos nucleicos são imediatamente estabilizados, ficando protegidos (167). É importante referir que, por questões meramente logísticas, um número reduzido de amostras foi armazenado num outro tipo de cartões, nomeadamente *Whatman Non-Indicating FTA<sup>®</sup> Micro Cards* (GE Healthcare, Reino Unido), os quais possuem uma tecnologia muito semelhante à dos cartões acima referidos. As amostras de sangue positivas foram centrifugadas, em condições padrão, para a separação do soro, sendo este armazenado a -20 °C. Já nas instalações do Instituto de Higiene e Medicina Tropical (IHMT), foi realizado um terceiro teste imunocromatográfico rápido, de modo a confirmar os resultados obtidos em Angola e a discriminar possíveis infecções por HIV-1 e HIV-2. Este terceiro teste foi o *Hexagon HIV-1/2* (Human Diagnostics, Alemanha), também de terceira geração, utilizado para a detecção de anticorpos anti-HIV-1 e -2 em soro, plasma ou sangue total, permitindo a distinção entre os dois tipos virais, através da utilização de duas linhas de teste, cada uma específica para um dos tipos de HIV, e ainda a detecção de anticorpos anti-HIV-1 do grupo O (168,169). Todos os testes foram utilizados segundo as recomendações dos respectivos fabricantes. Esta parte do trabalho foi realizada pela Dr.<sup>a</sup> Lesya Yefimenko, no âmbito do estudo transversal anteriormente referido.

### 2.2 Preparação de discos dos cartões FTA<sup>®</sup>

De modo a ter acesso ao DNA proviral presente nas amostras para a posterior amplificação das regiões codificantes da PR e da RT do HIV-1, foi necessária a preparação de discos a partir da mancha de sangue armazenada nos cartões *FTA<sup>®</sup>*, segundo um protocolo adaptado pelo Grupo de Virologia da Unidade de Microbiologia Médica do IHMT que permite remover os contaminantes e inibidores da PCR da matriz dos cartões permanecendo o DNA ligado à mesma. Após limpeza do perfurador utilizado e tapete de corte com etanol a 70% (v/v), iniciou-se o protocolo com a obtenção de dois círculos de 3 mm de diâmetro a partir do centro da mancha de sangue, que foram colocados num microtubo Eppendorf estéril de 1,5 mL devidamente identificado (entre

cada amostra, cortou-se uma vez o cartão numa zona limpa e, depois de limpar o perfurador com etanol a 70% (v/v), seis vezes em papel *Whatman*<sup>TM</sup> 3MM para evitar contaminações). De seguida, foram adicionados 200 µL de *FTA*<sup>®</sup> *purification reagent* (Whatman, EUA), incubando-se durante 5 min à TA com agitação ocasional e removendo-se o reagente no final da incubação – este passo foi repetido mais três vezes, sendo que na última repetição o tempo de incubação foi de 10 min. Depois foram adicionados 200 µL de tampão TE (Anexo 4), incubando-se durante 5 min à TA com agitação ocasional e removendo-se o tampão no final da incubação – este passo foi repetido mais uma vez. Para secar os discos, os microtubos foram depois colocados, com a tampa aberta, num evaporador (*Concentrator 5301*, Eppendorf, Alemanha), no programa para soluções aquosas, a 30 °C, durante 15 min. Uma vez secos, os discos foram conservados a -20 °C até ao momento da sua utilização.

É importante referir que, antes da aplicação deste protocolo, foi testado o protocolo *Whatman*<sup>TM</sup> *FTA*<sup>®</sup> *Elute*<sup>TM</sup> (Anexo 5), recomendado pelo fabricante para o tipo de cartões utilizados (*Whatman Indicating FTA*<sup>®</sup> *Elute Micro Cards*), cuja tecnologia permite eluir o DNA do cartão permanecendo as proteínas e contaminantes ligados ao mesmo (170,171). No entanto, não foi possível amplificar posteriormente as regiões da PR e da RT a partir do DNA obtido utilizando este protocolo, pelo que se optou pelo seu abandono e utilização do protocolo acima descrito, após ter sido testada a sua eficácia nas condições de realização do presente estudo.

### 2.3 Amplificação por PCR das regiões codificantes da protease e da transcriptase reversa no gene *pol*

#### 2.3.1 Caracterização dos *primers* utilizados

Para amplificar por PCR as regiões codificantes da PR e da RT, a partir do DNA total das diferentes amostras, foram utilizados *primers* previamente adaptados por Videira e Castro, em 2010 (172) a partir do protocolo estabelecido por Steegen *et al.* (173). As adaptações consistiram na introdução de alterações pontuais nas sequências nucleotídicas de alguns *primers*, por forma a aumentar a sua especificidade para os diferentes subtipos do grupo M do HIV-1. As combinações entre os diversos *primers* disponíveis, para cada fase do protocolo de amplificação, foram escolhidas com base nos resultados prévios de Simões, em 2016 (174), de modo a aumentar a eficiência (sensibilidade e especificidade)

da reacção de amplificação. As sequências nucleotídicas dos *primers* e respectivas posições, relativamente à estirpe de referência HXB2 do subtipo B, encontram-se descritas na Tabela 2. Três dos *primers* (PR2, RT137 e RT4) possuem posições degeneradas (Y=T/C, R=G/A, W=A/T, representadas a negrito) correspondentes às posições onde foi verificada maior variabilidade genética em sequências análogas depositadas nas bases de dados. Os *primers* foram sintetizados pela empresa STAB VIDA (Portugal) e a sua preparação, a partir do liofilizado, foi realizada numa câmara de fluxo laminar (*Laminar Flow Cabinet-class 100*, Flow Laboratories, EUA), na ausência de material genético, utilizando pontas de micropipeta com filtro *Biosphere® Filter Tip* (Sarstedt, Alemanha). Inicialmente, foram preparados *stocks* de cada um dos *primers* hidratando-os com água livre de nucleases (*PCR grade*, Roche, Alemanha) para uma concentração final de 100 µM, sendo posteriormente diluídos (1:10) para soluções de trabalho a uma concentração de 10 pmol/µL, utilizando água desionizada ultrapura livre de nucleases *Milli-Q®* (*Reagent Grade Ultrapure Systems*, Merck Millipore, EUA), mantidas a -20 °C.

Tabela 2: *Primers* utilizados na 1.<sup>a</sup> e 2.<sup>a</sup> *nested* PCRs – adaptados de Steegen *et al* (173).

| <i>Primers</i> | Posição<br>nucleotídica<br>em HXB2 | Sequência 5’- 3’                       | Região<br>amplificada | <i>nested</i><br>PCR       |
|----------------|------------------------------------|----------------------------------------|-----------------------|----------------------------|
| Gag2           | 1404 – 1426                        | CAA TGA GGA AGC TGC AGA ATG GG         | PR – RT               | <i>Primers</i><br>externos |
| PR2            | 1822 – 1844                        | GAA ATG ATG ACA GCA TGY CAG GG         |                       |                            |
| RT137          | 3301 – 3325                        | TTC TGT ATR TCA TTG ACA GTC CAG C      |                       |                            |
| PR3            | 2146 – 2166                        | CCA GAG CCA ACA GCC CCA CCA            | PR                    | <i>Primers</i><br>internos |
| PR4            | 2584 – 2605                        | GGG CCA TCC ATT CCT GGC TTT A          |                       |                            |
| RT1            | 2604 – 2632                        | CCA AAA GTT AAA CAA TGG CCA TTG ACA GA | RT                    |                            |
| RT4            | 3231 – 3253                        | TGG AGT TCA TAW CCC ATC CAA AG         |                       |                            |

### 2.3.2 Planificação do protocolo de amplificação

Neste estudo foram planeadas três PCRs sucessivas para a amplificação das regiões codificantes do gene *pol* correspondentes à PR e à RT, nas quais os fragmentos de DNA obtidos na primeira reacção foram utilizados como matriz para realização da segunda (*hemi-nested*), servindo os fragmentos nesta obtidos como matriz para a amplificação das regiões de interesse na terceira reacção (*nested*) (Figura 13). As condições de amplificação foram definidas considerando as temperaturas de desnaturação

e hibridação dos *primers*, o respectivo grau de degeneração (quando aplicável), o tamanho dos fragmentos de DNA a amplificar e as características intrínsecas da Taq DNA polimerase a utilizar nas reacções de PCR. Na primeira e segunda PCRs foi amplificada a extremidade 5' do gene *pol*, compreendendo a região codificante da PR e parte da região codificante da RT, utilizando-se, respectivamente, o par de *primers* Gag2/RT137, para amplificar um fragmento de DNA de 1922 pb, e PR2/RT137 (Figura 13), amplificando um fragmento interno de 1504 pb. Na terceira PCR foi realizada a amplificação, em separado, das regiões codificantes da PR e da RT utilizando, respectivamente, dois pares de *primers*, PR3/PR4 e RT1/RT4 (Figura 13). O *primer* PR3 hibrida na região codificante para a proteína p6 no gene *gag* e o *primer* PR4 já na região codificante para a actividade da RT (p51) no gene *pol*, resultando na amplificação de um fragmento de DNA com 460 pb, contendo a totalidade da região codificante para a PR. Os *primers* RT1 e RT4 hibridam ambos na região codificante para a RT, resultando na amplificação de um fragmento de DNA com 650 pb, contendo a parte 5' da região codificante da respectiva enzima.

Tendo em consideração os objectivos do estudo, foi utilizado nos três passos de PCR o sistema *Supreme NZYTaQ II 2× Green Master Mix* (NZYTech, Portugal) cuja Taq DNA polimerase possui capacidade *hot start*, o que a torna inactiva à TA, sendo a sua actividade restaurada a altas temperaturas, neste caso, por submissão a uma temperatura de 95 °C durante 5 min. Apesar desta polimerase não possuir actividade *proofreading*, foi sintetizada de forma a possuir uma elevada processividade – permitindo a amplificação de sequências mais longas, em menos tempo, com uma maior eficiência e qualidade dos produtos de PCR – e especificidade na amplificação de fragmentos com até 6 kpb, com a particularidade de os fragmentos amplificados possuírem extremidades não coesivas com resíduos protuberantes de adenosina que facilitam a sua posterior clonagem em vectores plasmídeos (175), se necessário. As misturas de reacção foram preparadas numa câmara de fluxo laminar (*Laminar Flow Cabinet-class 100*, Flow Laboratories, EUA), na ausência de material genético, utilizando pontas de micropipeta com filtro *Biosphere® Filter Tip* (Sarstedt, Alemanha), tendo sido incluídos controlos negativos em todas as reacções de amplificação com o objectivo de confirmar a inexistência de contaminações. A adição do disco previamente preparado, contendo o DNA das amostras em estudo, ao respectivo microtubo Eppendorf estéril foi realizada

num compartimento laboratorial distinto do local onde foram preparadas as misturas de reacção, as quais foram mantidas em gelo durante o procedimento.

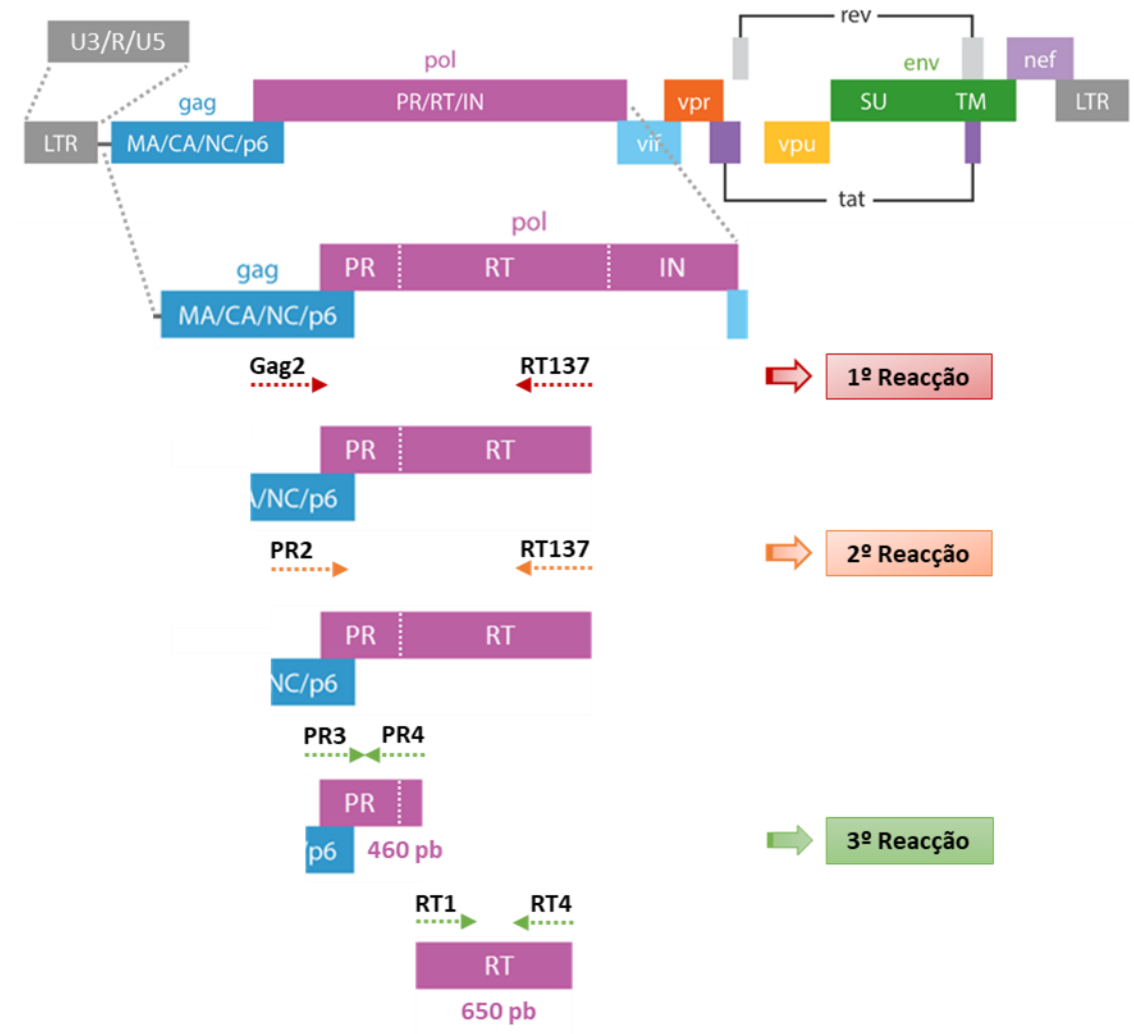


Figura 13: Esquema representativo da amplificação das regiões codificantes da protease e da transcriptase reversa do gene *pol* do HIV-1, por realização de três PCRs consecutivas, para a obtenção de fragmentos de 460 e 650 pb, respectivamente – o esquema não se encontra à escala (adaptado de [https://www.abmgood.com/marketing/knowledge\\_base/img/Lentivirus\\_System/lentivirus\\_genome\\_map.png](https://www.abmgood.com/marketing/knowledge_base/img/Lentivirus_System/lentivirus_genome_map.png)).

É ainda importante salientar que, numa fase inicial do estudo, o protocolo de amplificação consistia na realização de uma *nested* PCR recorrendo directamente ao protocolo de Steegen *et al.* (173). No entanto, após a realização de vários testes comparativos, verificou-se que a amplificação com recurso a três PCRs consecutivas era mais eficiente do que a amplificação por *nested* PCR clássica, pelo que se optou pela execução do protocolo previamente descrito. Além disso, para as duas primeiras reacções de PCR (ver Figura 13) foi inicialmente utilizado o sistema *Supreme NZYProof 2× Green Master Mix* (NZYTech, Portugal), o qual utiliza uma Taq DNA polimerase com

actividade *proofreading* 3'→5' exonuclease que permite evitar a introdução cumulativa de nucleótidos incorrectos e a consequente fixação artificial de mutações. Contudo, após a realização de vários testes, verificou-se que a eficiência do protocolo de amplificação era significativamente maior quando era utilizado o sistema *Supreme NZYTaq II 2× Green Master Mix* (NZYTech, Portugal) em todas as PCRs, pelo que a utilização do sistema com actividade *proofreading* foi abandonada, assumindo-se igualmente a baixa probabilidade de introdução de erros de transcrição durante a fase de polimerização, dado o tamanho relativamente pequeno dos amplicões obtidos. Finalmente, referir que, por questões logísticas e financeiras, foi ainda utilizado para algumas amostras o sistema *NZYTaq II 2× Green Master Mix* (NZYTech, Portugal), em tudo semelhante ao *Supreme NZYTaq II 2× Green Master Mix* (NZYTech, Portugal), com a excepção da ausência da capacidade *hot start* pela Taq DNA.

### 2.3.2.1 Reacções de PCR

Na generalidade das três PCRs foi utilizado o sistema *Supreme NZYTaq II 2× Green Master Mix* (NZYTech, Portugal), como referido anteriormente, e a mistura de reacção foi preparada de acordo com as recomendações do fabricante, adicionando-se, por amostra: 8,5 µL de água desionizada ultrapura livre de nucleases *Milli-Q®* (*Reagent Grade Ultrapure Systems*, Merck Millipore, EUA); 12,5 µL de *master mix*, contendo a enzima *Supreme NZYTaq II DNA polymerase* (NZYTech, Portugal) a uma concentração de 0,2 U/µL, MgCl<sub>2</sub> (concentração final 2,5 mM) e dNTPs, tampão de reacção e aditivos a concentrações optimizadas pelo fabricante; e 1 µL de cada *primer* (concentração final 0,4 µM, cada), perfazendo um volume final de 25 µL, após a adição do disco contendo o DNA matriz à mistura (primeira PCR) ou de 2 µL do produto de amplificação da reacção imediatamente anterior (segunda e terceira PCRs).

Todas as reacções de amplificação foram realizadas no termociclador *PCR Express* (Thermo Hybaid, Reino Unido) consistindo, cada uma, numa etapa de desnaturação inicial, seguida de vários ciclos de amplificação (incluindo os passos de desnaturação, hibridação e extensão) e uma etapa de extensão final, cujas respectivas condições experimentais se encontram descritas nas Tabelas 3 a 5. Os produtos de amplificação foram mantidos a 4 °C até à realização da reacção seguinte (segunda ou terceira) ou até à sua visualização em gel de agarose, no caso dos produtos da terceira reacção, sendo posteriormente armazenados a -20 °C. É importante referir que para as

amostras inicialmente não amplificáveis nas condições descritas para a primeira e segunda reacções, foram aplicados protocolos alternativos com temperaturas de hibridação inferiores, em diferentes combinações, nessas mesmas reacções, mantendo-se inalterado o protocolo da terceira PCR.

**Tabela 3: Condições de amplificação aplicadas na primeira PCR.**

| Etapas               | Condições de amplificação |              |                 | N.º de ciclos |
|----------------------|---------------------------|--------------|-----------------|---------------|
| Desnaturação inicial | 95°C – 5 min.             |              |                 | 1             |
| Amplificação         | Desnaturação              | Hibridação   | Extensão        | –             |
|                      | 94 °C – 30 s              | 45 °C – 30 s | 72 °C – 1 min   | 10            |
|                      | 94 °C – 30 s              | 50 °C – 30 s | 72 °C – 1,5 min | 10            |
|                      | 94 °C – 30 s              | 55 °C – 30 s | 72 °C – 2 min   | 15            |
| Extensão final       | 72 °C – 7 min             |              |                 | 1             |
| Manutenção           | 4 °C – ∞                  |              |                 | –             |

**Tabela 4: Condições de amplificação aplicadas na segunda PCR.**

| Etapas               | Condições de amplificação |              |                    | N.º de ciclos |
|----------------------|---------------------------|--------------|--------------------|---------------|
| Desnaturação inicial | 95°C – 5 min.             |              |                    | 1             |
| Amplificação         | Desnaturação              | Hibridação   | Extensão           | –             |
|                      | 94 °C – 30 s              | 45 °C – 30 s | 72 °C – 45 s       | 10            |
|                      | 94 °C – 30 s              | 50 °C – 30 s | 72 °C – 1 min 15 s | 10            |
|                      | 94 °C – 30 s              | 55 °C – 30 s | 72 °C – 1 min 45 s | 15            |
| Extensão final       | 72 °C – 7 min             |              |                    | 1             |
| Manutenção           | 4 °C – ∞                  |              |                    | –             |

**Tabela 5: Condições de amplificação aplicadas na terceira PCR.**

| Etapas               | Condições de amplificação |              |              | N.º de ciclos |
|----------------------|---------------------------|--------------|--------------|---------------|
| Desnaturação inicial | 95 °C – 5 min.            |              |              | 1             |
| Amplificação         | Desnaturação              | Hibridação   | Extensão     | –             |
|                      | 94 °C – 30 s              | 50 °C – 30 s | 72 °C – 30 s | 35            |
| Extensão final       | 72 °C – 7 min             |              |              | 1             |
| Manutenção           | 4 °C – ∞                  |              |              | –             |

### 2.4 Visualização dos produtos de PCR

Os fragmentos de DNA obtidos por PCR foram observados depois da sua migração por electroforese em gel de agarose (*SeaKem® LE Agarose*, Lonza, Suíça) a 1,2% (p/v), em tampão TAE 0,5 × e corado com brometo de etídeo (0,5 µg/mL), utilizando o equipamento *Power Pac 300* (Bio-Rad, EUA) como fonte de corrente eléctrica. Foram aplicados 5 µL de produto de amplificação por poço – não sendo necessária a adição de tampão de aplicação uma vez que a mistura reaccional do sistema de amplificação utilizado já continha dois corantes (xileno cianol e *Orange G*) – e aplicada uma voltagem de 90 V, durante um período de migração de 1 h e 20 min, findo o qual foi efectuada a exposição do gel a radiação ultravioleta ( $\lambda=302$  nm) e captação digital da imagem gerada utilizando o equipamento *Gel Doc™ XR* (Bio-Rad, EUA). De modo a poder estimar-se o tamanho aproximado dos fragmentos de DNA amplificados foi utilizado o marcador de massa molecular *NZYDNA Ladder V* (NZYTech, Portugal).

### 2.5 Purificação dos produtos de PCR

Após a visualização dos produtos de PCR obtidos, foram seleccionados para purificação aqueles que apresentavam tamanhos estimados de 460 e 650 pb, correspondentes aos tamanhos esperados para cada um dos amplicões de PR e RT, respectivamente. Dependendo das características da banda observada no gel de agarose – especificidade da amplificação e quantidade relativa de DNA em relação ao marcador de massa molecular – foi utilizado um de dois sistemas de purificação, ambos segundo as recomendações do fabricante.

Para produtos de PCR que apresentavam bandas específicas e quantidades de DNA aceitáveis, foi realizada purificação directa a partir da mistura reaccional, utilizando o sistema *DNA Clean & Concentrator™-5* (Zymo Research, EUA), começando por se adicionar aos produtos de PCR a purificar *DNA Binding Buffer* numa proporção de 5:1 (*e.g.*, para 19 µL de produto de PCR, adicionar 95 µL de *DNA Binding Buffer*). Após homogeneização, a mistura foi transferida para uma coluna *Zymo-Spin™* colocada num tubo colector e centrifugada a  $16\,000 \times g$  (*Centrifuge 5415D*, Eppendorf, Alemanha) durante 30 s, descartando depois o eluato. Seguiu-se um passo de lavagem (realizado duas vezes) que consistiu na adição de 200 µL de *DNA Wash Buffer* à coluna, seguida de centrifugação, nas mesmas condições, descartando novamente o eluato. A coluna foi

depois transferida para um microtubo Eppendorf devidamente identificado, sendo adicionados 20  $\mu\text{L}$  de *DNA Elution Buffer* e, findo um período de incubação de 1 a 2 min à TA, centrifugada a  $16\,000 \times g$  durante 60 s obtendo-se o DNA purificado.

Para produtos de PCR que apresentavam, além de bandas correspondentes a amplicões de tamanho esperado, bandas inespecíficas, foi realizada purificação a partir de gel de agarose utilizando o sistema *Zymoclean™ Gel DNA Recovery Kit* (Zymo Research, EUA), começando por se aplicar os produtos de PCR (aproximadamente 20  $\mu\text{L}$ ) num gel de agarose (*SeaKem® LE Agarose*, Lonza, Suíça) a 2% (p/v), em tampão TAE 0,5  $\times$  e corado com brometo de etídeo (0,5  $\mu\text{g/mL}$ ). Após um período de migração de aproximadamente 2 h a 40 V (*Power Pac 300*, Bio-Rad, EUA), foram excisados do gel, sob luz ultravioleta ( $\lambda=302\text{ nm}$ ), e transferidos para microtubos Eppendorf de 2,0 mL, os fragmentos de DNA de interesse utilizando-se um escalpelo descartável. De seguida, foram adicionados três volumes de *ADB Buffer* por cada volume (previamente estimado por pesagem) de gel excisado contendo o fragmento de DNA (*e.g.*, por cada 100  $\mu\text{L}$  (mg) de gel de agarose contendo DNA, adicionar 300  $\mu\text{L}$  de *ADB Buffer*). Esta mistura foi incubada a 55 °C até a agarose estar completamente dissolvida (5 a 10 min), sendo depois transferida para uma coluna *Zymo-Spin™* colocada num tubo colector e realizado o protocolo descrito no parágrafo anterior, com excepção da adição de *DNA Binding Buffer*.

De modo a verificar se a purificação dos produtos de PCR foi bem-sucedida, aplicou-se 1  $\mu\text{L}$  de produto de purificação, misturados com cerca de 1  $\mu\text{L}$  de tampão de aplicação (Anexo 4), num de gel de agarose (*SeaKem® LE Agarose*, Lonza, Suíça) a 1,2% (p/v), em tampão TAE 0,5  $\times$  e corado com brometo de etídeo (0,5  $\mu\text{g/mL}$ ), sendo visualizados os fragmentos obtidos nas condições descritas anteriormente, após um período de migração de 20 min a 100 V (*Power Pac 300*, Bio-Rad, EUA).

### 2.6 Sequenciação nucleotídica dos produtos de PCR

A sequenciação nucleotídica dos produtos de PCR purificados ficou a cargo da empresa STAB VIDA (Caparica, Portugal), tendo sido utilizado o método de terminação de cadeia desenvolvido por Sanger *et al.* (1977) implementado em plataforma automatizada. Para tal, foram enviados 16  $\mu\text{L}$  dos produtos de purificação, de modo a garantir a quantidade mínima estimada de 400 ng de DNA por reacção de sequenciação uma vez que cada fragmento seria sequenciado nos dois sentidos (directo e inverso),

sendo igualmente enviados 50 pmoles de cada um dos *primers* iniciadores (PR3, PR4, RT1 e RT4) por reacção de sequenciação.

### 2.7 Análise bioinformática das sequências nucleotídicas

#### 2.7.1 Edição das sequências e confirmação da sua identidade

As sequências nucleotídicas directas e complementares de cada molécula de DNA obtidas por sequenciação dos produtos de PCR foram manualmente editadas com base na análise dos respectivos cromatogramas, tendo como finalidade a construção de uma sequência consenso a ser utilizada nas análises posteriores. Para tal foi utilizado o *software BioEdit Sequence Alignment Editor* (versão 7.0.9.0), disponível em <http://www.mbio.ncsu.edu/bioedit/bioedit.html>.

Posteriormente, foi realizada a confirmação da identidade das sequências consenso obtidas por comparação com sequências depositadas no *GenBank* utilizando o programa *BLAST* (do inglês, *Basic Local Alignment Search Tool*), disponível em <http://blast.ncbi.nlm.nih.gov/Blast.cgi>. Este permite calcular (em percentagem) o grau de semelhança entre sequências e ainda estimar a probabilidade de as semelhanças encontradas serem devidas ao acaso (*E value*, 0-1).

#### 2.7.2 Genotipagem automatizada das sequências de HIV-1

A genotipagem automatizada das sequências de HIV-1 obtidas foi realizada recorrendo ao *software REGA HIV-1 Subtyping Tool* (versão 3.0), disponível em <http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/>. Este utiliza métodos filogenéticos e de *bootscanning* para a identificação do subtipo de uma determinada sequência e análise de possíveis eventos de recombinação. A classificação genotípica de sequências de HIV-1 utilizando este programa baseia-se na construção de árvores filogenéticas com um conjunto de sequências de referência do grupo M, sujeitas a análise complementar de *bootstrapping*, sendo apenas considerados na identificação do subtipo valores de *bootstrap* superiores a 70%. No entanto, a análise de eventos de recombinação genética só é efectuada se as sequências em estudo possuírem tamanho superior a 800 pb.

### 2.7.3 Análise filogenética

A genotipagem das sequências de HIV-1 codificantes para cada uma das enzimas em estudo (PR e RT) foi efectuada através da construção de árvores filogenéticas por métodos não automatizados. Começou-se pela construção de um alinhamento múltiplo para cada uma das regiões do gene *pol* em estudo, recorrendo ao programa *MAFFT* versão 7 (disponível em <https://mafft.cbrc.jp/alignment/server/>), contendo as sequências em estudo, sequências obtidas por pesquisa de semelhança local entre sequências (análise *BLAST*) e de referência do grupo M (subtipos, sub-subtipos e algumas CRFs – 02\_AG, 26\_A5U e 49\_cpx) e ainda uma sequência do grupo O (como *outgroup*), retiradas da *Los Alamos HIV sequence database* (<https://www.hiv.lanl.gov/>). Sempre que necessário, os alinhamentos obtidos foram editados manualmente com recurso ao *software BioEdit Sequence Alignment Editor* (versão 7.0.9.0). As respectivas árvores filogenéticas foram construídas utilizando o *software MEGA* versão X com base no método de *Neighbor-Joining* (176) e recorrendo ao modelo de dois parâmetros de Kimura (177) para o cálculo das matrizes de distância genética, seleccionando uma frequência igual de bases, uma taxa de transições igual à taxa de transversões e a opção *complete deletion* para as colunas contendo espaços abertos (*gaps*). A robustez das topologias sugeridas foi avaliada pelo método de *bootstrapping*, com base em 1000 amostragens aleatórias, tendo-se considerado os agrupamentos formados como sendo consistentes do ponto de vista filogenético quando os valores de *bootstrap* eram iguais ou superiores a 70%.

#### 2.7.3.1 Análise de sequências divergentes

No caso das sequências nucleotídicas cujos ramos de suporte nas árvores filogenéticas evidenciavam posicionamento ambíguo ou uma elevada divergência filogenética, foi realizada uma análise complementar para identificação de eventuais fenómenos de recombinação ou hipermutação genética. A identificação de fenómenos de recombinação foi realizada recorrendo a métodos de *bootscanning*, utilizando o *software SimPlot* versão 3.5.1 (disponível em: <http://sray.med.som.jhmi.edu/SCRsoftware/simplot/>), aplicando o método de *Neighbor-Joining* para a inferência filogenética, com robustez estimada por *bootstrapping*, com base em 1000 amostragens aleatórias, e considerada significativa quando os valores de *bootstrap* eram, no mínimo, 70%, e o modelo de dois parâmetros de Kimura para o

cálculo das matrizes de distância genética. A pesquisa de fenómenos de hipermutação por ocorrência de hiperadenilação (G→A) resultante da actividade das APOBEC (do inglês, *apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like*) foi realizada com base no relatório interpretativo da análise das sequências aminoacídicas obtido utilizando o programa *HIVdb – Genotypic Resistance Interpretation Algorithm* da *Stanford University HIV Drug Resistance Database* (versão 8.5), no qual são discriminadas este tipo de mutações (ver secção seguinte).

### 2.8 Análise bioinformática das sequências aminoacídicas

#### 2.8.1 Pesquisa de mutações associadas a resistência a anti-retrovirais

Após tradução das sequências nucleotídicas em estudo, foram obtidas sequências de aminoácidos putativas para a realização da pesquisa de mutações associadas a diminuição da susceptibilidade aos ARVs inibidores da PR e da RT e detecção de outros polimorfismos genéticos, com recurso ao programa *HIVdb – Genotypic Resistance Interpretation Algorithm* da *Stanford University HIV Drug Resistance Database* (versão 8.5), disponível em <https://hivdb.stanford.edu/hivdb/by-mutations/>. A utilização deste programa permite a identificação de mutações principais (*major*) e acessórias (*minor*), assim como outras mutações não associadas a resistência as ARVs (*other mutations*), por comparação das sequências em estudo com uma sequência consenso de referência do subtipo B, para cada uma das regiões codificantes em análise. Com base nessa comparação, é atribuída a cada mutação uma pontuação individual relativa a cada fármaco considerado na análise, que reflecte, com base em diversos estudos publicados, a probabilidade da mutação encontrada estar relacionada com a aquisição de resistência ao fármaco em questão. A soma das pontuações obtidas para um determinado fármaco permite, de acordo com o valor total obtido, estimar para cada sequência o nível de resistência fenotípica a esse fármaco, numa das cinco categorias seguintes:

- ❖ Susceptibilidade (pontuação entre 0 e 9) – sem evidência de redução da susceptibilidade ao ARV em comparação com a estirpe selvagem;
- ❖ Resistência potencial de nível baixo (pontuação entre 10 e 14) – o vírus não apresenta susceptibilidade diminuída ao fármaco, mas possui mutações

indicativas de exposição prévia a ARVs ou que apenas conferem resistência quando associadas a mutações adicionais;

- ❖ Resistência de nível baixo (pontuação entre 15 e 29) – o vírus pode apresentar uma diminuição da susceptibilidade ao fármaco *in vitro*, sendo que os indivíduos infectados com o mesmo podem desenvolver uma resposta virológica sub-ótima ao tratamento com esse fármaco;
- ❖ Resistência de nível intermédio (pontuação entre 30 e 59) – indicativo de elevada probabilidade de susceptibilidade reduzida ao fármaco, mas com retenção de um nível considerável da sua eficácia;
- ❖ Resistência de nível elevado (pontuação igual ou superior a 60) – indicativo de um nível de resistência semelhante ao observado *in vitro* em vírus altamente resistentes ao fármaco ou que se traduz numa resposta virológica muito reduzida, ou mesmo nula, em termos clínicos, nos indivíduos infectados.

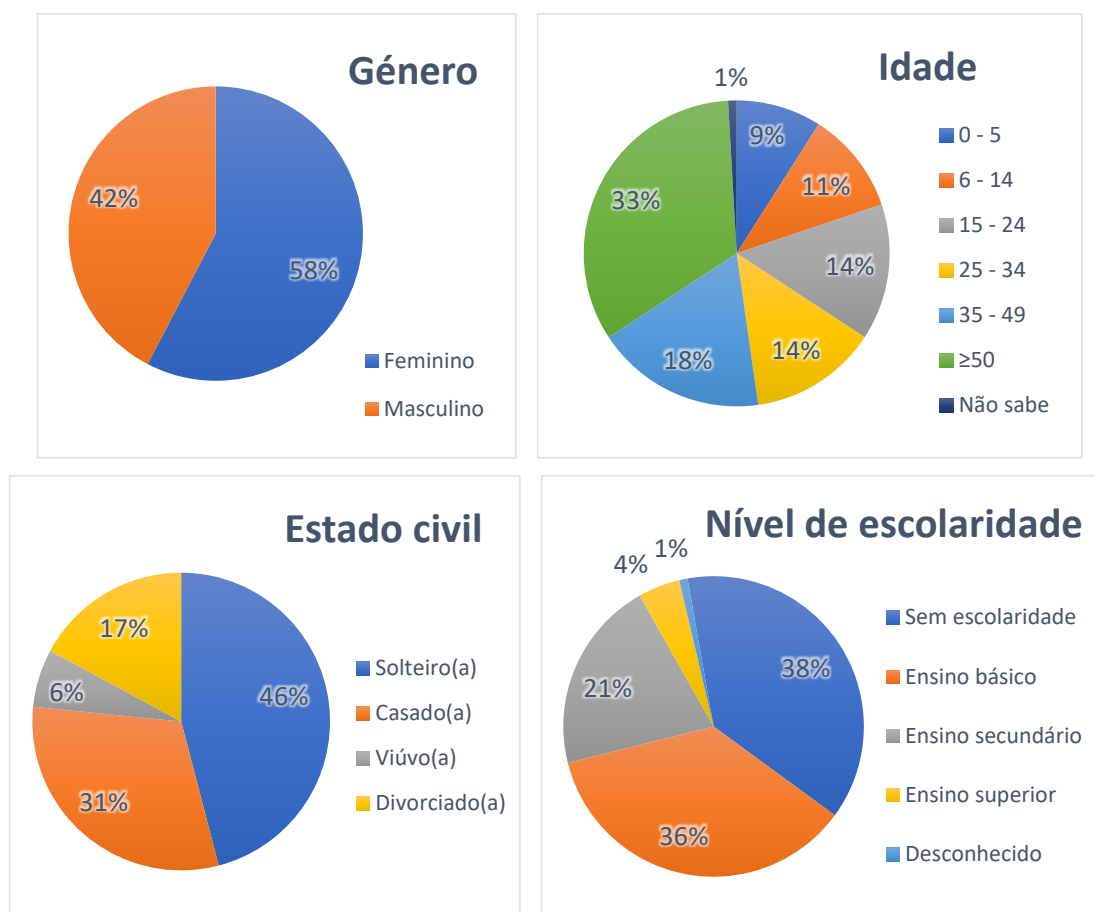
### 2.9 Análise estatística

Para os dados demográficos e clínicos que caracterizam a amostra populacional em estudo, como, género, estado civil, nível de escolaridade, local e área de residência e outras infecções diagnosticadas, foi realizada uma análise de distribuição, dada a sua natureza categórica. Em relação à idade, para além da análise de distribuição realizada, foi também calculada a média e o desvio padrão. Por sua vez, a associação dos diferentes polimorfismos genéticos com os genótipos de HIV-1 em circulação na amostra populacional em estudo (pesquisa de assinaturas genéticas) e da presença de mutações de hiperadenilação com a impossibilidade de genotipar as sequências nucleotídicas foram analisadas recorrendo ao teste exacto de Fisher bilateral aplicado à análise de tabelas de contingência 2×2, considerando um valor de significância de 1%. A pesquisa de associação entre a ocorrência de fenómenos de hipermutação e os genótipos encontrados foi realizada com recurso ao teste  $\chi^2$  aplicado à análise de tabelas n×2 (disponível em <http://www.physics.csbsju.edu/stats/>), com um nível de significância de 1%.

### 3 Resultados

#### 3.1 Caracterização da amostra populacional em estudo

O presente estudo incluiu amostras de sangue de um total de 111 indivíduos seropositivos para HIV-1 que se encontravam internados ou eram utentes de consultas externas/banco de urgência do HGB. Todos os indivíduos eram de nacionalidade angolana e, à data da colheita das amostras, tinham idades compreendidas entre 0 e 78 anos (idade média de  $35 \pm 20,3$  anos), sendo 42% do género masculino e 58% do género feminino. Os indivíduos em estudo eram maioritariamente solteiros (46%) ou casados (31%) e, apesar de a maioria possuir pelo menos o ensino básico (61%), existe uma percentagem considerável dos mesmos sem escolaridade (38%) – Figura 14, Anexo 3.



**Figura 14:** Representação gráfica, em percentagem, dos dados demográficos da amostra populacional em estudo, relativamente a género, idade, estado civil e nível de escolaridade.

Todos os indivíduos eram residentes na província de Benguela, a maioria na cidade de Benguela (46%) e os restantes noutras regiões da província, sendo que, de modo

geral, a maioria residia na zona urbana das respectivas localidades (62%). Uma vez que este estudo se insere num outro que tem também como objectos de estudo outras infecções de transmissão sexual, como sífilis e hepatite B (como referido na secção 1.15), foi possível verificar que 14% (16/111) dos indivíduos em estudo tinham marcadores serológicos positivos de outras infecções além da infecção pelo HIV-1, dos quais 63% (10/16) tinham sífilis (activa ou passada), 31% (5/16) estavam infectados com o vírus da hepatite B e um único indivíduo obteve resultados positivos para todas as infecções (Figura 15, Anexo 3).

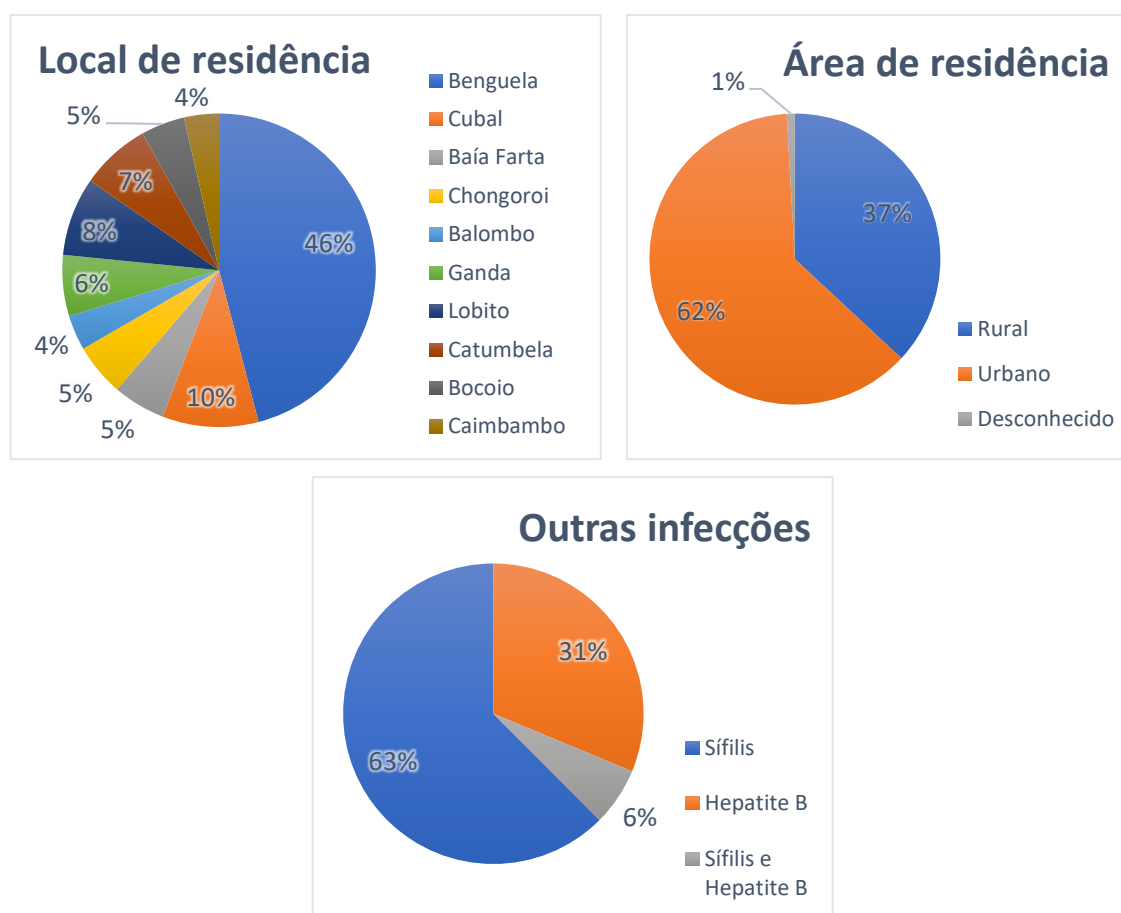
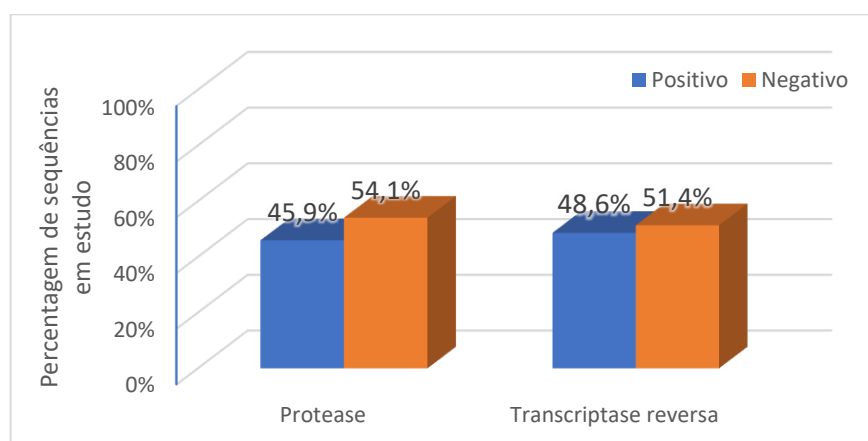


Figura 15: Representação gráfica, em percentagem, dos dados demográficos e clínicos da amostra populacional em estudo, relativamente a local de residência, área de residência e outras infecções diagnosticadas.

### 3.2 Amplificação por PCR das regiões codificantes da protease e da transcriptase reversa no gene *pol*

De modo a ser possível atingir os objectivos definidos, foi necessário introduzir alterações ao desenho do estudo inicialmente delineado, nomeadamente nos protocolos de preparação do DNA das amostras e de amplificação das regiões codificantes da PR e

da RT, como descrito nas secções 2.2 e 2.3.2. Foram preparados discos da mancha de sangue dos cartões *FTA*<sup>®</sup> das 111 amostras, procedendo-se depois à amplificação por dupla *nested* PCR das regiões codificantes da PR e da RT a partir do DNA proviral contido nos mesmos com recurso ao protocolo descrito na secção 2.3.2. Depois da realização do protocolo base, ou protocolos alternativos, sempre que as amostras não eram inicialmente amplificáveis (ver secção 2.3.2.1), obteve-se um total de 51 amostras positivas para a PR (45,9%) e 54 amostras positivas para a RT (48,6%), sendo que para as restantes amostras não foram obtidos produtos de amplificação apesar de terem sido testados vários protocolos (Figura 16). Cruzando as identificações das amostras positivas para cada uma das regiões amplificadas, constatou-se que foi possível amplificar pelo menos uma das regiões para um conjunto de 57 amostras (correspondendo a 51,4% do total das amostras incluídas no estudo), das quais, 48 (84,2%) foram positivas para ambas as regiões, 3 (5,3%) foram positivas apenas para a PR e 6 (10,5%) foram positivas somente para a RT.



**Figura 16:** Representação gráfica dos resultados obtidos (em percentagem) após amplificação por dupla *nested* PCR das regiões codificantes da protease e da transcriptase reversa.

A visualização dos produtos de PCR após migração por electroforese em gel de agarose permitiu a identificação das amostras que apresentavam fragmentos específicos de amplificação com cerca de 460 (Figura 17, poços 2 a 4) e 650 pb (Figura 17, poços 14 a 16), correspondendo aos produtos de amplificação da PR e da RT, respectivamente, assim como verificar a presença de fragmentos inespecíficos nalgumas amostras positivas (Figura 17, poços 2 e 14, por exemplo).

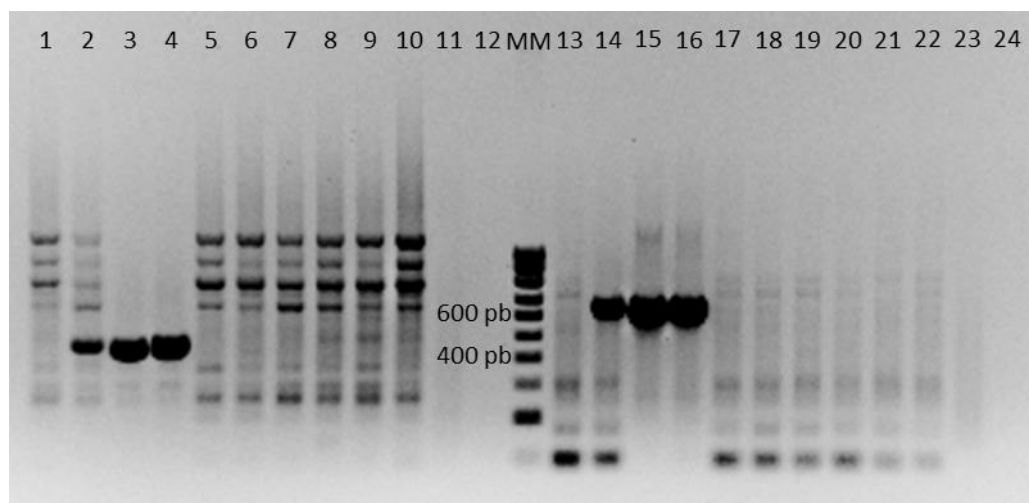


Figura 17: Visualização dos padrões de migração dos produtos de amplificação por dupla *nested* PCR das regiões codificantes da protease (à esquerda de MM) e transcriptase reversa (à direita de MM), após electroforese em gel de agarose a 1,2% (p/v), em tampão TAE 0,5 × e corado com brometo de etídeo (0,5 µg/mL). **Legenda:** MM – Marcador de massa molecular NZYDNA Ladder V (NZYTech, Portugal); 1 a 10 – Fragmentos de DNA amplificados da região da protease; 13 a 22 – Fragmentos de DNA amplificados da região da transcriptase reversa; 11, 12, 23 e 24 – Controlos negativos.

A purificação dos fragmentos amplificados de tamanho esperado foi realizada utilizando um de dois métodos, consoante as características da banda observada no gel de agarose, como descrito na secção 2.5. Assim, as amostras que apresentavam fragmentos de tamanho esperado com quantidades de DNA aceitáveis e ausência de inespecificidades [10/51 (19,6%), no caso da PR e 35/54 (64,8%), no caso da RT] foram purificadas directamente a partir dos produtos de PCR, enquanto amostras que apresentavam, além das bandas específicas com quantidades de DNA aceitáveis, várias bandas inespecíficas [41/51 (80,4%), no caso da PR e 19/54 (35,2%), no caso da RT] foram purificadas a partir de um gel de agarose, após separação electroforética. Uma vez obtidos os fragmentos de DNA purificados, estes foram enviados para sequenciação nucleotídica a cargo da empresa STAB VIDA (Caparica, Portugal).

### 3.3 Análise bioinformática das sequências nucleotídicas

#### 3.3.1 Edição das sequências e confirmação da sua identidade

Concluída a sequenciação nucleotídica dos 105 fragmentos purificados com os respectivos *primers* internos de amplificação, procedeu-se à comparação e edição das sequências directa e complementar obtidas para cada amostra, com base na análise dos respectivos cromatogramas, para a construção de sequências consenso. Em algumas sequências houve necessidade de atribuir bases degeneradas, segundo o código da IUPAC

(do inglês, *International Union of Pure and Applied Chemistry*) (Anexo 6, que também inclui o código para os aminoácidos), nas posições em que não era possível definir um único nucleótido com base na análise do cromatograma devido à existência de *quasispecies*. No total, foram obtidas 50 sequências consenso PR e 53 sequências consenso RT, sendo que para 2 dos 105 fragmentos de DNA sequenciados (1,9%) (ambos da mesma amostra, ou seja, um fragmento PR e um fragmento RT) não foi possível construir sequências consenso devido à qualidade das sequências nucleotídicas obtidas (por vezes truncadas) e respectivos cromatogramas, nos quais se verificou a presença de “ruído de fundo” e/ou um elevado número de degenerações que dificultaram a sua interpretação.

Uma vez obtidas as sequências consenso de cada amostra, procedeu-se à confirmação da sua identidade por comparação com sequências depositadas em bases de dados utilizando o programa *BLAST* como descrito na secção 2.7.1. Resumidamente, observaram-se valores de *E value* iguais ou muito próximos de zero ( $\leq 4 \times 10^{-115}$ ) e percentagens de identidade com sequências de HIV-1 depositadas em base de dados iguais ou superiores a 86% (dados não apresentados), permitindo confirmar a obtenção de sequências das regiões codificantes da PR e da RT no gene *pol*. As sequências HIV-1 depositadas em base de dados que apresentaram maior grau de semelhança com as sequências em estudo, assim como os respectivos *E value*, foram registados para posterior utilização na análise filogenética.

#### 3.3.2 Genotipagem automatizada das sequências de HIV-1

Após a confirmação da identidade das sequências em estudo, foi realizada uma genotipagem preliminar automatizada das mesmas, antes da construção de árvores filogenéticas, utilizando o *software REGA HIV-1 Subtyping Tool* (versão 3.0), por forma a adquirir um conhecimento inicial acerca da diversidade genética presente na amostra. Uma vez que a análise de eventos de recombinação genética só é efectuada neste *software* se as sequências submetidas possuírem tamanho superior a 800 pb, a mesma não foi realizada pois as sequências em estudo possuem um tamanho inferior. Na Tabela 6 apresentam-se os resultados desta análise, verificando-se um predomínio, em ambas as regiões codificantes (PR e RT), dos genótipos C [20/50 (40,0%), no caso da PR e 19/53, (35,8%), no caso da RT] e F1 [8/50 (16,0%), no caso da PR e 9/53 (17,0%), no caso da RT], sendo igualmente detectados os genótipos G [5/50 (10,0%), no caso da PR e 4/53

(7,5%), no caso da RT], A1 [2/50 (4,0%), no caso da PR e 2/53 (3,8%), no caso da RT] e J [1/50 (2,0%), no caso da PR e 1/53 (1,9%), no caso da RT], mas também o genótipo D numa única sequência da enzima RT (1,9%). De realçar que a distribuição dos diferentes genótipos não parece diferir substancialmente entre as duas regiões codificantes em estudo. Verifica-se ainda que não foi atribuído um genótipo a 28,0% (14/50) e 32,1% (17/53) das sequências PR e RT, respectivamente, o que pode indicar a presença de sequências com genótipo recombinante ou geneticamente muito divergentes das sequências de referência utilizadas para a realização da análise.

**Tabela 6: Resultados da genotipagem automatizada para as regiões codificantes da protease e da transcriptase reversa utilizando o software *REGA HIV-1 Subtyping Tool* (versão 3.0).**

| Genótipo      | Protease (n=50) | Transcriptase reversa (n=53) |
|---------------|-----------------|------------------------------|
| A (A1)        | 2 (4,0%)        | 2 (3,8%)                     |
| C             | 20 (40,0%)      | 19 (35,8%)                   |
| D             | —               | 1 (1,9%)                     |
| F (F1)        | 8 (16,0%)       | 9 (17,0%)                    |
| G             | 5 (10,0%)       | 4 (7,5%)                     |
| J             | 1 (2,0%)        | 1 (1,9%)                     |
| Não atribuído | 14 (28,0%)      | 17 (32,1%)                   |

### 3.3.3 Análise filogenética

Concluída a análise preliminar descrita na secção anterior, procedeu-se à análise filogenética das sequências em estudo para genotipagem das mesmas, por meio da construção de árvores filogenéticas com recurso a métodos não automatizados (como descrito na secção 2.7.3) que permitem a selecção das sequências e algoritmos a utilizar, de modo a obter, para cada uma das regiões codificantes do gene *pol* em estudo (PR e RT), a árvore filogenética com maior robustez topológica possível. Para tal, foi construído um alinhamento múltiplo para cada uma das regiões, incluindo sequências de referência disponíveis na *Los Alamos HIV Sequence Database* (<https://www.hiv.lanl.gov/>), sequências obtidas por pesquisa de semelhança local entre sequências (análise *BLAST*) e as sequências em estudo (Tabela 7), a partir dos quais foram construídas as árvores filogenéticas recorrendo ao software *MEGA* versão X pelo método *Neighbor-Joining* (176), utilizando o modelo de dois parâmetros de Kimura (177) para o cálculo das matrizes de distância genética. A robustez das topologias sugeridas foi avaliada pelo

método de *bootstrapping* com base em 1000 amostragens aleatórias, tendo-se considerado os agrupamentos formados como sendo consistentes do ponto de vista filogenético quando os valores de *bootstrap* eram iguais ou superiores a 70%.

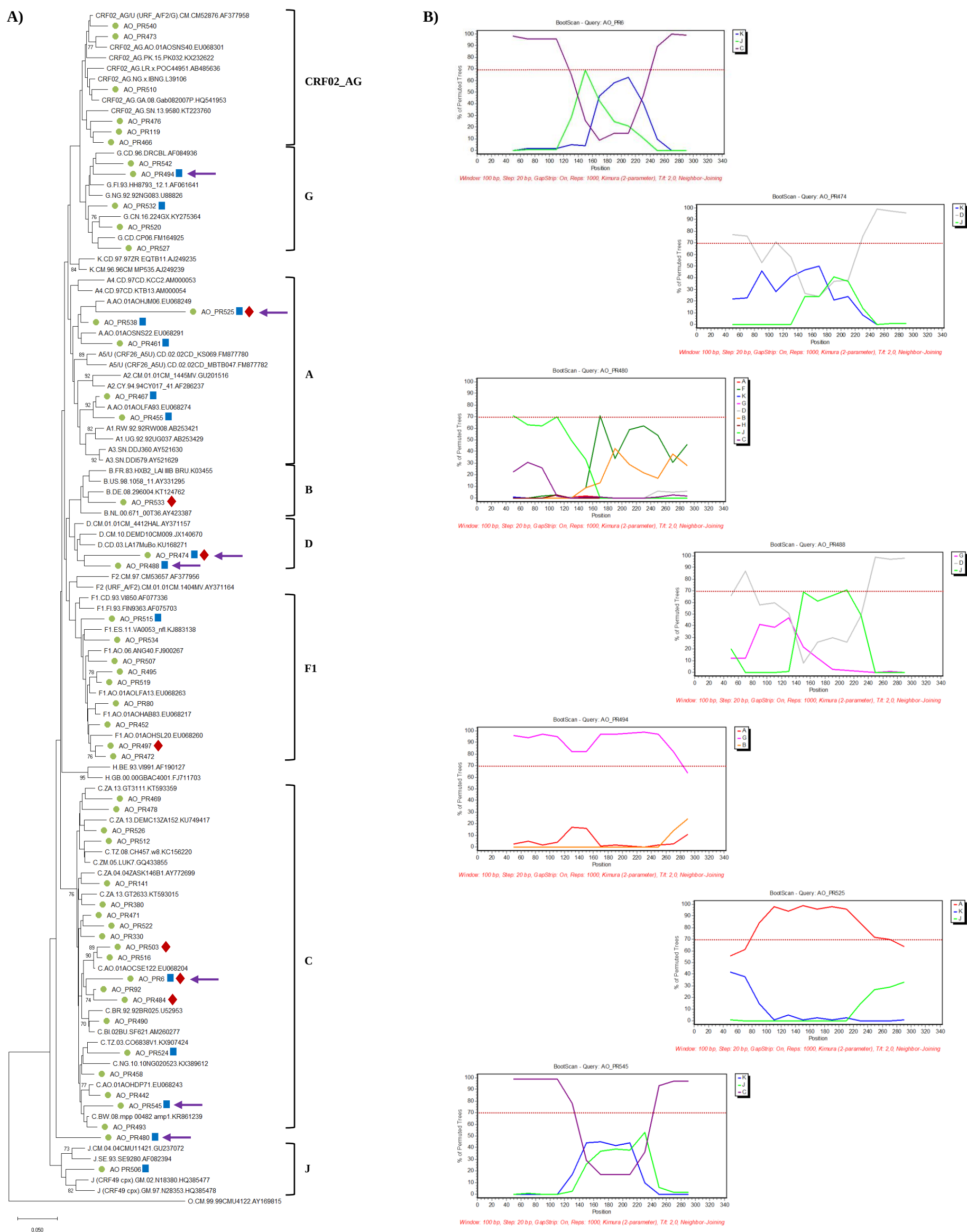
Tabela 7: Número total de sequências e tamanho do alinhamento múltiplo (em nucleótidos) utilizados na construção das árvores filogenéticas PR e RT.

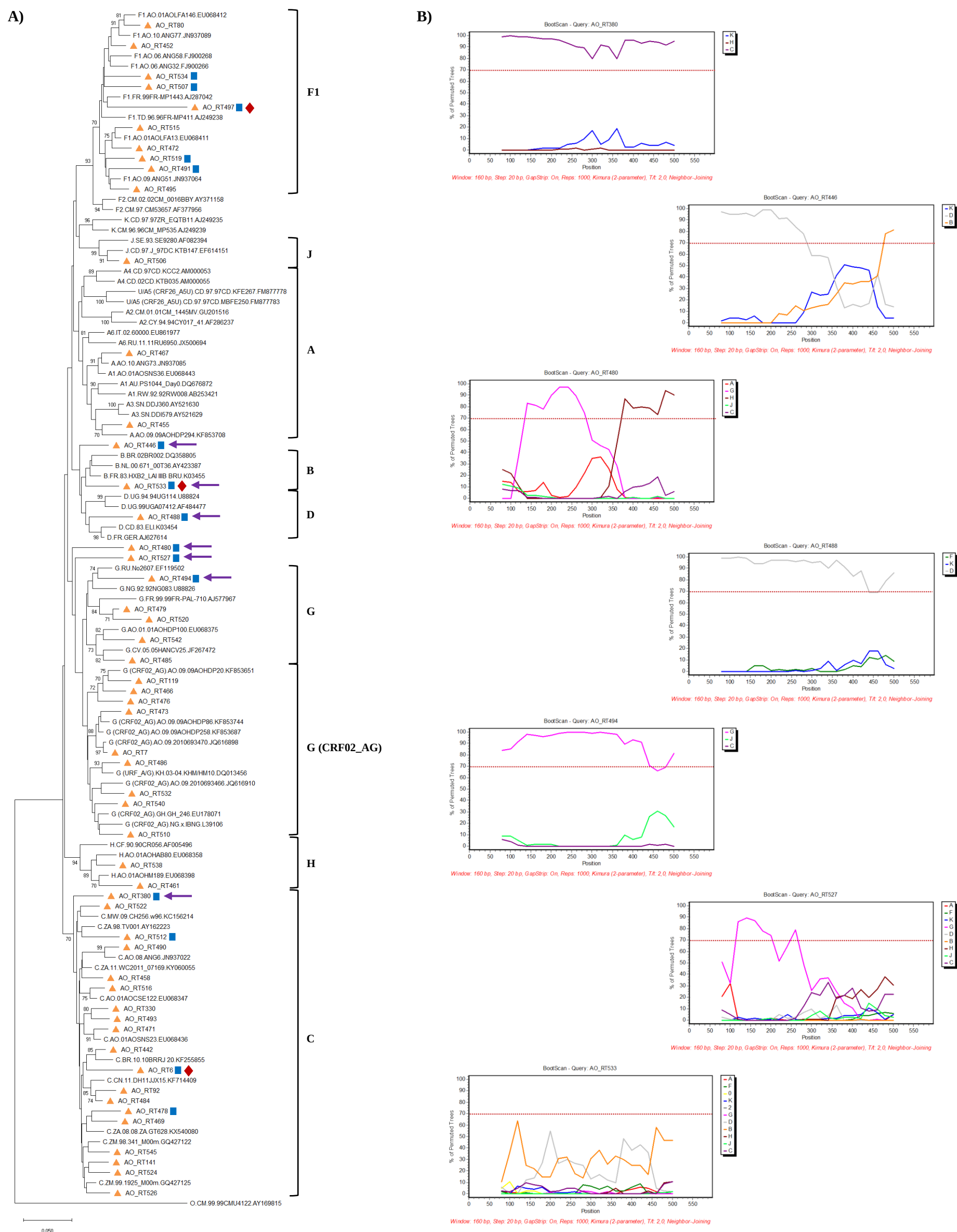
| Árvore filogenética   | Sequências de referência | Sequências análise <i>BLAST</i> | Sequências em estudo | Total de sequências | Total de nucleótidos |
|-----------------------|--------------------------|---------------------------------|----------------------|---------------------|----------------------|
| Protease              | 42                       | 21                              | 50                   | 113                 | 307                  |
| Transcriptase reversa | 33                       | 31                              | 53                   | 117                 | 532                  |

Após a construção das árvores filogenéticas para as regiões codificantes da PR e da RT, verificou-se que, em ambas, as sequências em estudo formavam grupos monofiléticos juntamente com as sequências de referência do HIV-1 (recolhidas da *Los Alamos HIV Sequence Database* e obtidas por análise *BLAST*) com topologias, de um modo geral, bem definidas, sendo que no caso da região RT a maioria dos agrupamentos formados eram suportados por valores de *bootstrap* mínimos de 70% (Figura 18 A). No entanto, o mesmo não se verificou para PR, em que a maioria dos agrupamentos formados era suportada por valores de *bootstrap* inferiores a 70%, com exceção do agrupamento do subtipo C, suportado por um valor de 76% (Figura 18 A). Desse modo, foi possível identificar inequivocamente o genótipo da maioria das sequências em estudo, ainda que com um grau de certeza menor na classificação genotípica da PR. Os resultados da genotipagem das sequências em estudo obtidos unicamente através da análise das árvores filogenéticas construídas encontram-se resumidos na Tabela 8, na qual genótipo “não atribuído” corresponde a sequências com elevada divergência genética ou com um posicionamento ambíguo na árvore filogenética e que, por isso, requerem uma análise adicional para a determinação do respectivo genótipo.

Finalmente, dada a estrutura genômica das sequências de referência das três CRFs do HIV-1 utilizadas na construção das árvores filogenéticas PR e RT (genomas mosaico, Anexo 7), é importante referir que a região codificante da PR da CRF02\_AG é classificada como A/G, enquanto a região codificante em estudo da RT é classificada como subtipo G; a região codificante da PR da CRF26\_A5U é classificada como A5/U, enquanto a região codificante em estudo da RT é classificada como U/A5; e a região codificante da PR da CRF49\_cpx é classificada como subtipo J com base nos resultados

da genotipagem rápida automatizada das respectivas sequências de referência utilizadas recorrendo ao *software REGA HIV-1 Subtyping Tool* (versão 3.0), com valores de *bootstrap* de cerca de 100% (Anexo 8), apesar do seu mapa genómico indicar A/J para a mesma região (Anexo 7). Com base nesta informação foi possível genotipar correctamente as sequências PR e RT em estudo que formam grupos monofiléticos com as sequências de referência do HIV-1 das referidas CRFs utilizadas nas respectivas árvores filogenéticas.





**Figura 19: A)** Relações filogenéticas entre 117 sequências nucleotídicas da região codificante da transcriptase reversa, baseadas num alinhamento múltiplo de 532 nucleótidos. Na árvore filogenética foram incluídas as 53 sequências em estudo (indicadas com o símbolo ▲), 33 sequências de referência do grupo M (A1-A4, A6, B-D, F1, F2, G, H, J, K, CRF02\_AG e CRF26\_A5U) e O (utilizada como *outgroup*) e 31 sequências obtidas por pesquisa de semelhança local entre sequências (análise *BLAST*). Com excepção das sequências em estudo, a designação de cada sequência inicia-se com o respectivo genótipo, seguido do código do país de origem (AO – Angola, AU – Austrália, BR – Brasil, CD – República Democrática do Congo, CF – República Centro-Africana, CM – Camarões, CN – China, CV – Cabo Verde, CY – Chipre, FR – França, GH – Gana, IT – Itália, KH – Camboja, MW – Malawi, NG – Nigéria, NL – Holanda, RU – Rússia, RW – Ruanda, SE – Suécia, SN – Senegal, TD – Chade, UG – Uganda, ZA – África do Sul, ZM – Zâmbia), ano (quando conhecido), nome da estirpe/isolado viral e número de acesso no *GenBank*. A inferência filogenética foi realizada com recurso ao *software* *MEGA* versão X, tendo por base o modelo de *Neighbor-Joining* e o método de dois parâmetros de Kimura para o cálculo das distâncias genéticas (a escala indica 5% de distância genética). A robustez da inferência filogenética foi analisada por *bootstrapping* de 1000 amostragens aleatórias, considerando-se significativos valores  $\geq 70\%$ . As sequências seleccionadas para análise de *bootscanning* estão indicadas com o símbolo ■ e as sequências nas quais foi detectada a ocorrência de fenómenos de hipermutação estão indicadas com o símbolo ◆. **B)** Gráficos seleccionados da análise de *bootscanning*, das sequências indicadas em A) com uma seta roxa, utilizando os mesmos modelos aplicados à árvore filogenética, com recurso ao *software* *SimPlot* (versão 3.5.1) com janela 100 pb e passo 20 pb. As linhas horizontais tracejadas indicam o *cut-off* de 70%.

**Tabela 8: Resultados da genotipagem para as regiões codificantes da protease e da transcriptase reversa após construção de árvores filogenéticas.**

| Genótipo      | Protease (n=50) | Transcriptase reversa (n=53) |
|---------------|-----------------|------------------------------|
| A             | –               | 2 (3,8%)                     |
| B             | 1 (2,0%)        | –                            |
| C             | 17 (34,0%)      | 15 (28,3%)                   |
| F (F1)        | 8 (16,0%)       | 5 (9,4%)                     |
| G             | 3 (6,0%)        | 4 (7,5%)                     |
| CRF02_AG      | 6 (12,0%)       | 9 (17,0%)                    |
| H             | –               | 2 (3,8%)                     |
| J             | –               | 1 (1,9%)                     |
| Não atribuído | 15 (30,0%)      | 15 (28,3%)                   |

Como seria de esperar, verificou-se um elevado grau de concordância entre os resultados da análise filogenética manual (Tabela 8) e os obtidos previamente utilizando o *software REGA HIV-1 Subtyping Tool* (versão 3.0) (Tabela 6) para ambas as regiões em estudo. No entanto, foram encontradas algumas diferenças, a ser mencionadas seguidamente. Em relação à região codificante da PR, um conjunto de sete sequências (AO\_PR119, AO\_PR466, AO\_PR476, AO\_PR495, AO\_PR510, AO\_PR533 e AO\_PR540) não subtipadas recorrendo à genotipagem automatizada foi classificado genotipicamente com base na análise da árvore filogenética construída, enquanto para um grupo de oito sequências (AO\_PR6, AO\_PR455, AO\_PR467, AO\_PR506, AO\_PR515, AO\_PR524, AO\_PR532 e AO\_PR545) ocorreu exactamente o contrário. Relativamente à região codificante da RT, à semelhança do que foi verificado para a região da PR, doze sequências (AO\_RT7, AO\_RT119, AO\_RT461, AO\_RT466, AO\_RT473, AO\_RT476, AO\_RT479, AO\_RT486, AO\_RT510, AO\_RT532, AO\_RT538 e AO\_RT542) sem um genótipo atribuído previamente pela genotipagem automatizada, foram classificadas através da análise da árvore filogenética, enquanto outras dez sequências (AO\_RT6, AO\_RT380, AO\_RT478, AO\_RT488, AO\_RT491, AO\_RT494, AO\_RT497, AO\_RT512, AO\_RT519 e AO\_RT534), subtipadas através do método automatizado, não foram possíveis de classificar inequivocamente através da análise filogenética manual.

No decorrer da análise das árvores filogenéticas observou-se que algumas sequências apresentavam um posicionamento ambíguo na respectiva árvore filogenética, o que dificultava, ou impossibilitava mesmo, a sua classificação genotípica, ou eram

suportadas por ramos longos, indicativos de uma elevada divergência genética. Ambos os casos podem dever-se a eventos de recombinação genética ou de hipermutação por ocorrência de hiperadenilação (G→A) resultante da actividade das APOBEC, sendo que para verificar a primeira hipótese foi realizada uma análise de *bootscanning* utilizando o *software SimPlot* (versão 3.5.1), sendo a segunda possibilidade analisada com base no relatório interpretativo da análise das sequências aminoacídicas obtido utilizando o programa *HIVdb – Genotypic Resistance Interpretation Algorithm* da *Stanford University HIV Drug Resistance Database* (versão 8.5), no qual é discriminado esse tipo de mutações. Foram seleccionadas as seguintes sequências para análise de *bootscanning*: AO\_PR6, AO\_PR455, AO\_PR461, AO\_PR467, AO\_PR474, AO\_PR480, AO\_PR488, AO\_PR494, AO\_PR506, AO\_PR515, AO\_PR524, AO\_PR525, AO\_PR532, AO\_PR538 e AO\_PR545, relativamente à região codificante da PR; e AO\_RT6, AO\_RT380, AO\_RT446, AO\_RT478, AO\_RT480, AO\_RT488, AO\_RT491, AO\_RT494, AO\_RT497, AO\_RT507, AO\_RT512, AO\_RT519, AO\_RT527, AO\_RT533 e AO\_RT534, em relação à região codificante da RT, totalizando 30 sequências (assinaladas nas árvores filogenéticas das Figuras 18 e 19 com quadrados de cor azul).

Para a região da PR, a análise de *bootscanning* permitiu classificar as sequências AO\_PR494 e AO\_PR532 como do subtipo G e as sequências AO\_PR525, AO\_PR538, AO\_PR461, AO\_PR467 e AO\_PR455 como do subtipo A, não sendo possível especificar neste caso um sub-subtipo e tratando-se a primeira (AO\_PR525) de uma sequência divergente desse subtipo (Figura 18 B). A sequência AO\_PR515 é do subtipo F, mais especificamente do sub-subtipo F1; as sequências AO\_PR6, AO\_PR524 e AO\_PR545 são do subtipo C (Figura 18 B); e a sequência AO\_PR506 trata-se de uma sequência do subtipo J. É importante salientar que os gráficos de *bootscan* das sequências AO\_PR6 e AO\_PR545 são indicativos de algum grau de divergência relativamente ao subtipo C, no entanto, essa divergência não é sugerida pela topologia da árvore filogenética ou pelo posicionamento das sequências na mesma, não estando, também, necessariamente relacionada com o tamanho dos respectivos ramos (Figura 18). Por fim, não foi possível classificar genotipicamente as sequências AO\_PR474, AO\_PR480 e AO\_PR488 através da análise de *bootscanning*, tendo-lhes sido atribuída a designação U (Figura 18 B), sendo que a sua análise *SimPlot* (que permite avaliar o grau de similaridade entre as sequências) foi igualmente inconclusiva (dados não apresentados). No entanto, as sequências

AO\_PR474 e AO\_PR488 apresentam, nos respectivos gráficos de *bootscan*, homologia parcial com o subtipo D, pelo que podem eventualmente tratar-se de sequências muito divergentes desse subtipo. Contudo, por uma questão de cautela, não foram consideradas como tal, apesar das suas posições na árvore filogenética sugerirem o mesmo (Figura 18). Concluindo, não foram observados indícios de recombinação genética nas sequências da região codificante da PR analisadas.

Na região da RT, após a análise de *bootscanning*, as sequências AO\_RT534, AO\_RT507, AO\_RT497, AO\_RT519 e AO\_RT491 foram identificadas como do subtipo F, mais especificamente do sub-subtipo F1, e as sequências AO\_RT380, AO\_RT512, AO\_RT6 e AO\_RT478 foram classificadas como do subtipo C (Figura 19 B). A sequência AO\_RT488 revelou tratar-se do subtipo D, sendo a sequência AO\_RT494 do subtipo G (Figura 19 B). Já as sequências AO\_RT446, AO\_RT480, AO\_RT527 e AO\_RT533 foram consideradas não genotipáveis (U) com base na análise de *bootscanning* (sendo a sua análise *SimPlot* igualmente inconclusiva – dados não apresentados), apesar de as sequências AO\_RT446 e AO\_RT527 apresentarem, nos respectivos gráficos de *bootscan*, homologia parcial com o subtipo D e o subtipo G, respectivamente, podendo eventualmente tratar-se de sequências muito divergentes desses subtipos. Contudo, e tendo em conta as suas posições na árvore filogenética, não foram assim consideradas (Figura 19). A sequência AO\_RT480 poderia também ser considerada como uma possível forma recombinante entre os subtipos G e H, com base na sua análise de *bootscanning*. No entanto, não foi classificada como tal, pois o respectivo gráfico de *bootscan* não permite tirar essa conclusão de forma inequívoca (Figura 19 B). Desse modo, tal como para a PR, não foram observadas formas recombinantes evidentes nas sequências da região codificante da RT analisadas.

No que se refere à pesquisa de fenómenos de hipermutação por ocorrência de hiperadenilação (G→A) resultante da actividade das APOBEC, a análise dos relatórios interpretativos emitidos pelo programa *HIVdb* para cada uma das sequências em estudo (cuja análise detalhada se encontra nas secções seguintes) – e não apenas das que apresentavam posicionamentos ambíguos nas árvores filogenéticas – permitiu verificar a presença de mutações deste tipo em sete sequências da região codificante da PR e três da região codificante da RT. No caso da PR, foi encontrado um total de nove mutações

APOBEC na sequência AO\_PR525; seis mutações nas sequências AO\_PR6 e AO\_PR484; e quatro, três, duas e uma mutação nas sequências AO\_PR503, AO\_PR474, AO\_PR497 e AO\_PR533, respectivamente. No caso da RT, foram detectadas oito mutações APOBEC na sequência AO\_RT6, sete mutações na sequência AO\_RT497 e uma única mutação na sequência AO\_RT533. Digno de realce, o elevado número total de mutações de hiperadenilação encontrado nas amostras AO\_6 e AO\_497 (14 e 9, respectivamente), quando consideradas as duas regiões do genoma viral em análise. De modo a verificar a existência de algum tipo de associação entre a presença destas mutações e a impossibilidade de genotipar as sequências em estudo, para ambas as regiões codificantes (PR e RT), foi realizado um teste de Fisher bilateral aplicado à análise de tabelas de contingência 2×2 com nível de significância de 1%, cujo resultado indicou não haver qualquer associação entre as duas variáveis ( $p=0,370$  e  $p=0,214$ , respectivamente, para PR e RT). Foi ainda realizada uma análise análoga à anterior, mas desta vez para verificar a possível existência de associação entre a ocorrência de fenómenos de hipermutação e os genótipos encontrados, recorrendo ao teste  $\chi^2$  aplicado à análise de tabelas  $n \times 2$  com um nível de significância de 1%, sendo que não foi encontrada qualquer associação entre genótipo viral e hipermutação nas sequências de ambas as regiões estudadas ( $p=0,235$ , no caso da PR, e  $p=0,823$ , no caso da RT).

Concluída a análise filogenética e a pesquisa de eventos de recombinação genética ou de hipermutação por ocorrência de hiperadenilação (G→A), verificou-se que os genótipos mais frequentes, para ambas as regiões estudadas, foram o subtipo C (40,0% das sequências PR e 35,8% das sequências RT) e o sub-subtipo F1 (18,0% das sequências PR e 18,9% das sequências RT), sendo classificadas como não genotipáveis (U) três sequências da região codificante da PR e quatro da região RT. Observou-se ainda que não foram identificadas sequências do subtipo B no conjunto das sequências RT, nem sequências dos subtipos D e H no grupo das sequências PR, tendo-se identificado pelo menos uma sequência dos subtipos referidos no conjunto das sequências codificantes da PR e da RT, respectivamente. Na Tabela 9 encontram-se resumidos os resultados finais da classificação genotípica de todas as sequências em estudo ( $n=103$ ).

**Tabela 9: Resultados finais da genotipagem para as regiões codificantes da protease e da transcriptase reversa após conclusão da respectiva análise filogenética e da pesquisa de eventos de recombinação genética e/ou de hipermutação por ocorrência de hiperadenilação.**

| <b>Genótipo</b> | <b>Protease (n=50)</b> | <b>Transcriptase reversa (n=53)</b> |
|-----------------|------------------------|-------------------------------------|
| A               | 5 (10,0%)              | 2 (3,8%)                            |
| B               | 1 (2,0%)               | –                                   |
| C               | 20 (40,0%)             | 19 (35,8%)                          |
| D               | –                      | 1 (1,9%)                            |
| F (F1)          | 9 (18,0%)              | 10 (18,9%)                          |
| G               | 5 (10,0%)              | 5 (9,4%)                            |
| CRF02_AG        | 6 (12,0%)              | 9 (17,0%)                           |
| H               | –                      | 2 (3,8%)                            |
| J               | 1 (2,0%)               | 1 (1,9%)                            |
| U               | 3 (6,0%)               | 4 (7,5%)                            |

Comparando os resultados de genotipagem obtidos para ambas as regiões em estudo (Anexo 9), verificou-se um elevado nível de concordância entre os genótipos das regiões da PR e da RT nas amostras para as quais foi possível amplificar as duas (47/56), das quais 87,2% (41/47) apresentavam genótipos concordantes, sendo importante referir que uma dessas amostras (AO\_480) foi classificada como não genotipável (U) para ambas as regiões. Foram encontradas duas amostras com genótipo possivelmente recombinante entre as regiões da PR e da RT, nomeadamente as amostras AO\_461 e AO\_538, tendo-se verificado para ambas que o seu genótipo combinado é do tipo A/H. Observou-se também que uma das amostras (AO\_532) apresentava um genótipo combinado das duas regiões estudadas do tipo G/G (CRF02\_AG), o que pode ter diversas explicações, entre as quais tratar-se de um vírus recombinante de segunda geração. Constatou-se ainda que três das amostras para as quais foi possível amplificar ambas as regiões em estudo (AO\_488, AO\_527 e AO\_533) apresentavam uma das regiões estudadas classificada como não genotipável (U) e, finalmente, que não foi possível atribuir um sub-subtipo a nenhuma das sequências genotipadas como subtipo A (AO\_455 e AO\_467). É importante referir que os genótipos combinados das 47 amostras para as quais foram obtidas sequências consenso de ambas as regiões codificantes em estudo não incluem uma pequena região de cerca de 50 nucleótidos entre as regiões codificantes da PR e da RT analisadas, uma vez que a mesma não foi amplificada por PCR. Não foram ainda

consideradas nesta análise as amostras para as quais apenas foi possível amplificar uma das regiões codificantes (n=9).

De um modo geral, constatou-se que no conjunto das 47 amostras analisadas para o genótipo combinado (Anexo 9), os genótipos predominantes são o subtipo C e o subsubtipo F1, correspondendo a 40,4% (19/47) e 19,1% (9/47) das amostras, respectivamente, sendo a forma recombinante CRF02\_AG o terceiro genótipo mais frequente (12,8%, 6/47). Verificou-se também que nenhuma amostra possuía um genótipo combinado puro do subtipo H e que as sequências de três amostras (6,4%) apresentavam um genótipo combinado do tipo recombinante único (URF, A/H, para as amostras AO\_461 e AO\_538, e G/CRF02\_AG, para a amostra AO\_532, sendo que o genótipo desta última foi considerado recombinante por ser a hipótese mais provável, embora não a única). Para as amostras AO\_488, AO\_527 e AO\_533 considerou-se que o respectivo genótipo combinado correspondia ao genótipo da região codificante que foi possível classificar. No caso das duas primeiras, as sequências da região classificada como não genotipável (U) apresentam nos gráficos de *bootscan* homologia parcial com os genótipos das sequências da respectiva região genotipada (Figuras 18 B e 19 B), sendo que no caso da última a decisão teve por base o posicionamento da sequência AO\_RT533 (U) na respectiva árvore filogenética, pois encontra-se filogeneticamente próxima das sequências de referência do subtipo B (Figura 19 A). No cômputo geral, estas podem constituir sequências divergentes dos respectivos genótipos considerados. Na Figura 20 encontram-se resumidos os resultados da análise de genótipos combinados das 47 amostras para as quais foi possível amplificar ambas as regiões codificantes em estudo.

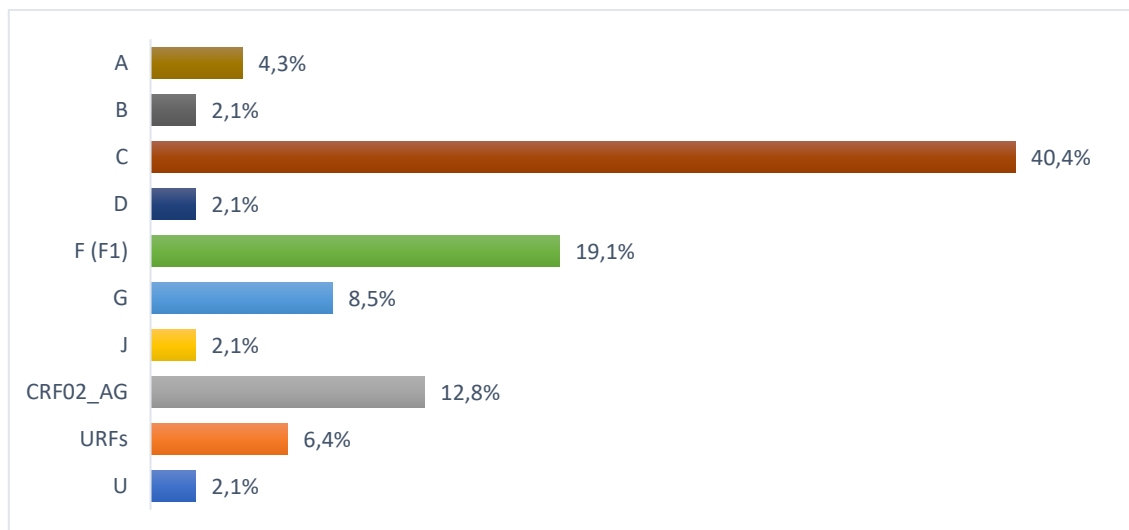


Figura 20: Representação gráfica da proporção relativa dos genótipos do HIV-1 obtidos a partir da análise filogenética das sequências nucleotídicas de um total de 47 amostras (PR+RT) [o genótipo “URF” corresponde a sequências cujo genótipo PR é diferente do genótipo RT, não correspondendo a nenhuma CRF já descrita].

### 3.4 Análise bioinformática das sequências aminoacídicas

#### 3.4.1 Pesquisa de mutações associadas a resistência a anti-retrovirais

Conforme descrito na secção 2.8.1, a análise da presença de mutações associadas a diminuição da susceptibilidade aos ARVs inibidores da PR e da RT (e detecção de outros polimorfismos genéticos) nas sequências correspondentes às regiões codificantes de cada uma destas enzimas foi realizada utilizando o programa *HIVdb – Genotypic Resistance Interpretation Algorithm* da *Stanford University HIV Drug Resistance Database* (versão 8.5). Com base na comparação com uma sequência consenso de referência do subtipo B, este *software* permite obter para cada uma das sequências em estudo um relatório interpretativo com as mutações associadas a resistência aos ARVs e outros polimorfismos genéticos, no qual é também apresentado o respectivo perfil fenotípico de resistência mais provável para cada um dos ARVs considerados, com base no método descrito na secção 2.8.1.

As sequências em estudo correspondiam à enzima PR completa, num total de 99 aminoácidos, e à enzima RT incompleta, num total de 199 dos 560 aminoácidos que constituem a sua subunidade maior (p66) (posições 29-227, verificando-se a ausência de 28 aminoácidos da extremidade N-terminal e dos 333 aminoácidos que compõem a extremidade carboxílica terminal). Embora parcial, a região da RT analisada neste estudo é suficiente para cumprir o objectivo definido, estando de acordo com as recomendações

internacionais mais recentes que aconselham a inclusão “da primeira metade da RT (pelo menos até ao [aminoácido] 215) [...] para todos os testes [de resistência] genotípicos” (178). Em quatro sequências da região codificante da RT (AO\_RT7, AO\_RT119, AO\_RT467 e AO\_RT507) estavam também ausentes 9, 5 e 3 aminoácidos da extremidade 3’, respectivamente. Não foram observadas mutações *frameshift* (resultantes da inserção ou deleção de nucleótidos) em qualquer das sequências em estudo. No entanto, verificou-se a presença de codões de terminação e/ou mutações associadas a actividade APOBEC, inclusive, nalguns casos, em posições associadas a mutações de resistência, em algumas sequências em estudo (como se viu anteriormente, sete no caso da PR: AO\_PR6, AO\_PR474, AO\_PR484, AO\_PR497, AO\_PR503, AO\_PR525 e AO\_PR533; e três no caso da RT: AO\_RT6, AO\_RT497 e AO\_RT533), além de algumas mutações pouco comuns, sem impacto na susceptibilidade aos ARVs (Anexo 10).

#### 3.4.1.1 Região codificante da protease

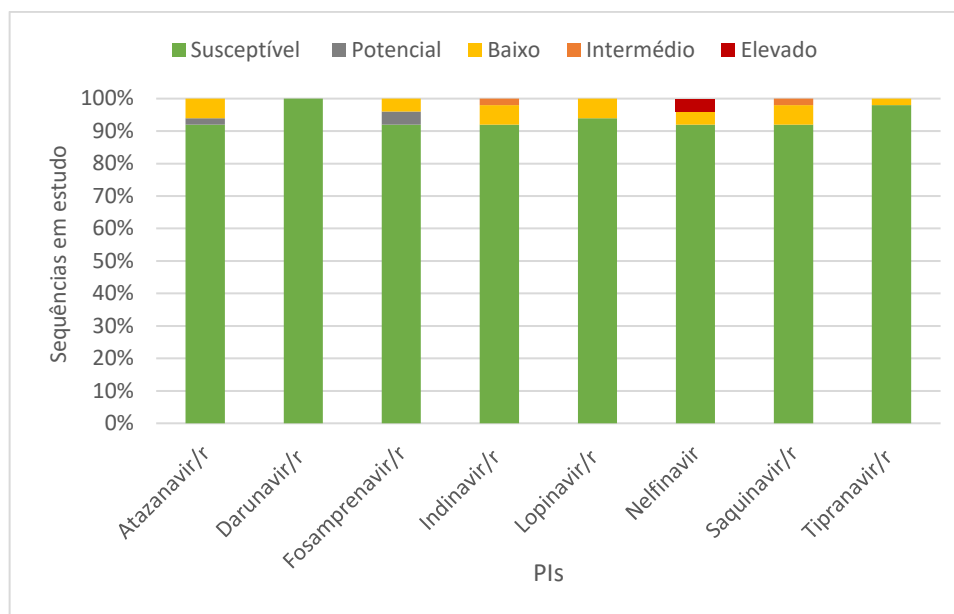
No conjunto das sequências da região codificante da PR foram encontradas quatro mutações principais (D30N, M46I, I54V e L90M) e uma mutação acessória (G73S), num total de quatro sequências (AO\_PR474, AO\_PR494, AO\_PR525 e AO\_PR542), sendo que todas estas mutações têm importância para a vigilância epidemiológica da resistência transmitida aos ARVs no HIV-1 (179,180) (Figura 21, Anexo 10). Cada uma das sequências mencionadas apresentava somente uma destas mutações, com excepção da sequência AO\_PR525 que apresentava um total de três mutações, duas principais (D30N e M46I) e uma acessória (G73S). Foram também encontradas outras mutações associadas a resistência somente na presença de outras mutações, nomeadamente L10I/V, K20I, A71I, T74S e V82I. No entanto, somente a L10I, a K20I e a V82I foram encontradas associadas a mutações principais em três sequências – todas associadas à mutação I54V na sequência AO\_PR542, sendo que as duas últimas surgem também associadas à mutação L90M na sequência AO\_PR494 e a mutação K20I está associada às mutações D30N e M46I na sequência AO\_PR525.

|     |    |    |    |    |    |    |    |    |    |    |
|-----|----|----|----|----|----|----|----|----|----|----|
| PIs | L  | K  | D  | M  | I  | A  | G  | T  | V  | L  |
|     | 10 | 20 | 30 | 46 | 54 | 71 | 73 | 74 | 82 | 90 |
|     | I  | I  | N  | I  | V  | I  | S  | S  | I  | M  |
|     | V  |    |    |    |    |    |    |    |    |    |

**Figura 21:** Esquema representativo de todas as mutações associadas a resistência encontradas nas sequências em estudo da região codificante da protease, incluindo principais acessórias e outras mutações. As letras em cima representam o aminoácido presente na sequência consenso do subtipo B (susceptível), o número identifica a respectiva posição na mesma e as letras em baixo representam o aminoácido introduzido decorrente da mutação.

Verificou-se que duas das sequências que apresentavam mutações associadas a diminuição da susceptibilidade aos PIs, AO\_PR474 e AO\_PR525, possuíam também várias mutações associadas à actividade APOBEC (três e nove mutações, respectivamente, uma das quais, G48E, localizada numa posição relacionada com resistência aos PIs na sequência AO\_PR525), incluindo a introdução de um codão de terminação (W42\*) em ambas as sequências e outras mutações não usuais (o relatório interpretativo das mesmas indica também que as mutações associadas a resistência aos PIs poderão reflectir actividade APOBEC). Entre as mutações principais e acessórias, a mais frequente foi G73S (4,0%), sendo que, no referente a outras mutações, a mutação K20I surgiu em 28,0% das sequências e as mutações L10V e L10I, em 22,0% e 14,0%, respectivamente.

Em relação ao perfil fenotípico putativo de resistência aos diferentes PIs, a esmagadora maioria das sequências analisadas apresenta um perfil fenotípico de total susceptibilidade (92,0%), enquanto as quatro sequências que possuem mutações associadas a resistência aos mesmos apresentavam um perfil fenotípico de resistência concordante com as respectivas mutações (ou combinação das mesmas, no caso da sequência AO\_PR525). De modo geral, verificou-se que todas as sequências em estudo apresentavam susceptibilidade ao darunavir/r (em que “/r” indica a utilização do ritonavir como potenciador farmacocinético) e que somente uma sequência (AO\_PR542) apresentava algum tipo de resistência, ainda que de nível baixo, ao tipranavir/r. No que se refere aos restantes PIs, observou-se que as sequências que possuem mutações associadas a resistência apresentavam níveis de resistência entre o potencial de nível baixo e o intermédio, sendo que apenas se verificaram níveis de resistência elevados em duas delas (AO\_PR494 e AO\_PR525) relativamente ao nelfinavir (Figura 22, Anexo 11).



**Figura 22: Representação gráfica dos níveis relativos de resistência (em porcentagem) para os diferentes inibidores da protease, referentes às sequências em estudo da respectiva região codificante.**

Foi ainda realizada a pesquisa de assinaturas genéticas para cada uma das mutações associadas a resistência aos PIs encontradas (principais, acessórias e outras) com o propósito de verificar a possível existência de associação com os diferentes genótipos observados. Para tal, recorreu-se ao teste de Fisher bilateral aplicado à análise de tabelas de contingência 2×2 com um nível de significância de 1% ( $\alpha=0,01$ ) e os resultados obtidos revelaram correlações positivas (i.e., associação de determinada mutação a um genótipo específico) entre a mutação K20I e os genótipos G e CRF02\_AG, a mutação L10V e o sub-subtipo F1 e a mutação V82I e o subtipo G (Anexo 12). Para as restantes mutações, incluindo todas as mutações principais e a mutação acessória encontradas nas sequências em estudo, não foi observada qualquer associação com os diferentes genótipos (dados não apresentados).

#### 3.4.1.2 Região codificante da transcriptase reversa

No conjunto das sequências da região codificante da RT foram encontradas catorze mutações associadas a diminuição de susceptibilidade aos NRTIs num total de doze sequências (Anexo 10), sendo que apenas duas (E44D e K70T) não constituem mutações com importância para a vigilância epidemiológica da resistência transmitida aos ARVs no HIV-1, e doze mutações associadas a diminuição de susceptibilidade aos NNRTIs num total de dezassete sequências, sendo que sete dessas mutações (K101E, K103N, V106A/M, Y181C, Y188L e G190A) têm importância para a vigilância

epidemiológica (179,180) (Figura 23, Anexo 10). No geral, verificou-se que 18 das 53 sequências em estudo possuem uma ou mais mutações associadas a diminuição de susceptibilidade aos NRTIs e/ou NNRTIs e que, em todas elas, pelo menos uma dessas mutações tem importância para a vigilância epidemiológica da resistência transmitida aos ARVs. Foram também encontradas outras mutações associadas a resistência mas na presença de outras mutações, nomeadamente, S68G, T69S, V90I, K101N, V106I, V118I e V179I, as quais (com excepção da T69S) surgem, em diferentes proporções, associadas a mutações com impacto na resistência aos NRTIs e/ou NNRTIs em uma ou mais sequências (sete no total – AO\_RT92, AO\_RT461, AO\_RT478, AO\_RT497, AO\_RT507, AO\_RT533 e AO\_RT534).

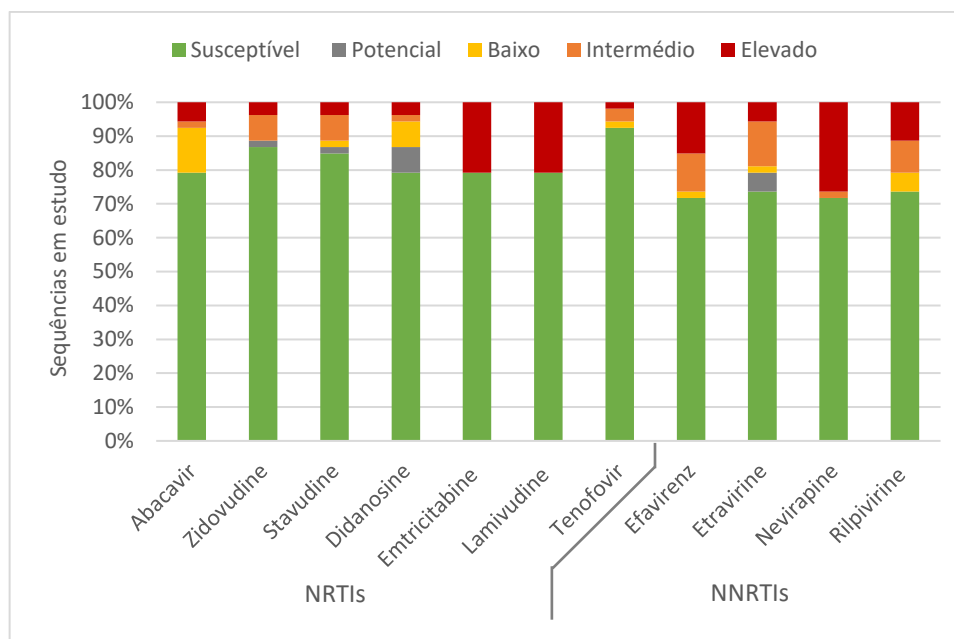
|        |    |    |             |     |             |        |     |     |     |        |     |        |        |
|--------|----|----|-------------|-----|-------------|--------|-----|-----|-----|--------|-----|--------|--------|
| NRTIs  | M  | E  | D           | S   | T           | K      | V   | Y   | V   | M      | L   | T      | K      |
|        | 41 | 44 | 67          | 68  | 69          | 70     | 75  | 115 | 118 | 184    | 210 | 215    | 219    |
|        | L  | D  | N           | G   | S           | E<br>T | M   | F   | I   | I<br>V | W   | F<br>Y | N<br>R |
| NNRTIs | V  | A  | K           | K   | V           | V      | E   | V   | Y   | Y      | G   | H      |        |
|        | 90 | 98 | 101         | 103 | 106         | 108    | 138 | 179 | 181 | 188    | 190 | 221    |        |
|        | I  | G  | E<br>H<br>N | N   | A<br>I<br>M | I      | A   | I   | C   | L      | A   | Y      |        |

**Figura 23:** Esquema representativo de todas as mutações associadas a resistência encontradas nas sequências em estudo da região codificante da transcriptase reversa, incluindo mutações associadas a resistência apenas na presença de outras mutações. As letras em cima representam o aminoácido presente na sequência consenso do subtipo B (susceptível), o número identifica a respectiva posição na mesma e as letras em baixo representam o aminoácido introduzido decorrente da mutação.

Verificou-se que três das sequências que apresentavam mutações associadas a diminuição da susceptibilidade aos NRTIs e/ou NNRTIs, AO\_RT6, AO\_RT497 e AO\_RT533, mostravam também várias mutações associadas à actividade APOBEC (oito, sete e uma mutações, respectivamente), sendo que na primeira, uma dessas mutações (G190R) surge localizada numa posição relacionada com resistência aos NNRTIs, mas não está associada a resistência aos mesmos, tratando-se de uma mutação altamente incomum nesta posição. Na sequência AO\_RT497 também foi encontrada uma mutação extremamente incomum, não APOBEC (Y188M), numa posição relacionada com resistência aos NNRTIs, embora não esteja associada a resistência aos mesmos. Nas mesmas duas sequências verificou-se ainda a introdução de vários codões de terminação [três na sequência RT6 (W88\*, W153\*, W212\*), e cinco na sequência RT497 (W71\*,

W88\*, W153\*, L210\*, W212\*)] e de outras mutações não usuais. O relatório interpretativo obtido no programa *HIVdb* indica também que uma mutação associada a resistência aos NRTIs, de cada uma das sequências, poderá reflectir actividade APOBEC (M184I, na sequência AO\_RT6, e D67N, na sequência AO\_RT497). Das mutações associadas à diminuição da susceptibilidade aos NRTIs a mais frequente foi M184V (17,0%), seguida de T215Y (9,4%), enquanto no caso das mutações associadas à diminuição da susceptibilidade aos NNRTIs, as mais frequentes foram Y181C (18,9%) e V108I (9,4%). No que se refere a outras mutações, V179I foi a mais frequente, surgindo em 7,5% das sequências.

Em relação ao perfil fenotípico putativo de resistência aos diferentes NRTIs e NNRTIs, a maioria das sequências analisadas apresenta um perfil fenotípico de total susceptibilidade (66,0%), enquanto as restantes sequências que possuem mutações associadas a resistência aos mesmos apresentam um perfil fenotípico de resistência concordante com as respectivas mutações ou combinação das mesmas. Digno de nota que um nível de resistência elevado a pelo menos um dos RTIs foi associado a 17 sequências diferentes. Uma das sequências em questão (AO\_RT520) apresenta níveis de resistência elevados a todos os NRTIs e a dois dos NNRTIs analisados (nevirapine e rilpivirine), possuindo um nível de resistência intermédio aos restantes (efavirenz e etravirine), tratando-se da sequência com o perfil fenotípico de resistência de nível elevado mais abrangente no conjunto das sequências analisadas. Não há, neste caso, qualquer associação à actividade mutagénica das enzimas APOBEC do hospedeiro. De modo geral, verificou-se que as sequências com mutações associadas a resistência aos RTIs apresentavam níveis de resistência entre o potencial de nível baixo e o elevado, para os diferentes fármacos, sendo que o nevirapine é o ARV com maior número de sequências com nível de resistência elevado (14 em 18 sequências, correspondendo a 26,4% do total das sequências da região codificante da RT analisadas) e que apenas quatro sequências (AO\_RT92, AO\_RT497, AO\_RT520 e AO\_RT542) apresentavam algum tipo de resistência ao tenofovir, entre os níveis baixo e elevado (Figura 24, Anexo 11).



**Figura 24:** Representação gráfica dos níveis relativos de resistência (em percentagem) para os diferentes inibidores da transcriptase reversa, em percentagem, nucleosídicos e não nucleosídicos, referentes às sequências em estudo da respectiva região codificante.

Foi também realizada a pesquisa de assinaturas genéticas para cada uma das mutações associadas a resistência aos diferentes RTIs encontradas, com o propósito de verificar a possível existência de associação com os diferentes genótipos observados. Recorrendo ao teste de Fisher bilateral aplicado à análise de tabelas de contingência 2×2 com um nível de significância de 1% ( $\alpha=0,01$ ), verificou-se a existência de apenas uma correlação positiva entre a mutação H221Y (associada a diminuição da susceptibilidade aos NNRTIs) e o subtipo G (Anexo 12). Para as restantes mutações descritas, não foi observada qualquer associação significativa entre as mesmas e os diferentes genótipos (dados não apresentados).

### 4 Discussão e conclusão

---

A elevada diversidade genética apresentada pelo HIV-1 tem-se revelado um dos maiores obstáculos ao controlo da pandemia. Nas últimas décadas, assistiu-se a uma complexidade crescente no padrão de circulação de formas genéticas virais à escala global, nomeadamente, através do aparecimento de novos sub-subtipos, CRFs e URFs, sendo já perto de 100 as CRFs descritas e cada vez maior o número de URFs relatadas nos diferentes estudos realizados. Esta vasta diversidade deve-se primeiramente a características do próprio vírus e respectivo ciclo replicativo, nomeadamente, a elevada taxa de mutação durante a retrotranscrição do genoma viral, resultante da ausência de actividade *proofreading* da RT, a elevada taxa de replicação viral *in vivo*, relacionada com o curto tempo médio de semi-vida dos viriões e a ocorrência de recombinação genética (90,93). Além destes aspectos intrínsecos ao vírus, existem ainda outros factores que influenciam e contribuem para a diversidade genética, em particular, elevadas pressões selectivas por parte do sistema imunitário do indivíduo infectado, a acção de sistemas enzimáticos envolvidos em mecanismos de imunidade inata antiviral (como a APOBEC3G) e a terapia anti-retroviral, mas também a migração das populações ao longo das décadas (181). É por isso importante manter a vigilância epidemiológica sobre as estirpes virais circulantes, além de outras características da pandemia, uma vez que a sua crescente diversidade tem um impacto considerável em vários aspectos, sendo um dos mais relevantes o tratamento da infecção.

Apesar do sucesso da HAART, que permitiu diminuições consideráveis na mortalidade e morbilidade associadas à infecção e uma melhoria generalizada da qualidade de vida dos indivíduos infectados, tornando esta infecção numa patologia crónica passível de ser controlada, a emergência de mutações de resistência aos ARVs utilizados nos planos terapêuticos acarreta consequências potencialmente negativas, sendo o cenário de falência terapêutica alvo de preocupação do ponto de vista clínico e da saúde pública (91,128). Por esse motivo, a pesquisa de mutações associadas a diminuição da susceptibilidade aos ARVs tornou-se uma ferramenta de monitorização fundamental no controlo da infecção pelo HIV que permite, não só, ajustar o tratamento a cada indivíduo, especialmente em casos de resistência primária aos ARVs, mas também fornecer informações para a pesquisa e desenvolvimento de novos fármacos (91,182).

As questões que relacionam a diversidade genética do HIV-1 e a terapia anti-retroviral são particularmente relevantes nos países da África Central, em especial na RDC e países vizinhos, uma vez que é nestes países que é encontrada a maior diversidade de genótipos em circulação. Por outro lado, os testes de susceptibilidade aos ARVs não são aqui facilmente aplicados, com a agravante de os regimes terapêuticos normalmente utilizados apresentarem uma baixa barreira genética para aquisição de resistência, criando desse modo condições favoráveis à selecção e circulação de estirpes virais resistentes (128,178,181–183). Em Angola, país fronteiro com a RDC, os estudos realizados revelaram um provável aumento da complexidade da epidemia do HIV-1, que se caracteriza pelo aparente incremento da diversidade das estirpes virais em circulação e pela crescente frequência de mutações de resistência aos ARVs, em particular aos RTIs, constatando-se ainda a existência de casos de resistência transmitida (ou resistência primária) (154,155,157–162). Contudo, aquilo que é conhecido sobre a epidemia em Angola é baseado em estudos com um número limitado de indivíduos infectados e coberturas geográficas restritas, dado que a maioria das amostras é proveniente da cidade de Luanda, havendo por isso um potencial enviesamento da informação devido à escassez de dados sobre outras regiões do país. Pretendeu-se então com este estudo caracterizar filogeneticamente as estirpes do HIV-1 em circulação na província de Benguela e estimar o respectivo grau de resistência primária aos ARVs inibidores da PR e da RT, com o intuito de colmatar a falta de informação encontrada e contribuir para um melhor conhecimento acerca das características epidemiológicas moleculares da epidemia do HIV-1 em Angola.

Considerando os objectivos definidos, foi realizada a amplificação por dupla *nested* PCR das regiões codificantes da PR e da RT do gene *pol* de um total de 111 amostras de sangue periférico de indivíduos seropositivos para HIV-1 que se encontravam internados ou eram utentes de consultas externas/banco de urgência do HGB, todos de nacionalidade angolana e residentes na província de Benguela, sendo importante referir que a totalidade dos indivíduos incluídos neste estudo nunca foi submetida a terapia anti-retroviral. Relativamente aos resultados da amplificação, mesmo com a aplicação de protocolos alternativos com temperaturas de hibridação inferiores às do protocolo base em diferentes combinações, apenas 45,9% (51/111) das amostras foram positivas para a PR e 48,6% (54/111) para a RT (Figura 16), considerando-se que estes

resultados estarão, muito provavelmente, relacionados com o suporte para o qual foram colhidas e armazenadas as amostras sanguíneas, designadamente, os cartões *Whatman Indicating FTA<sup>®</sup> Elute Micro Cards* (GE Healthcare, Reino Unido). De facto, só foi possível obter resultados de amplificação positivos para qualquer uma das regiões em estudo quando foi abandonado o protocolo recomendado pelo fabricante para a extracção do DNA proviral (Anexo 5) e se optou por utilizar o protocolo de preparação de discos desenvolvido no IHMT (secção 2.2). Estes resultados vão de encontro ao que está descrito nos estudos de Denniff & Spooner (2010) e Hearps *et al.* (2010), nos quais se verificou que o tipo de cartões utilizado neste estudo (*FTA<sup>®</sup> Elute*) é mais sensível a condições extremas de temperatura e humidade relativa, sendo que a integridade da matriz dos cartões e da própria mancha de sangue ficam comprometidas (184). Nestas condições, a recuperação de ácidos nucleicos é menos eficiente, em comparação com cartões sem a tecnologia *Elute*, sendo necessários protocolos e reagentes alternativos para aumentar a eficiência da recuperação do material genético para análise (185). Assim, concluiu-se que a relativa baixa taxa de sucesso do protocolo de amplificação aplicado (47,3%, quando consideradas ambas as regiões do genoma em estudo) se deverá, muito provavelmente, ao tipo de cartões utilizados, apesar de terem sido controladas as condições do seu armazenamento (mantidos à TA, recorrendo-se a sílica gel para o controlo da humidade relativa). No entanto, dadas as características intrínsecas do HIV-1, e a região geográfica em causa, a elevada diversidade genética que era expectável encontrar nas amostras em estudo pode ter também afectado significativamente outros aspectos inerentes à reacção de amplificação, nomeadamente, a hibridação dos *primers* com as regiões alvo, apesar destas últimas serem regiões do genoma do HIV-1 relativamente bem conservadas.

Os produtos de amplificação com tamanhos esperados correspondentes às regiões da PR e da RT (460 e 650 pb, respectivamente) foram purificados e enviados para sequenciação nucleotídica em ambos os sentidos. Ao tomar esta opção pretendeu-se diminuir a probabilidade de ocorrência de erros de sequência associados à própria técnica e aumentar a fiabilidade dos resultados obtidos, pois um dos objectivos deste trabalho consiste na pesquisa de mutações de resistência aos ARVs. Do total de 105 fragmentos purificados, foram obtidas 50 sequências consenso PR e 53 sequências consenso RT após a edição das sequências obtidas com base nos respectivos cromatogramas, sendo que para dois dos fragmentos sequenciados (1,9%) não foi possível obter sequências consenso,

pelas razões já enumeradas na secção 3.3.1. Uma vez confirmada a identidade das 103 sequências em estudo por pesquisa de semelhança local entre sequências recorrendo ao programa *BLAST*, foi realizada a sua genotipagem automatizada utilizando o *software REGA HIV-1 Subtyping Tool* (versão 3.0), por forma a adquirir um conhecimento preliminar acerca da diversidade genética presente na amostra. A escolha deste programa deve-se ao facto de o mesmo apresentar os resultados com base em análise filogenética e não na pesquisa de simples similaridade entre sequências, como outros *softwares* disponíveis *online*, sendo a sua análise mais significativa, apesar de possuir algumas limitações (186,187). Esta análise preliminar permitiu, desde logo, verificar o predomínio dos genótipos C e F1, em ambas as regiões codificantes (Tabela 6). No entanto, através desta análise, não foram atribuídos genótipos a um número considerável de sequências, indicando a possível existência de sequências recombinantes ou geneticamente muito divergentes.

Para uma caracterização mais consistente das sequências nucleotídicas em estudo, foi efectuada a sua genotipagem definitiva através da construção de árvores filogenéticas por métodos não automatizados, cuja vantagem é permitir a selecção das sequências de referência e dos métodos e algoritmos utilizados aumentando a fiabilidade dos resultados. A construção das árvores filogenéticas para cada uma das regiões codificantes do gene *pol* em estudo (PR e RT) foi realizada utilizando o *software MEGA* versão X, com base no método de *Neighbor-Joining* (176) e recorrendo ao modelo de dois parâmetros de Kimura (177) para o cálculo das matrizes de distâncias genéticas, tendo-se considerado os agrupamentos formados como filogeneticamente consistentes quando suportados por valores mínimos de *bootstrap* de 70%.

Numa análise inicial das árvores filogenéticas construídas para cada uma das regiões codificantes, verificou-se que ambas apresentavam, de um modo geral, topologias bem definidas, embora mais consistente no caso da RT. Por seu lado, quase todos os agrupamentos da árvore filogenética para a PR são suportados por valores de *bootstrap* inferiores a 70% (Figura 18 A), limitando a robustez da inferência genética realizada e tornando-a menos fiável. Estes resultados deverão estar relacionados com o menor tamanho da região da PR analisada (307 nucleótidos), quando comparado com o da RT (532 nucleótidos), e com o facto da região codificante da PR ser relativamente conversada entre subtipos (188), resultando na obtenção de um baixo sinal filogenético, dificultando

a sua análise. No caso da região RT, a maioria dos agrupamentos formados eram suportados por valores de *bootstrap* iguais ou superiores a 70% (Figura 19 A), pelo que o grau de certeza da inferência filogenética e, como consequência, da classificação genotípica com base nesta árvore filogenética é bastante mais elevado. Tendo estes aspectos em consideração, procedeu-se a uma análise detalhada dos resultados obtidos e foi possível classificar genotipicamente a maioria das sequências de ambas as regiões codificantes, apenas com base na análise das árvores filogenéticas construídas, inclusive no caso da região PR, para um total de 35 sequências (70,0%) da PR e 38 sequências (71,7%) da RT (Tabela 8).

Os agrupamentos de sequências do subtipo C são os únicos grupos monofiléticos formados contendo sequências em estudo suportados, em ambas as árvores filogenéticas, por valores de *bootstrap* significativos – 76% no caso da PR e 70% no caso da RT. Verificou-se que as sequências em estudo de ambas as regiões codificantes genotipadas como C com base no respectivo posicionamento nas árvores filogenéticas (17 sequências PR e 15 sequências RT) apresentavam relações filogenéticas próximas com sequências de referência com origem em Angola, Zâmbia, África do Sul e Brasil. No caso da PR, foram ainda encontradas relações filogenéticas com sequências de referência originárias do Burundi, Botswana, Nigéria e Tanzânia (Figura 18 A), e, para a RT, com sequências de referência oriundas da China e Malawi (Figura 19 A). Estas observações vão de encontro aos resultados do estudo de Afonso *et al.* (2012), demonstrando que as estirpes do subtipo C circulantes em Angola resultaram de múltiplas introduções de vírus deste subtipo, oriundos principalmente de países do sul do continente africano, mas também da África Central e Oriental, entre o final dos anos 70 do século passado e meados dos anos 2000, em consequência de fenómenos migratórios populacionais (153). O facto de ambas as árvores filogenéticas incluírem uma sequência de referência proveniente do Brasil, com as quais uma das sequências em estudo de cada região codificante apresenta relações filogenéticas próximas (AO\_PR490 e AO\_RT442), pode ser explicado com base na introdução no Brasil do subtipo C, no início da década de 1980, a partir da disseminação de uma única estirpe fundadora filogeneticamente muito próxima de estirpes originárias do Burundi (189). Este país é igualmente apontado como a origem mais provável de uma das linhagens do subtipo C circulantes em Angola (153), verificando-se inclusive a elevada proximidade filogenética da sequência AO\_PR490 com as sequências de

referência oriundas de ambos os países (Figura 18 A). Por outro lado, a proximidade filogenética entre as sequências AO\_RT92 e AO\_RT484 e uma sequência de referência com origem na China [onde existe uma prevalência considerável do subtipo C e respectivas formas recombinantes (190)] poderá explicar-se pelas relações filogenéticas próximas entre vírus deste subtipo oriundos da China, Índia e África do Sul (191), uma vez que o HIV-1 do subtipo C na Índia tem origem na África do Sul e o primeiro surto epidémico do HIV-1 na China resultou da introdução de duas estirpes indianas desse subtipo na província de Yunnan, cerca de 1993 (192). Dado que a África do Sul foi igualmente apontada como uma das origens mais prováveis de cinco linhagens do HIV-1 do subtipo C angolanas (153), compreende-se as relações filogenéticas observadas na árvore da RT, podendo a sequência de referência utilizada com origem na China corresponder a um vírus geneticamente próximo daqueles que foram directa ou indirectamente introduzidos em ambos os países a partir da África do Sul. Estes dados permitem fortalecer a robustez da inferência filogenética realizada para a classificação das sequências deste subtipo, reforçando também os achados de estudos anteriores sobre as relações filogenéticas próximas entre as estirpes circulantes em Angola e noutros países africanos e Brasil.

Os agrupamentos do sub-subtipo F1 de cada árvore filogenética incluem maioritariamente sequências de referência de Angola, com as quais as sequências em estudo de ambas as regiões codificantes apresentam relações filogenéticas muito próximas, como seria expectável. As estirpes deste genótipo circulantes em Angola terão resultado da introdução de um único vírus fundador que se disseminou pelo país, o qual deverá ser originalmente da RDC e terá chegado a Angola através de migrantes infectados no período pós-Guerra Colonial, no início de 1980 (193,194). Curiosamente, apesar de o agrupamento das sequências F1 codificantes para a RT ser suportado por um valor de *bootstrap* significativo (70%), apresenta um maior número de sequências em estudo em posições ambíguas (n=5) (Figura 19 A), dificultando a sua genotipagem, comparativamente com o agrupamento correspondente para a região codificante da PR (n=1) (Figura 18 A).

Dos grupos monofiléticos das sequências do subtipo G, nenhum era suportado por valores de *bootstrap* significativos (Figuras 18 A e 19 A). No entanto, para a RT, todos os grupos contendo sequências em estudo e sequências de referência eram

suportados por valores de *bootstrap* mínimos de 74%, permitindo a genotipagem das primeiras, com exceção da sequência AO\_RT494, que apresenta um grau de divergência genética muito elevado. Ambos os agrupamentos do subtipo G (PR e RT) incluíam sequências de referência cuja proximidade filogenética com as sequências em estudo – indicada pelo comprimento dos ramos das árvores filogenéticas – permitiu classificar algumas das sequências PR como sendo do subtipo G, apesar de apenas um dos subagrupamentos ser suportado por valores de *bootstrap* significativos (AO\_PR520, Figura 18 A). De referir que uma das sequências utilizada como referência para o genótipo G na árvore filogenética da região codificante da RT [G (URF\_A/G).KH.03-04.KHM/HM10.DQ013456] se trata, de facto, de uma URF com um genoma mosaico resultante da recombinação entre os subtipos A e G, tendo, no entanto, sido mantida na árvore filogenética, uma vez que apresenta o genótipo G na região em estudo, segundo a análise de *bootscanning* realizada (dados não apresentados).

Em relação aos agrupamentos dos genótipos CRF02\_AG e G (CRF02\_AG), das árvores filogenéticas PR e RT, respectivamente, apesar de ambos serem suportados por valores de *bootstrap* não significativos (embora alguns subagrupamentos o sejam, particularmente no caso da RT), a sua topologia permitiu classificar genotipicamente todas as sequências em estudo (de ambas as regiões), uma vez que os seus posicionamentos não se apresentam ambíguos (Figuras 18 A e 19 A). Contudo, dada a estrutura mosaico do genoma da CRF02\_AG e as regiões do genoma viral analisadas, não é possível inferir com precisão sobre a circulação desta forma recombinante na amostra estudada, em particular para a região RT. Para tal, seria necessário analisar, idealmente, o genoma viral na totalidade, pelo que apenas nos será possível com este trabalho fazer inferências relativamente à região do gene *pol* estudada (ver à frente). É de salientar que uma das sequências utilizadas como referência para a CRF02\_AG na árvore filogenética da PR é uma URF resultante da recombinação entre os subtipos A, F2 e G, apresentando um genótipo recombinante CRF02\_AG/U na região analisada, de acordo com os resultados da análise de *bootscanning* (dados não apresentados).

Para as restantes sequências em estudo cuja classificação genotípica foi realizada de modo relativamente inequívoco, a sequência AO\_PR533 foi genotipada como sendo do subtipo B dado o seu posicionamento no respectivo agrupamento, apesar deste não ser

suportado por valores de *bootstrap* significativos, tornando a sua classificação menos consistente (Figura 18 A); as sequências AO\_RT467 e AO\_RT455 foram genotipadas como sendo do subtipo A com valores de *bootstrap* significativos, não sendo possível precisar, para ambas, um sub-subtipo; a sequência AO\_RT506 foi classificada do subtipo J, tratando-se de uma inferência filogenética muito robusta (valores de *bootstrap* acima dos 90%); e as sequências AO\_RT538 e AO\_RT461 foram genotipadas com sendo do subtipo H, de modo também altamente significativo (Figura 19 A). Nestas últimas, de notar que as sequências PR correspondentes não se encontram filogeneticamente relacionadas com o subtipo H (Figura 18 A), indiciando, desde logo, a existência de formas virais recombinantes.

Nesta análise mais detalhada constatou-se também que algumas sequências em estudo (15 em cada árvore filogenética) apresentam posicionamentos ambíguos nas respectivas árvores filogenéticas, algumas delas não estando incluídas num grupo monofilético significativo, dificultando a sua classificação genotípica, ou são suportadas por ramos muito longos, sugerindo uma elevada divergência genética. Assim, na tentativa de clarificar a sua natureza, foi realizada a pesquisa de eventos de recombinação genética e de fenómenos de hipermutação, em particular a ocorrência de hiperadenilação (G→A). A pesquisa de possíveis eventos de recombinação utilizando o *software SimPlot* (versão 3.5.1) não permitiu confirmar de forma evidente a sua ocorrência nas sequências analisadas de ambas as regiões estudadas. Como descrito na secção 3.3.3, a análise de *bootscanning* permitiu, no entanto, clarificar a genotipagem de algumas sequências em estudo, embora sete delas (AO\_PR474, AO\_PR480, AO\_PR488, AO\_RT446, AO\_RT480, AO\_RT527 e AO\_RT533) tenham sido classificadas como não genotipáveis (U) (Tabela 9).

Há, no entanto, algumas observações mais específicas a destacar decorrentes da análise de *bootscanning*:

- ❖ a divergência sugerida pelo comprimento do ramo da árvore filogenética que suporta a sequência AO\_RT497 (Figura 19 A) não foi associada a qualquer fenómeno de recombinação (dados não apresentados);
- ❖ a sequência AO\_PR525 foi considerada divergente do subtipo A, dado o comprimento do ramo que a suporta na respectiva árvore filogenética e o gráfico de *bootscan* obtido (Figura 18);

- ❖ os gráficos de *bootscan* das sequências AO\_PR6 e AO\_PR545 sugerem algum grau de divergência relativamente ao subtipo C, apesar de tal não ser sugerido pela topologia da árvore filogenética ou pelos seus posicionamentos na mesma, nem estar necessariamente relacionado com o tamanho dos ramos que as suportam (Figura 18);
- ❖ a análise de *bootscanning* das sequências AO\_PR474 e AO\_PR488 revelou identidade parcial com o subtipo D, o que, tendo em conta a suas posições na respectiva árvore filogenética, poderia indicar que se tratam de sequências muito divergentes desse subtipo, mas, dado que a homologia é apenas verificada num número reduzido de nucleótidos em ambas as extremidades (Figura 18), estas foram classificadas como não genotipáveis;
- ❖ embora as sequências AO\_RT446 e AO\_RT527 apresentem, na análise de *bootscanning*, maior identidade com os subtipos D e G, respectivamente, não foram classificadas como sequências muito divergentes desses subtipos, mas sim como não genotipáveis, dadas as suas posições na árvore filogenética (Figura 19);
- ❖ a análise de *bootscanning* da sequência AO\_RT480 poderá sugerir a ocorrência de recombinação entre os subtipos G e H, embora não de forma inequívoca, pelo que aquela foi classificada como não genotipável (Figura 19 B).

A pesquisa de fenómenos de hipermutação por ocorrência de hiperadenilação (G→A) resultante da actividade das APOBEC foi efectuada, para todas as sequências em estudo, com base no relatório interpretativo da análise das sequências aminoacídicas obtido (analisado à frente, em maior detalhe), como já mencionado, utilizando o programa *HIVdb – Genotypic Resistance Interpretation Algorithm* da *Stanford University HIV Drug Resistance Database* (versão 8.5). É sabido que as enzimas da família APOBEC, activas nas células-alvo do HIV-1, em particular a APOBEC3G, participam nos mecanismos de imunidade inata contra a infecção por este vírus. São capazes de se inserir nas partículas virais durante o seu processo de montagem e actuar no ciclo replicativo seguinte, promovendo a ocorrência de mutações que podem resultar na introdução de codões de terminação no genoma viral, tornando as partículas virais produzidas defectivas, não replicativas (93,195). Esta é uma importante defesa celular baseada na conversão de resíduos de desoxicitidina em resíduos de desoxiuridina na molécula de DNA complementar, de polaridade negativa, por desaminação, durante a retrotranscrição,

resultando na substituição de resíduos de desoxiguanidina por resíduos de desoxiadenosina na molécula correspondente de DNA proviral formada. No entanto, é de salientar que a proteína viral acessória Vif é capaz de contrariar a actividade destas enzimas, sinalizando-as para degradação no proteassoma após ubiquitinação, mas também actuando directamente ao nível da sua expressão nas células infectadas (93).

A análise dos relatórios interpretativos emitidos pelo programa *HIVdb* revelou (como descrito na secção 3.3.3) a presença de mutações características da actividade das APOBEC em sete sequências da região codificante da PR (AO\_PR6, AO\_PR474, AO\_PR484, AO\_PR497, AO\_PR503, AO\_PR525 e AO\_PR533) e três sequências da região codificante da RT (AO\_RT6, AO\_RT497 e AO\_RT533), não tendo sido verificadas associações estatisticamente significativas entre a sua presença e a impossibilidade de atribuir um genótipo às sequências em estudo ou, quando possível, qualquer um dos genótipos encontrados. A presença de um número considerável deste tipo de mutações nas sequências AO\_PR6, AO\_PR474, AO\_PR525, AO\_RT6 e AO\_RT497 (6, 3, 9, 8 e 7, respectivamente) poderá justificar, em parte, o longo comprimento dos ramos que as suportam nas respectivas árvores filogenéticas (Figuras 18 A e 19 A).

Para uma caracterização mais completa da região do gene *pol* estudada, foram comparados os resultados da análise filogenética das sequências codificantes de ambas as enzimas e obtidos os genótipos combinados (PR+RT) para cada uma das 47 amostras para as quais foram obtidas sequências consenso de ambas as regiões codificantes (em nove amostras só foi possível amplificar uma das regiões codificantes) (Anexo 9). Esta análise não inclui uma pequena região de 50 nucleótidos entre as regiões codificantes da PR e da RT, uma vez que, não considerando os *primers*, as regiões amplificadas não eram contíguas. Verificou-se um elevado nível de concordância entre os genótipos das regiões da PR e da RT, sendo que 87,2% (41/47) das amostras apresentavam genótipos concordantes, o que permitiu concluir, desde logo, que as amostras AO\_119, AO\_466, AO\_473, AO\_476, AO\_510 e AO\_540 apresentam, de facto, para a região do gene *pol* estudada, o genótipo CRF02\_AG.

Foram identificadas duas possíveis URFs, nomeadamente nas amostras AO\_461 e AO\_538, cujo genótipo combinado é do tipo A/H, sendo que já foram reportadas URFs com padrões de recombinação genética entre estes dois subtipos, nas regiões codificantes

da PR e da RT, em amostras de Angola (154,158,161,162). Num desses estudos, Bártolo *et al.* (2009) descreveram em dois isolados este mesmo padrão de recombinação, nomeadamente, A1/H, mas não é possível fazer inferências sobre a dimensão da sua expansão geográfica, uma vez que não é reportada com exactidão a origem das amostras estudadas. Dado que são já quatro as estirpes angolanas que apresentam um padrão de recombinação A/H entre as regiões codificantes PR/RT, poderia colocar-se a hipótese de estarmos perante uma nova CRF, com eventual *fitness* aumentada. Contudo, seria necessário analisar e comparar o genoma completo destes vírus para se comprovar esta hipótese e, nesse caso, garantir que os indivíduos infectados não estão epidemiologicamente relacionados. Estas observações poderão estar relacionadas com o elevado potencial de recombinação do subtipo H, embora lhe seja reportada baixa *fitness* biológica e transmissibilidade, suportado pelo facto de surgir em várias CRFs, apesar de ser um dos subtipos puros mais raros a nível global (163).

As sequências da amostra AO\_532 apresentam um genótipo combinado do tipo G/G (CRF02\_AG), o qual pode ser explicado de diversas formas: pode tratar-se de um vírus com um genótipo do subtipo G puro algo divergente, dado que a região codificante da RT é do subtipo G na CRF02\_AG (Anexo 7); tendo em conta este último aspecto, poderia também tratar-se de um dos vírus do subtipo G ancestral, ou próximo do ancestral, da CRF02\_AG, embora haja pelo menos um estudo que indique que a evolução das duas formas genéticas poderá ter ocorrido exactamente em sentido contrário (196), o que implica que a actual classificação do HIV-1 não reflecte a sua história evolutiva; ou poderá ser um vírus recombinante de segunda geração entre o subtipo G puro e a CRF02\_AG, sendo esta a hipótese mais provável dado que ambos os genótipos foram reportados em estudos anteriores como estando envolvidos em eventos de recombinação genética que deram origem a novas URFs em Angola, cujo elevado número é sugestivo de cenários frequentes de co- e superinfecção na população angolana (161,162). Por esse motivo, considerou-se que o genótipo combinado desta amostra é do tipo recombinante entre o subtipo G e a CRF02\_AG. Contudo, a confirmação de qualquer uma das hipóteses apresentadas apenas poderia ser realizada após a sequenciação e genotipagem do restante genoma viral (não analisado neste trabalho).

Constatou-se ainda que as sequências de quatro amostras apresentam proximidade filogenética parcial com os subtipos B e D, nomeadamente, das amostras AO\_446,

AO\_474, AO\_488 (que apresentam proximidade com o subtipo D numa das regiões analisadas, tendo inclusive a última sido classificada como sendo desse subtipo) e AO\_533 (classificada como subtipo B para a região codificante da PR, mas não genotipável para a região codificante da RT) (Figuras 18 e 19), podendo tratar-se de casos de infecção com vírus possuindo genótipos desses subtipos, mas muito divergentes. Esta elevada divergência genética poderá ser responsável pela menor eficiência na amplificação, considerando que, para duas destas amostras (AO\_446 e AO\_474), não foi possível obter amplificação para uma das regiões em estudo (Anexo 9). Assim, considerou-se que os genótipos combinados das amostras AO\_488 e AO\_533 correspondiam ao genótipo da região em estudo que foi possível classificar inequivocamente – subtipos D e B, respectivamente. O mesmo critério foi aplicado à amostra AO\_527, cuja sequência da região codificante da RT foi considerada não genotipável (Figura 19 B), embora apresentasse alguma identidade, muito parcial, com o subtipo G (o genótipo da região codificante da PR correspondente), pelo que pode também tratar-se de um caso de infecção com um vírus do subtipo G mas muito divergente. Finalmente, no caso das sequências da amostra AO\_480, não foi possível atribuir nenhum genótipo, com um grau de certeza elevado. As sequências PR e RT surgem em posições muito ambíguas em ambas as árvores filogenéticas (Figuras 18 A e 19 A) e a análise de *bootscanning* efectuada também não foi elucidativa, em particular para a região codificante da PR. Deste modo, optou-se por uma abordagem conservadora, classificando ambas as regiões em estudo como não genotipáveis (U). Estes exemplos assemelham-se a outros reportados na literatura para Angola (154,162), nos quais se descrevem sequências de HIV-1 que formam agrupamentos filogenéticos com topologias mal definidas e subestruturas pouco organizadas, não permitindo uma distinção clara entre subtipos, sugerindo que a actual classificação do HIV-1 não permite retratar a extensão da diversidade genética encontrada em Angola. Considerando estes aspectos, seria interessante estudar o genoma completo deste vírus, pois pode tratar-se de uma nova forma genética, rara, ou inclusive de um novo subtipo do HIV-1, uma vez que não apresenta proximidade filogenética com nenhum dos subtipos e sub-subtipos virais ou CRFs conhecidas.

Em suma, os resultados obtidos revelaram a existência de uma elevada diversidade genética do HIV-1 na amostra de indivíduos infectados residentes em Benguela

diagnosticados no decorrer deste projecto (e do estudo no qual o mesmo se insere) entre Agosto de 2016 e Janeiro de 2017. Os subtipos mais frequentes na amostra populacional são o subtipo C (40,4%) e o sub-subtipo F1 (19,1%), seguidos da forma recombinante CRF02\_AG (12,8%). Foram também encontrados os subtipos G (8,5%), A (4,3%), B, D e J (2,1%, cada), e ainda uma pequena percentagem de URFs (6,4%), sendo que apenas uma amostra (2,1%) foi considerada não genotipável (Figura 20). Estes dados globais são relativos às 47 amostras para as quais foi possível amplificar ambas as regiões codificantes em estudo, estando resumidos na Tabela 9 os resultados da genotipagem da totalidade das sequências das regiões da PR e da RT. De um modo geral, estes resultados são concordantes com os obtidos em estudos anteriores, que demonstram que a epidemia do HIV-1 em Angola é caracterizada por uma elevada diversidade genética das estirpes circulantes (154,155,158,159,161,162). Era por isso expectável encontrar uma grande diversidade de genótipos nas sequências em estudo, tendo-se observado uma quase exclusividade dos subtipos não-B, assim como um número apreciável de sequências U e algumas URFs, também frequentemente encontradas em amostras angolanas. Estas últimas têm vindo a aumentar em frequência ao longo do tempo, enquanto o número de CRFs encontrado tem permanecido mais ou menos estável (162). O predomínio do subtipo C nas amostras em estudo (40,4%) pode dever-se à movimentação das populações entre Benguela e outras regiões do país (ou estrangeiro) dada a sua localização geográfica, uma vez que a província fica próxima da região sul do país onde este subtipo é mais frequente devido, muito provavelmente, ao influxo de estirpes virais com origem nos países que fazem fronteira com o sul e sudeste de Angola, ou se situam próximo, nomeadamente, a Zâmbia, o Botswana e o Zimbabué, onde a prevalência do subtipo C é muito elevada, assim como na região sul da RDC (161,162). Por outro lado, são conhecidas as relações sociais e comerciais estreitas entre as províncias do sul de Angola e a África do Sul, país com elevada prevalência de infecção pelo HIV-1 do subtipo C. À semelhança de estudos anteriores (154,162), também se verificou neste trabalho a presença de sequências em posições basilares (antes da ramificação dos grupos monofiléticos dos diferentes subtipos) na respectiva árvore filogenética, nomeadamente as sequências AO\_PR480, AO\_RT380, AO\_RT446, AO\_RT480 e AO\_RT527 (Figuras 18 A e 19 A), reforçando a ideia de que o HIV-1 circula há já muito tempo em Angola. Por fim, verificam-se diferenças entre os resultados obtidos e a distribuição de

genótipos para Angola na *Los Alamos HIV Sequence Database* (<http://www.hiv.lanl.gov/>), onde as percentagens de sequências dos genótipos C e CRF02\_AG (18,1% e 6,4%, respectivamente) são cerca de metade do que aqui foi encontrado, existindo ainda um número superior de sequências do subtipo A (14,5%). Estas diferenças poderão estar relacionadas com o facto de estarmos a estudar uma amostra relativamente reduzida, de apenas uma província, que está sujeita à influência epidemiológica dos países vizinhos a sul e sudeste, como já foi referido. Além disso, não se encontram, na mesma base de dados, sequências angolanas do subtipo B, embora já tenham sido descritos dois isolados deste subtipo para a região codificante da RT (159) e um recombinante B/D entre as regiões codificantes PR/RT (155), todos provenientes de Luanda, pelo que a sequência AO\_PR533 não é a primeira do subtipo B a ser reportada para Angola. As análises adicionais realizadas não permitiram chegar a conclusões sobre a provável origem geográfica desta estirpe do subtipo B. Finalmente, deve realçar-se que a distribuição de genótipos para Angola na *Los Alamos HIV Sequence Database* pode não representar fielmente a situação epidemiológica molecular no terreno.

Concluída a análise filogenética das sequências nucleotídicas, procedeu-se ao estudo das respectivas sequências aminoácídicas putativas e pesquisa de mutações associadas a diminuição da susceptibilidade aos ARVs inibidores da PR e da RT. Para tal, foi utilizado o *software HIVdb*, já mencionado, o qual gera para cada uma das sequências analisadas um relatório interpretativo com as diferentes mutações encontradas baseando-se na comparação com uma sequência consenso de referência do subtipo B. Neste relatório é indicado o perfil fenotípico mais provável para cada sequência e cada um dos ARVs considerados com base no método descrito na secção 2.8.1. Convém, desde logo, salientar que nesta análise deverá ser tido em consideração o facto de que, neste trabalho, as sequências em estudo foram obtidas a partir da amplificação do DNA proviral presente nas células mononucleadas de sangue periférico infectadas existentes na mancha de sangue armazenada nos cartões *FTA*<sup>®</sup>. Nestas circunstâncias, o DNA proviral pode não se encontrar transcricionalmente activo, podendo antes representar populações virais latentes que não são necessariamente representativas da população viral presente no plasma (vírus replicativos). É por este motivo que o RNA genómico viral isolado a partir do plasma continua a ser o material biológico preferencialmente utilizado na pesquisa de mutações de resistência (128,197). Porém, os vírus que se encontram latentes nas células

mononucleadas, nomeadamente, nos linfócitos T CD4<sup>+</sup>, podem reemergir quando há activação dessas células mediante determinadas condições e influenciar a resposta à terapia anti-retroviral, sendo que, caso se tratem de estirpes resistentes, podem inclusive levar a falência virológica (197,198). A pesquisa de mutações de resistência no DNA proviral apresenta, no entanto, algumas vantagens, em particular quando os níveis da carga viral são indetectáveis ou quando há interrupções/alterações na terapia anti-retroviral, dada a possível reemergência de vírus latentes. A análise dos seus resultados exige alguma atenção, pois as mutações encontradas nem sempre coincidem com as detectadas no RNA viral (apesar de alguns estudos reportarem elevados níveis de concordância entre o RNA viral e o DNA proviral) e, por outro lado, o DNA proviral apresenta mais frequentemente hipermutações e, consequentemente, genomas defectivos (como mencionado anteriormente) (197–200). Por fim, há ainda a possibilidade de existirem estirpes minoritárias, clinicamente relevantes, com mutações associadas a resistência aos ARVs no conjunto das *quasispecies* virais, as quais podem não ser detectadas, independentemente do tipo de material biológico de partida, devido ao limiar de detecção (>20%) dos testes genotípicos utilizados para a pesquisa de mutações de resistência. Apesar desta, e outras limitações, estes testes são, no entanto, preferencialmente utilizados, por serem mais rápidos e económicos (126,128,132).

De um modo geral, verificou-se a ausência de mutações de *frameshift* (causadas por inserções ou deleções simples ou duplas de nucleótidos) nas sequências em estudo de ambas as regiões analisadas. Contudo, constatou-se que algumas mutações do tipo APOBEC identificadas resultaram na introdução de codões de terminação (na posição 42, nas sequências AO\_PR6, AO\_PR474, AO\_PR484, AO\_PR497 e AO\_PR525, e nas posições 88, 153 e 212, nas sequências AO\_RT6 e AO\_RT497, sendo que esta última possuía ainda outros codões de terminação nas posições 71 e 210, dos quais o último, segundo o relatório interpretativo do programa *HIVdb*, não estará relacionado com actividade APOBEC), indiciando que as estirpes nas quais foram encontradas corresponderão a vírus defectivos para pelo menos uma das enzimas estudadas e, como tal, não replicativos, dada a importância das mesmas no ciclo replicativo do HIV-1. Teoricamente, os provírus em causa podem dar origem a partículas virais imaturas e não infecciosas, no caso de a PR ser não funcional, impedindo dessa forma que ocorra a maturação das partículas virais produzidas, ou viriões que, apesar de infecciosos,

possuem ciclos replicativos abortivos durante a retrotranscrição do genoma viral, no caso de a RT ser não funcional, sendo igualmente possível que nem se cheguem a formar quaisquer partículas virais, dependendo da localização na grelha de leitura aberta do gene *pol* do primeiro codão de terminação introduzido. Outras mutações APOBEC foram encontradas em posições associadas a mutações de resistência (na posição 48, nas sequências AO\_PR503 e AO\_PR525, e na posição 190, na sequência AO\_RT6). Embora consideradas nos relatórios interpretativos como mutações acessórias de resistência aos PIs ou associadas a resistência aos NNRTIs, trata-se, de acordo com a *Stanford University HIV Drug Resistance Database* (201), de mutações muito pouco comuns nas respectivas posições, aparentando não possuir grande impacto na susceptibilidade aos ARVs.

Começando pelo conjunto das sequências codificantes da PR, foram encontradas, num total de quatro sequências, as mutações principais D30N (2,0%), M46I (2,0%), I54V (2,0%) e L90M (2,0%) e a mutação acessória G73S (4,0%), todas associadas a resistência aos PIs, sendo que em cada sequência foi somente observada uma das mutações mencionadas, com excepção da sequência AO\_PR525, que apresentava em simultâneo as mutações D30N, M46I e G73S (esta última foi também encontrada na sequência AO\_PR474). Todas as mutações referidas apresentam importância para a vigilância epidemiológica da resistência transmitida aos PIs no HIV-1, pois cumprem os critérios adoptados pela OMS para a sua definição, a saber: as mutações devem ser reconhecidas como causadoras ou contribuintes para a resistência aos ARVs (definido como constarem em pelo menos 3/5 das listas de mutações de resistência elaboradas por especialistas); devem ser não polimórficas e não ocorrer em posições altamente polimórficas; a lista onde constam deve ser aplicável aos oito subtipos mais comuns; e a mesma deve ser o mais reduzida possível, excluindo mutações raras resultantes da pressão selectiva exercida pelos fármacos (180). Além destas, foram também identificadas outras mutações secundárias, associadas a resistência, mas na presença de outras mutações, nomeadamente, K20I (28%), L10V (22,0%), L10I (14,0%), V82I (12,0%), T74S (4,0%) e A71I (2,0%), das quais, apenas L10I, K20I e V82I foram encontradas associadas a mutações principais, em três sequências. Quando comparados com estudos anteriores igualmente realizados com amostras de indivíduos nunca submetidos a terapia anti-retroviral, estes resultados revelam a presença de mutações principais e acessórias associadas a diminuição da susceptibilidade aos PIs nunca antes descritas na população

angolana, nomeadamente, D30N, M46I, I54V, L90M e G73S, mas também a presença de mutações já anteriormente descritas (L10I/V, K20I e T74S) (155,160–162), sendo que estas últimas surgiram em proporções semelhantes às já relatadas. No entanto, o relatório interpretativo das sequências AO\_PR474 e AO\_PR525 refere que as mutações de resistência aos PIs nelas encontradas (D30N, M46I e G73S) podem resultar da actividade das enzimas APOBEC, o que, associado ao facto de apresentarem também outras mutações APOBEC, em particular a introdução de um codão de terminação na posição 42, indica que estas mutações, presentes unicamente nestas sequências, poderão não ter verdadeiro impacto na terapêutica, uma vez que foram identificadas em estirpes possivelmente defectivas para a PR e, como tal, muito provavelmente, não replicativas. Este facto vem assim conferir relevo particular à presença das mutações principais I54V e L90M, em duas das sequências PR estudadas.

A análise dos perfis fenotípicos putativos de resistência de cada sequência da PR permitiu concluir que a sua esmagadora maioria (92,0%) possui perfis fenotípicos de susceptibilidade aos diferentes PIs, sendo que as restantes (8,0%) apresentam algum tipo de mutações associadas a resistência. De salvaguardar que no caso das sequências AO\_PR474 e AO\_PR525 os respectivos perfis fenotípicos putativos podem não ter relevância na prática clínica, pelas razões já discutidas. As mutações principais I54V e L90M, encontradas em uma sequência cada, estão associadas a resistência (em níveis diferentes) a todos os PIs, com excepção do darunavir/r, daí o respectivo perfil de susceptibilidade total observado para o conjunto das sequências analisadas (Figura 22). Estas apresentavam ainda um perfil de quase susceptibilidade total para o tipranavir/r, verificando-se para este fármaco apenas um perfil de resistência, ainda que de nível baixo, na sequência com a mutação I54V (AO\_PR542). Segundo a definição para este nível de resistência, isto significa que o vírus poderá apresentar uma diminuição da susceptibilidade ao tipranavir/r *in vitro* e que o indivíduo infectado poderá desenvolver uma resposta virológica sub-ótima, se tratado com esse fármaco. Nessa mesma sequência, foi também observada a mutação L10I, a qual, em combinação com a mutação I54V, pode conferir um maior nível de resistência ou melhorar a capacidade replicativa do vírus, aumentando a sua *fitness* (202,203). O facto de terem sido atribuídos perfis de resistência de nível elevado ao nelfinavir a duas sequências (AO\_PR494 e AO\_PR525, pese embora o impacto da actividade APOBEC nesta última) é indicativo que o

tratamento com este fármaco nestes casos será, muito provavelmente, ineficaz, devendo optar-se por outro PI. Dado que estamos a trabalhar sobre amostras de indivíduos nunca submetidos a terapia, o mais provável é que estejamos perante cenários de resistência transmitida, embora, aparentemente, ainda a níveis não preocupantes para os PIs.

Foi ainda realizada a pesquisa de assinaturas genéticas relativamente aos genótipos observados para a região da PR, tendo-se constatado correlações positivas (ou seja, a associação significativa da mutação analisada a um determinado genótipo) entre a mutação L10V e o sub-subtipo F1, a mutação K20I e os genótipos G e CRF02\_AG e a mutação V82I e o subtipo G, sendo que, em nenhum dos casos, se tratam de mutações principais, ou mesmo acessórias, de resistência aos PIs. No caso das mutações K20I e V82I, as correlações observadas são facilmente justificadas pelo facto de o resíduo mutado corresponder ao aminoácido consenso nos subtipos com os quais foi confirmada a associação (204), o que vem confirmar o nosso achado de que se trata de assinaturas genéticas desses mesmos subtipos. Já no caso da mutação L10V, atendendo ao facto de que as estirpes F1 circulantes na Roménia e no Brasil são filogeneticamente próximas das estirpes circulantes do mesmo subtipo em Angola (193), compreende-se a semelhança com resultados descritos em amostras provenientes destes dois países, nas quais a mutação L10V foi também muito frequente no sub-subtipo F1 (205,206).

Passando para o conjunto das sequências codificantes da RT, foram encontradas as mutações M184I/V (3,8/17,0%), T215F/Y (1,9/9,4%), M41L (5,7%), D67N (3,8%), E44D (1,9%), K70E/T (1,9%, cada), V75M (1,9%), Y115F (1,9%), L210W (1,9%) e K219N/R (1,9%, cada) associadas a resistência aos NRTIs, num conjunto de doze sequências, e as mutações Y181C (18,9%), V108I (9,4%), G190A (7,5%), H221Y (7,5%), A98G (5,7%), K103N (5,7%), E138A (3,8%), V106A/M (1,9/3,8%), K101E/H (1,9% cada) e Y188L (1,9%) associadas a resistência aos NNRTIs, num conjunto de dezassete sequências. No cômputo geral, constatou-se a presença de mutações associadas a diminuição de susceptibilidade aos NRTIs e/ou NNRTIs num total de 18 das 53 sequências RT em estudo. Pelo menos uma das mutações em cada sequência tem importância para a vigilância epidemiológica da resistência transmitida aos RTIs no HIV-1, pois apenas as mutações E44D e K70T, associadas a resistência aos NRTIs, e as mutações A98G, K101H, V108I, E138A e H221Y, associadas a resistência aos NNRTIs, não cumprem os critérios enunciados pela OMS para aquela definição (180). À

semelhança do que se verificou para as sequências da PR, foram ainda identificadas mutações associadas a resistência aos RTIs na presença de outras mutações, nomeadamente, V179I (7,5%), V106I (5,7%), S68G (3,8%), V90I (3,8%), V118I (3,8%), T69S (1,9%) e K101N (1,9%), sendo que todas (com excepção da T69S) surgem, em diferentes proporções, associadas a mutações com impacto na resistência aos NRTIs e/ou NNRTIs, num conjunto de sete sequências. Porém, estas mutações possuem, em geral, um impacto mínimo na susceptibilidade aos RTIs, sendo que algumas têm até os seus efeitos atenuados na presença de mutações específicas (201). Comparando os resultados obtidos com os de estudos anteriores, nos quais também apenas foram incluídos indivíduos nunca submetidos a terapia anti-retroviral, constatou-se que algumas das mutações encontradas já tinham sido descritas na população angolana, mas não a sua totalidade (155,160–162). Destaca-se o facto de as mutações M184V – associada a resistência a emtricitabine e lamivudine, mas também ao aumento da susceptibilidade a zidovudine, stavudine e tenofovir (207,208) – e Y181C – associada a resistência (em diferentes níveis) a efavirenz, etravirine, nevirapine e rilpivirine (209) – serem as mais frequentes (17,0% e 18,9%, respectivamente), à semelhança de outros estudos (153), ainda que em diferentes proporções. Verificou-se também que a mutação T215Y – TAM associada a resistência (em diferentes níveis) a ziduvine, stavudine, didanosine, abacavir e tenofovir (210) –, que apenas tinha sido encontrada num único indivíduo no estudo de Bártolo *et al.* (160), constitui a segunda mutação associada a resistência aos NRTIs mais frequente na amostra populacional estudada (9,4%), e que a mutação V108I – associada a resistência (em diferentes níveis) a nevirapine e efavirenz, aparentando também reduzir a susceptibilidade a etravirine e rilpivirine (211,212) –, presente em 9,4% das amostras em estudo, foi pela primeira vez observada na população angolana. Finalmente, é importante salientar que os relatórios interpretativos das sequências AO\_RT6 e AO\_RT497, nas quais foi encontrado um número considerável de hipermutações e, em particular, de codões de terminação, indicam, especificamente, que as mutações M184I e D67N, presentes, cada uma, respectivamente, em cada uma daquelas sequências, podem resultar de actividade APOBEC. Assim, poderá especular-se que as diferentes mutações associadas a diminuição da susceptibilidade aos diferentes RTIs encontradas nas referidas sequências poderão não ter impacto na terapêutica, correspondendo aquelas a possíveis estirpes defectivas para RT e, como tal, não replicativas.

Através da análise dos perfis fenotípicos putativos de resistência, constatou-se que, para aproximadamente um terço do total de sequências (18/53, 34,0%), foi identificado um perfil de resistência, de nível diverso, a pelo menos um dos RTIs. No caso das sequências AO\_RT6 e AO\_RT497, os respectivos perfis fenotípicos putativos não terão relevância na prática clínica pelas razões anteriormente apresentadas. Verificou-se que dos diferentes RTIs considerados nesta análise, o nevirapine (um NNRTI) é aquele para o qual um maior número de sequências está associado a níveis elevados de resistência putativa (14 das 18 sequências nas quais foram identificadas mutações de resistência aos RTIs, correspondendo a 26,4% do total das sequências RT analisadas), logo seguido de emtricitabine e de lamivudine (dois NRTIs), para os quais foi reportada uma prevalência de 20,8% de sequências contendo mutações associadas a um perfil fenotípico de resistência de nível elevado. Ao invés, o maior grau de susceptibilidade foi encontrado para o tenofovir (um NRTI), para o qual foi sugerido algum tipo de resistência em apenas quatro sequências, entre os níveis baixo e elevado (Figura 24). É importante destacar o caso de uma sequência em particular (AO\_RT520), na qual se verificaram níveis de resistência putativos de nível elevado a todos os NRTIs e dois dos NNRTIs (nevirapine e rilpivirine) considerados na análise, apresentando níveis de resistência intermédios para os restantes (efavirenz e etravirine), tratando-se da sequência com o espectro fenotípico de resistência mais alargado do conjunto das sequências analisadas. Este facto resulta da presença na sequência AO\_RT520 de dez mutações de resistência aos RTIs (sete associadas a resistência aos NRTIs e três aos NNRTIs) (Anexo 10), pelo que um eventual tratamento da infecção pelo HIV-1 com esta classe de ARVs não é, neste caso, recomendável, considerando as definições para estes níveis de resistência apresentados pela *Stanford University HIV Drug Resistance Database* (201). Para além da sequência mencionada, outras apresentam perfis fenotípicos putativos de resistência de nível elevado em simultâneo para diversos fármacos do grupo dos RTIs, na ausência de qualquer indício de efeitos da actividade APOBEC (*e.g.*, AO\_RT92, AO\_RT458, AO\_RT461, AO\_RT478, AO\_RT488, AO\_RT494, AO\_RT507, AO\_RT534, AO\_RT542). Há inclusive algumas sequências que apresentam mutações de resistência cruzada para todos os fármacos representantes de uma determinada classe terapêutica (*e.g.*, AO\_RT520, para todos os NRTIs, e AO\_RT92, AO\_RT494 e AO\_RT542, para os NNRTIs), antecipando grande dificuldade

na implementação de uma terapia anti-retroviral combinada de alta eficácia bem sucedida nos indivíduos infectados, considerando igualmente a ausência de monitorização pré-tratamento, através da realização de testes genotípicos, presentemente em Angola.

Tal como anteriormente, para as sequências da PR, realizou-se a pesquisa de assinaturas genéticas em relação aos genótipos observados para a região da RT, tendo-se verificado somente uma correlação positiva significativa entre a mutação H221Y e o subtipo G. Esta mutação não possui, por si só, um grande impacto na resistência aos NNRTIs, mas surge comumente associada à mutação Y181C, potenciando a resistência causada por esta última ao nevirapine (213,214), o que é consistente com os resultados obtidos, dado que ambas surgem sempre associadas nas sequências nas quais foi encontrada a mutação H221Y.

Genericamente, os resultados para cada uma das regiões codificantes em análise revelam um conjunto diversificado de mutações associadas a resistência aos diferentes PIs e RTIs na amostra em estudo com origem em Benguela, tendo-se constatado que uma parte considerável das mutações identificadas apresentam relevância para a vigilância epidemiológica da resistência transmitida aos ARVs no HIV-1 (cerca de 45,5% (5/11) e 57,6% (19/33) da totalidade das mutações associadas a resistência aos PIs e RTIs encontradas, respectivamente). Embora a maioria das amostras em estudo (para as quais foi possível amplificar pelo menos uma das regiões codificantes) apresente um perfil fenotípico putativo de susceptibilidade aos ARVs considerados, 35,7% (20/56) apresentava algum tipo de resistência aos mesmos (duas amostras apresentavam em simultâneo mutações de resistência a PIs e RTIs, nomeadamente, AO\_494 e AO\_542). Este aspecto vai de encontro a achados anteriores (161), que referem uma maior prevalência de mutações de resistência na região centro de Angola, possivelmente devida à movimentação dos indivíduos para outros países onde a terapia anti-retroviral está disponível ou ao seu mais fácil acesso nesta região do país, não sendo por isso surpreendente encontrar perfis fenotípicos putativos de resistência em 21,4% (22/103) da totalidade das sequências em estudo. Verificou-se a emergência, ainda que numa baixa frequência, de mutações principais e acessórias associadas a diminuição da susceptibilidade aos PIs que ainda não tinham sido descritas na população angolana, reforçando a hipótese, anteriormente levantada (160,162), de que algumas estirpes circulantes em Angola apresentam uma baixa barreira genética para a aquisição de

resistência a alguns PIs. Também se verificou que um número considerável de amostras, aproximadamente 32,1% (18/56), apresenta mutações associadas a diminuição da susceptibilidade a diferentes RTIs, sendo que os níveis putativos de resistência observados são particularmente elevados no caso dos NNRTIs (Figura 24), o que pode ser facilmente justificado, entre outros factores, pela baixa barreira genética para aquisição de resistência característica deste tipo de RTIs (120,128,147).

Estes aspectos são particularmente relevantes, uma vez que a amostra populacional em estudo inclui exclusivamente indivíduos nunca submetidos a qualquer tipo de terapia anti-retroviral, significando que estamos perante um cenário de considerável resistência primária aos ARVs, em particular aos RTIs. Esta observação é preocupante tendo em consideração os esquemas terapêuticos de primeira linha que são normalmente aplicados em Angola, os quais incluem a utilização de dois NRTIs e um NNRTI ou PI (215), sendo que, no caso de indivíduos nunca tratados, as directrizes angolanas indicam ainda que deverá optar-se pela utilização de um NNRTI, em detrimento de um PI, a não ser que haja contra-indicação. Além disso, a realização da pesquisa de mutações de resistência está, segundo essas mesmas directrizes, apenas prevista em situações de falência virológica (215), pelo que a utilização destes esquemas de primeira linha pode ser altamente problemática em casos como os encontrados na amostra estudada, pois a sua eficácia estará, à partida, substancialmente reduzida devido à presença de mutações de resistência de nível elevado aos RTIs. No entanto, é importante salientar que os resultados da pesquisa de mutações de resistência e os perfis fenotípicos putativos decorrentes da sua interpretação não devem ditar, de modo imediato, a escolha de um esquema terapêutico específico. A decisão terapêutica deve ser tomada considerando igualmente todos os aspectos relacionados com a infecção e o indivíduo infectado, desde a carga viral e a contagem de linfócitos T CD4<sup>+</sup>, até às interacções farmacológicas e estádios da função renal e hepática, além de factores como o grau previsível de adesão à terapêutica, cujo impacto na eficácia do tratamento já foi comprovado e descrito em diversos estudos (178).

Concluindo, com a realização deste estudo foi possível confirmar que as características moleculares das estirpes do HIV-1 circulantes em Benguela são semelhantes às que foram descritas em estudos anteriores para a epidemia do mesmo em Angola (154,155,157–162). As estirpes circulantes na província de Benguela presentes

na amostra populacional estudada revelaram a existência de uma elevada diversidade genética, com um claro predomínio do subtipo C, mas também um considerável grau de resistência primária aos ARVs (principalmente aos RTIs) indicativo de uma frequente transmissão de vírus resistentes. Deste modo, confirma-se a crescente complexidade da epidemia do HIV-1 em Angola, que aparenta caracterizar-se pela circulação de um número crescente de vírus com mutações associadas a diminuição da susceptibilidade aos ARVs utilizados nos regimes terapêuticos de primeira linha, em resultado de diversos factores, como o uso intermitente dos ARVs, o uso não regularizado e não monitorizado dos mesmos (que muitas vezes não é reportado por questões sociais), ou os movimentos migratórios das populações, que levam à introdução no país de novas estirpes virais (155,160–162). Estes achados vão também de encontro aos resultados de Gupta *et al.* (183), que demonstram o aumento da prevalência da resistência aos ARVs em vários países de médio e baixo rendimento, em particular a resistência aos NNRTIs, sendo a transmissão de vírus resistentes apontada como a causa mais plausível de falência terapêutica no caso dos indivíduos não-experimentados.

É importante referir que este estudo possui algumas limitações, entre as quais o número relativamente reduzido de amostras incluídas e a relativamente baixa taxa de sucesso de amplificação das regiões codificantes do gene *pol* em estudo, tendo já sido apresentadas as razões por detrás das mesmas e que se prendem, essencialmente, com dificuldades na optimização inicial do protocolo laboratorial para a obtenção dos fragmentos de DNA a analisar, e, como consequência, constrangimentos ao nível do cronograma, que ditaram a exclusão de algumas amostras, por forma a concluir o estudo nos prazos estipulados. Como tal, os resultados aqui apresentados poderão não retratar com inteira precisão aquela que é a realidade actual da epidemia do HIV-1 em Benguela, até porque a amostragem foi realizada durante um curto período de seis meses, podendo apenas ser retiradas ilações sobre as características da epidemia durante esse período. Inclusivamente, os resultados obtidos não incluem informação sobre as estirpes minoritárias presentes na amostra populacional estudada, dadas as características do método de genotipagem e pesquisa de mutações de resistência aos ARVs utilizado, o qual não permite a sua detecção.

Ainda assim, reforça-se com este trabalho a importância de manter constante a vigilância dos vários aspectos inerentes à infecção pelo HIV-1 em Angola, especialmente

das mutações de resistência aos diferentes ARVs e sua transmissão, devido ao seu impacto directo na terapêutica e controlo da infecção. Devem ser realizados esforços no sentido de melhorar as ferramentas de prevenção, diagnóstico, tratamento e gestão da infecção pelo HIV-1 disponíveis nos países de médio e baixo rendimento, uma vez que ainda são verificadas claras diferenças nos padrões da pandemia entre estes países e os países de alto rendimento (178,182). No futuro, seria útil a continuação deste estudo, através da análise das restantes amostras que não foi possível incluir neste trabalho, mas também da realização de protocolos de amplificação alternativos, com objectivo de aumentar as taxas de positividade para ambas as regiões do genoma analisadas. Seria também importante a realização de mais projectos como este (e o estudo no qual o mesmo se insere), não só em Benguela, como noutras regiões de Angola, de modo a poder obter informação mais completa sobre a epidemia do HIV-1 neste país, contribuindo de um modo activo para um melhor controlo e monitorização da mesma.

## 5 Referências bibliográficas

---

1. Hemelaar J. The origin and diversity of the HIV-1 pandemic. *Trends Mol Med*. 2012;18(3):182–92. Disponível em: <http://dx.doi.org/10.1016/j.molmed.2011.12.001>
2. Becerra JC, Bildstein LS, Gach JS. Recent Insights into the HIV/AIDS Pandemic. *Microb Cell*. 2016;3(9):451–75. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/28357381>
3. Sierra S, Kupfer B, Kaiser R. Basics of the virology of HIV-1 and its replication. *J Clin Virol*. 2005;34(4):233–44. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/16198625>
4. Hutchinson JF. THE BIOLOGY AND EVOLUTION OF HIV. *Annu Rev Anthropol*. 2001;30:85–108. Disponível em: <http://www.annualreviews.org/doi/pdf/10.1146/annurev.anthro.30.1.85>
5. Centers for Disease Control (CDC). Pneumocystis pneumonia--Los Angeles. *MMWR Morb Mortal Wkly Rep*. 5 de Junho de 1981;30(21):250–2. Disponível em: [https://www.cdc.gov/mmwr/preview/mmwrhtml/june\\_5.htm](https://www.cdc.gov/mmwr/preview/mmwrhtml/june_5.htm)
6. Centers for Disease Control (CDC). Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR Morb Mortal Wkly Rep*. 3 de Julho de 1981;30(25):305–8. Disponível em: <http://www.jwatch.org/ac200106010000008/2001/06/01/kaposi-sarcoma-and-pneumocystis-pneumonia-among>
7. Hymes KB, Cheung T, Greene JB, Prose NS, Marcus A, Ballard H, et al. Kaposi's sarcoma in homosexual men-a report of eight cases. *Lancet (London, England)*. 19 de Setembro de 1981;2(8247):598–600. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/6116083>
8. Masur H, Michelis MA, Greene JB, Onorato I, Vande Stouwe RA, Holzman RS, et al. An Outbreak of Community-Acquired Pneumocystis carinii Pneumonia. N

- Engl J Med. 10 de Dezembro de 1981;305(24):1431–8. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/6975437>
9. Centers for Disease Control (CDC). A cluster of Kaposi's sarcoma and Pneumocystis carinii pneumonia among homosexual male residents of Los Angeles and Orange Counties, California. MMWR Morb Mortal Wkly Rep. 18 de Junho de 1982;31(23):305–7. Disponível em:  
<https://www.cdc.gov/mmwr/preview/mmwrhtml/00001114.htm>
  10. Centers for Disease Control (CDC). Epidemiologic Notes and Reports Persistent, Generalized Lymphadenopathy among Homosexual Males. MMWR Morb Mortal Wkly Rep. 1982;31(19):249–51. Disponível em:  
<https://www.cdc.gov/mmwr/preview/mmwrhtml/00001096.htm>
  11. Centers for Disease Control (CDC). Current Trends Update on Acquired Immune Deficiency Syndrome (AIDS) --United States. MMWR Morb Mortal Wkly Rep. 24 de Setembro de 1982;31(37):507–8, 513–4. Disponível em:  
<https://www.cdc.gov/mmwr/preview/mmwrhtml/00001163.htm>
  12. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science. 20 de Maio de 1983;220(4599):868–71. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/6189183>
  13. Montagnier L, Chermann JC, Barré-Sinoussi F, Klatzmann D, Wain-Hobson S, Alizon M, et al. Lymphadenopathy associated virus and its etiological role in AIDS. Princess Takamatsu Symp. 1984;15:319–31. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/6100650>
  14. Gallo RC, Salahuddin SZ, Popovic M, Shearer GM, Kaplan M, Haynes BF, et al. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science. 4 de Maio de 1984;224(4648):500–3. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/6200936>

15. Broder S, Gallo RC. A Pathogenic Retrovirus (HTLV-III) Linked to AIDS. *N Engl J Med*. 15 de Novembro de 1984;311(20):1292–7. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/6208484>
16. Skaland I, Nordhus M, Gudlaugsson E, Klos J, Kjellekvold KH, Janssen E a M, et al. Evaluation of 5 different labeled polymer immunohistochemical detection systems. *Appl Immunohistochem Mol Morphol*. 2010;18(1):90–6.
17. Popovic M, Sarngadharan MG, Read E, Gallo RC. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science*. 4 de Maio de 1984;224(4648):497–500. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/6200935>
18. Coffin J, Haase A, Levy JA, Montagnier L, Oroszlan S, Teich N, et al. Human immunodeficiency viruses. *Science*. 9 de Maio de 1986;232(4751):697. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/3008335>
19. CASE K. Nomenclature: Human Immunodeficiency Virus. *Ann Intern Med*. 1 de Julho de 1986;105(1):133. Disponível em: <http://annals.org/article.aspx?doi=10.7326/0003-4819-105-1-133>
20. Clavel F, Guétard D, Brun-Vézinet F, Chamaret S, Rey MA, Santos-Ferreira MO, et al. Isolation of a new human retrovirus from West African patients with AIDS. *Science*. 18 de Julho de 1986;233(4761):343–6. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/2425430>
21. Clavel F, Guyader M, Guétard D, Sallé M, Montagnier L, Alizon M. Molecular cloning and polymorphism of the human immune deficiency virus type 2. *Nature*. 18 de Dezembro de 1986;324(6098):691–5. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/3025743>
22. Clavel F, Mansinho K, Chamaret S, Guetard D, Favier V, Nina J, et al. Human Immunodeficiency Virus Type 2 Infection Associated with AIDS in West Africa. *N Engl J Med*. 7 de Maio de 1987;316(19):1180–5. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/3472076>

23. Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med. 2011;1(1):1–22. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3234451/>
24. Tebit DM, Arts EJ. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. Lancet Infect Dis. 2011;11(1):45–56. Disponível em: [http://dx.doi.org/10.1016/S1473-3099\(10\)70186-9](http://dx.doi.org/10.1016/S1473-3099(10)70186-9)
25. Ekouevi DK, Balestre E, Coffie PA, Minta D, Messou E, Sawadogo A, et al. Characteristics of HIV-2 and HIV-1/HIV-2 Dually Seropositive Adults in West Africa Presenting for Care and Antiretroviral Therapy: The IeDEA-West Africa HIV-2 Cohort Study. PLoS One. 2013;8(6):1–8. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23824279>
26. Tchounga B, Ekouevi DK, Balestre E, Dabis F. Mortality and survival patterns of people living with HIV-2. Curr Opin HIV AIDS. Setembro de 2016;11(5):537–44. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/27254747>
27. Hahn BH, Shaw GM, De Cock KM, Sharp PM. AIDS as a zoonosis: scientific and public health implications. Science. 28 de Janeiro de 2000;287(5453):607–14. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/10649986>
28. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, et al. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature. 4 de Fevereiro de 1999;397(6718):436–41. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/9989410>
29. Huet T, Cheynier R, Meyerhans A, Roelants G, Wain-Hobson S. Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature. 24 de Maio de 1990;345(6273):356–9. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/2188136>
30. D'arc M, Ayoub A, Esteban A, Learn GH, Boué V, Liegeois F, et al. Origin of the HIV-1 group O epidemic in western lowland gorillas. Proc Natl Acad Sci. 2015;112(11):E1343–52. Disponível em:

<http://www.pnas.org/lookup/doi/10.1073/pnas.1502022112>

31. Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR. An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature*. 1 de Junho de 1989;339(6223):389–92. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/2786147>
32. Sousa JD de, Müller V, Lemey P, Vandamme A-M. High GUD Incidence in the Early 20th Century Created a Particularly Permissive Time Window for the Origin and Initial Spread of Epidemic HIV Strains. Martin DP, editor. *PLoS One*. 1 de Abril de 2010;5(4):e9936. Disponível em:  
<http://dx.plos.org/10.1371/journal.pone.0009936>
33. Chen Z, Telfer P, Gettie A, Reed P, Zhang L, Ho DD, et al. Genetic Characterization of New West African Simian Immunodeficiency Virus SIVsm: Geographic Clustering of Household-Derived SIV Strains with Human Immunodeficiency Virus Type 2 Subtypes and Genetically Diverse Viruses from a Single Feral Sooty Mangabey T. *J Virol*. 1996;70(6):3617–27. Disponível em:  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC190237/pdf/703617.pdf>
34. Santiago ML, Range F, Keele BF, Li Y, Bailes E, Bibollet-Ruche F, et al. Simian immunodeficiency virus infection in free-ranging sooty mangabeys (*Cercopithecus atys atys*) from the Taï Forest, Côte d’Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. *J Virol*. 1 de Outubro de 2005;79(19):12515–27. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/16160179>
35. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science*. 2014;346(6205):56–61. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/25278604>
36. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, et al. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature*. 2 de Outubro de 2008;455(7213):661–4. Disponível em:

- <http://www.ncbi.nlm.nih.gov/pubmed/18833279>
37. Pépin J. The origins of AIDS: from patient zero to ground zero. *J Epidemiol Community Health*. 1 de Junho de 2013;67(6):473–5. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23322854>
  38. UNAIDS. 2017 Global HIV Statistics Fact Sheet. 2018. Disponível em: [http://www.unaids.org/sites/default/files/media\\_asset/UNAIDS\\_FactSheet\\_en.pdf](http://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf)
  39. UNAIDS. UNAIDS DATA 2018. 2018. Disponível em: [http://www.unaids.org/sites/default/files/media\\_asset/unaid-data-2018\\_en.pdf](http://www.unaids.org/sites/default/files/media_asset/unaid-data-2018_en.pdf)
  40. International Committee on Taxonomy of Viruses. ICTV Taxonomy history: Human immunodeficiency virus 1. Disponível em: [https://talk.ictvonline.org/taxonomy/p/taxonomy-history?taxnode\\_id=20164256](https://talk.ictvonline.org/taxonomy/p/taxonomy-history?taxnode_id=20164256)
  41. International Committee on Taxonomy of Viruses. ICTV Taxonomy history: Human immunodeficiency virus 2. Disponível em: [https://talk.ictvonline.org/taxonomy/p/taxonomy-history?taxnode\\_id=20164257](https://talk.ictvonline.org/taxonomy/p/taxonomy-history?taxnode_id=20164257)
  42. Shaw GM, Hunter E. HIV transmission. *Cold Spring Harb Perspect Med*. 1 de Novembro de 2012;2(11):1–23. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23043157>
  43. Fearon M. The laboratory diagnosis of HIV infections. *Can J Infect Dis Med Microbiol = J Can des Mal Infect la Microbiol medicale*. Janeiro de 2005;16(1):26–30. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/18159524>
  44. Daar ES, Pilcher CD, Hecht FM. Clinical presentation and diagnosis of primary HIV-1 infection. *Curr Opin HIV AIDS*. Janeiro de 2008;3(1):10–5. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19372938>
  45. Zulfiqar HF, Javed A, Sumbal, Afroze B, Ali Q, Akbar K, et al. HIV Diagnosis and Treatment through Advanced Technologies. *Front public Heal*. 2017;5:32.

- Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/28326304>
46. Gonzalo-gil E, Ikediobi U, Sutton RE. Mechanisms of Virologic Control and Clinical Characteristics of HIV + Elite / Viremic Controllers. *Yale J Biol Med.* 2017;90:245–59. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/28656011>
  47. Geijtenbeek TB, Kwon DS, Torensma R, van Vliet SJ, van Duijnhoven GC, Middel J, et al. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell.* 3 de Março de 2000;100(5):587–97. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/10721995>
  48. Kwon DS, Gregorio G, Bitton N, Hendrickson WA, Littman DR. DC-SIGN-mediated internalization of HIV is required for trans-enhancement of T cell infection. *Immunity.* Janeiro de 2002;16(1):135–44. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/11825572>
  49. Azevedo-Pereira JM, Santos-Costa Q. HIV Interaction With Human Host: HIV-2 As a Model of a Less Virulent Infection. *AIDS Rev.* 18(1):44–53. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/26936760>
  50. Nyamweya S, Hegedus A, Jaye A, Rowland-Jones S, Flanagan KL, Macallan DC. Comparing HIV-1 and HIV-2 infection: Lessons for viral immunopathogenesis. *Rev Med Virol.* Julho de 2013;23(4):221–40. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23444290>
  51. Visseaux B, Damond F, Matheron S, Descamps D, Charpentier C. Hiv-2 molecular epidemiology. *Infect Genet Evol.* Dezembro de 2016;46:233–40. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/27530215>
  52. Maartens G, Celum C, Lewin SR. HIV infection: Epidemiology, pathogenesis, treatment, and prevention. *Lancet.* 2014;384(9939):258–71. Disponível em: [http://dx.doi.org/10.1016/S0140-6736\(14\)60164-1](http://dx.doi.org/10.1016/S0140-6736(14)60164-1)
  53. Miguel J. HIV-2 Interaction with Target Cell Receptors, or Why HIV-2 is Less Pathogenic than HIV-1. Em: *Current Perspectives in HIV Infection.* InTech;

2013. p. 411–45. Disponível em: <http://www.intechopen.com/books/current-perspectives-in-hiv-infection/hiv-2-interaction-with-target-cell-receptors-or-why-hiv-2-is-less-pathogenic-than-hiv-1>
54. Lever AML. Molecular Biology of HIV-2. Em: Encyclopedia of AIDS. New York, NY: Springer New York; 2013. p. 1–11. Disponível em: [http://link.springer.com/10.1007/978-1-4614-9610-6\\_43-1](http://link.springer.com/10.1007/978-1-4614-9610-6_43-1)
55. Fanales-Belasio E, Raimondo M, Suligoi B, Buttò S. HIV virology and pathogenetic mechanisms of infection: a brief overview. doi.org. 2010;46(1):5–14. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/20348614>
56. Greene WC, Peterlin BM. Charting HIV's remarkable voyage through the cell: Basic science as a passport to future therapy. Nat Med. 2002;8(7):673–80. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/12091904>
57. Frankel AD, Young JAT. HIV-1: Fifteen Proteins and an RNA. Annu Rev Biochem. 28 de Junho de 1998;67(1):1–25. Disponível em: <http://www.annualreviews.org/doi/10.1146/annurev.biochem.67.1.1>
58. Nisole S, Saïb A. Early steps of retrovirus replicative cycle. Retrovirology. 14 de Maio de 2004;1(1):1–20. Disponível em: <http://retrovirology.biomedcentral.com/articles/10.1186/1742-4690-1-9>
59. Swanson CM, Malim MH. SnapShot: HIV-1 Proteins. Cell. 16 de Maio de 2008;133(4):742-742.e1. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/18485880>
60. Freed EO. HIV-1 replication. Somat Cell Mol Genet. 2001;26(1–6):13–33. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/12465460>
61. Bukrinsky M. A hard way to the nucleus. Mol Med. 2004;10(1–6):1–5. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/15502876>
62. Suzuki Y, Craigie R. The road to chromatin — nuclear entry of retroviruses. Nat Rev Microbiol. Março de 2007;5(3):187–96. Disponível em:

<http://www.ncbi.nlm.nih.gov/pubmed/17304248>

63. Levin A, Rosenbluh J, Hayouka Z, Friedler A, Loyter A. Integration of HIV-1 DNA is regulated by interplay between viral rev and cellular LEDGF/p75 proteins. *Mol Med*. 2010;16(1–2):34–44. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19855849>
64. Wu Y, Marsh JW. Selective Transcription and Modulation of Resting T Cell Activity by Preintegrated HIV DNA. *Science* (80- ). 24 de Agosto de 2001;293(5534):1503–6. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/11520990>
65. Jablonski JA, Buratti E, Stuani C, Caputi M. The secondary structure of the human immunodeficiency virus type 1 transcript modulates viral splicing and infectivity. *J Virol*. Agosto de 2008;82(16):8038–50. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/18550660>
66. Pollard VW, Malim MH. THE HIV-1 REV PROTEIN. *Annu Rev Microbiol*. Outubro de 1998;52(1):491–532. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/9891806>
67. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, et al. HIV-1 Nomenclature Proposal - A Reference Guide to HIV-1 Classifications. Los Alamos National Laboratory; 2000. p. 492–505. Disponível em: <https://www.hiv.lanl.gov/content/immunology/pdf/1999/4/nomenclature.pdf>
68. Ayoub A, Souquière S, Njinku B, Martin PM, Müller-Trutwin MC, Roques P, et al. HIV-1 group N among HIV-1-seropositive individuals in Cameroon. *AIDS*. 10 de Novembro de 2000;14(16):2623–5. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/11101082>
69. Delaugerre C, De Oliveira F, Lascoux-Combe C, Plantier J-C, Simon F. HIV-1 group N: travelling beyond Cameroon. *Lancet*. 26 de Novembro de 2011;378(9806):1894. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/22118443>

70. Plantier J-C, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemée V, et al. A new human immunodeficiency virus derived from gorillas. *Nat Med*. 2 de Agosto de 2009;15(8):871–2. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19648927>
71. Vallari A, Holzmayer V, Harris B, Yamaguchi J, Ngansop C, Makamche F, et al. Confirmation of putative HIV-1 group P in Cameroon. *J Virol*. 1 de Fevereiro de 2011;85(3):1403–7. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21084486>
72. Peeters M, Gueye A, Mboup S, Bibollet-Ruche F, Ekaza E, Mulanga C, et al. Geographical distribution of HIV-1 group O viruses in Africa. *AIDS*. 15 de Março de 1997;11(4):493–8. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/9084797>
73. Kluge SF, Mack K, Iyer SS, Pujol FM, Heigle A, Learn GH, et al. Nef proteins of epidemic HIV-1 group O strains antagonize human tetherin. *Cell Host Microbe*. 12 de Novembro de 2014;16(5):639–50. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/25525794>
74. Lau KA, Wong JLL. Current Trends of HIV Recombination Worldwide. *Infect Dis Rep*. 6 de Junho de 2013;5 (Suppl 1(e4)):15–20. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/24470968>
75. Takebe Y, Uenishi R, Li X. Global Molecular Epidemiology of HIV: Understanding the Genesis of AIDS Pandemic. *Adv Pharmacol*. 1 de Janeiro de 2008;56:1–25. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1054358907560011>
76. Foley BT, Leitner T, Paraskevis D, Peeters M. Primate immunodeficiency virus classification and nomenclature: Review. *Infect Genet Evol*. 2016;46:150–8. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/27789390>
77. Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM. The challenge of HIV-1 subtype diversity. *N Engl J Med*. 10 de Abril de 2008;358(15):1590–602.

- Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/18403767>
78. Buonaguro L, Tornesello ML, Buonaguro FM. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J Virol.* Outubro de 2007;81(19):10209–19. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/17634242>
  79. Fang G, Weiser B, Kuiken C, Philpott SM, Rowland-Jones S, Plummer F, et al. Recombination following superinfection by HIV-1. *AIDS.* 23 de Janeiro de 2004;18(2):153–9. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/15075531>
  80. Los Alamos National Laboratory. HIV Sequence Database: Nomenclature Overview. Disponível em: <https://www.hiv.lanl.gov/content/sequence/HelpDocs/subtypes-more.html>
  81. Camacho R. The significance of subtype-related genetic variability: controversies and unanswered questions. *Antiretroviral Resistance in Clinical Practice.* Mediscript; 2006. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21249770>
  82. Los Alamos National Laboratory. HIV Circulating Recombinant Forms (CRFs). Disponível em: <https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html>
  83. Peeters M, D’Arc M, Delaporte E. Origin and diversity of human retroviruses. *AIDS Rev.* 2014;16(1):23–34. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/24584106>
  84. Hemelaar J, Gouws E, Ghys PD, Osmanov S, WHO-UNAIDS Network for HIV Isolation and Characterisation W-UN for HI and. Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS.* 13 de Março de 2011;25(5):679–89. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21297424>
  85. Los Alamos National Laboratory. Distribution of all HIV-1 sequences.

Disponível em:

<https://www.hiv.lanl.gov/components/sequence/HIV/geo/geo.comp>

86. Hemelaar J. Implications of HIV diversity for the HIV-1 pandemic. *J Infect.* Maio de 2013;66(5):391–400. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/23103289>
87. Beloukas A, Psarris A, Giannelou P, Kostaki E, Hatzakis A, Paraskevis D. Molecular epidemiology of HIV-1 infection in Europe: An overview. *Infect Genet Evol.* Dezembro de 2016;46:180–9. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/27321440>
88. Thomson MM, Nájera R. Molecular epidemiology of HIV-1 variants in the global AIDS pandemic: an update. *AIDS Rev.* 7(4):210–24. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/16425961>
89. Lihana RW, Ssemwanga D, Abimiku A, Ndembu N. Update on HIV-1 diversity in Africa: a decade in review. *AIDS Rev.* 2012;14(2):83–100. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/22627605>
90. Smyth RP, Davenport MP, Mak J. The origin of genetic diversity in HIV-1. *Virus Res.* 2012;169(2):415–29. Disponível em:  
<http://dx.doi.org/10.1016/j.virusres.2012.06.015>
91. Santoro MM, Perno CF. HIV-1 Genetic Variability and Clinical Implications. *ISRN Microbiol.* 2013;2013(August):1–20. Disponível em:  
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3703378&tool=pmcentrez&rendertype=abstract>
92. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science.* 15 de Março de 1996;271(5255):1582–6. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/8599114>
93. Sosin D V, Tchurikov NA. Molecular Mechanisms of HIV-1 Genetic Diversity. *Mol Biol.* 2017;51(4):483–95. Disponível em:

<https://link.springer.com/article/10.1134%2FS0026893317030153>

94. Jetzt AE, Yu H, Klarmann GJ, Ron Y, Preston BD, Dougherty JP. High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J Virol*. Fevereiro de 2000;74(3):1234–40. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/10627533>
95. Lal RB, Chakrabarti S, Yang C. Impact of genetic diversity of HIV-1 on diagnosis, antiretroviral therapy & vaccine development. *Indian J Med Res*. Abril de 2005;121(4):287–314. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/15817945>
96. Aldrich C, Hemelaar J. Global HIV-1 diversity surveillance. *Trends Mol Med*. 2012;18(12):691–4. Disponível em: <http://dx.doi.org/10.1016/j.molmed.2012.06.004>
97. Foley B, Leitner T, Apetrei C, Hahn B, Mizrachi I, Mullins J, et al. HIV Sequence Compendium 2017. New Mexico: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory; 2017. Disponível em: <https://www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/2017/sequence2017.pdf>
98. Hill M, Tachedjian G, Mak J. The packaging and maturation of the HIV-1 Pol proteins. *Curr HIV Res*. Janeiro de 2005;3(1):73–85. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/15638725>
99. Cassan M, Delaunay N, Vaquero C, Rousset J-P. Translational Frameshifting at the gag-pol Junction of Human Immunodeficiency Virus Type 1 Is Not Increased in Infected T-Lymphoid Cells. *J Virol*. 1994;68(3):1501–8. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC236606/pdf/jvirol00012-0249.pdf>
100. Hung M, Patel P, Davis S, Green SR. Importance of Ribosomal Frameshifting for Human Immunodeficiency Virus Type 1 Particle Assembly and Replication. *J Virol*. 1998;72(6):4819–24. Disponível em:

<http://jvi.asm.org/content/72/6/4819.full.pdf>

101. Konvalinka J, Kräusslich H-G, Müller B. Retroviral proteases and their roles in virion maturation. *Virology*. 1 de Maio de 2015;479–480:403–17. Disponível em:  
<https://www.sciencedirect.com/science/article/pii/S0042682215001579?via%3Di> hub
102. Todd S, Anderson C, Jolly DJ, Craik CS. HIV protease as a target for retrovirus vector-mediated gene therapy. *Biochim Biophys Acta*. 7 de Março de 2000;1477(1–2):168–88. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/10708857>
103. Wlodawer A, Gustchina A. Structural and biochemical studies of retroviral proteases. *Biochim Biophys Acta*. 7 de Março de 2000;1477(1–2):16–34. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/10708846>
104. Hou T, McLaughlin WA, Wang W. Evaluating the potency of HIV-1 protease drugs to combat resistance. *Proteins*. 15 de Maio de 2008;71(3):1163–74. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/18004760>
105. Hornak V, Simmerling C. Targeting structural flexibility in HIV-1 protease inhibitor binding. *Drug Discov Today*. Fevereiro de 2007;12(3–4):132–8. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/17275733>
106. Kohl NE, Emini EA, Schleif WA, Davis LJ, Heimbach JC, Dixon RA, et al. Active human immunodeficiency virus protease is required for viral infectivity. *Proc Natl Acad Sci U S A*. Julho de 1988;85(13):4686–90. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/3290901>
107. Sarafianos SG, Marchand B, Das K, Himmel D, Parniak MA, Hughes SH, et al. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *J Mol Biol*. 23 de Janeiro de 2009;385(3):693. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19022262>
108. Sluis-Cremer N, Kempner E, Parniak MA. Structure-activity relationships in

- HIV-1 reverse transcriptase revealed by radiation target analysis. *Protein Sci.* Setembro de 2003;12(9):2081–6. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/12931006>
109. Sluis-Cremer N, Temiz NA, Bahar I. Conformational changes in HIV-1 reverse transcriptase induced by nonnucleoside reverse transcriptase inhibitor binding. *Curr HIV Res.* Outubro de 2004;2(4):323–32. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/15544453>
110. Basavapathruni A, Anderson KS. Reverse transcription of the HIV-1 pandemic. *FASEB J.* 20 de Julho de 2007;21(14):3795–808. Disponível em:  
<http://www.fasebj.org/cgi/doi/10.1096/fj.07-8697rev>
111. Yu Q, Ottmann M, Pechoux C, Le Grice S, Darlix JL. Mutations in the primer grip of human immunodeficiency virus type 1 reverse transcriptase impair proviral DNA synthesis and virion maturation. *J Virol.* Setembro de 1998;72(9):7676–80. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/9696874>
112. Gutiérrez-Rivas M, Menéndez-Arias L. A mutation in the primer grip region of HIV-1 reverse transcriptase that confers reduced fidelity of DNA synthesis. *Nucleic Acids Res.* 15 de Dezembro de 2001;29(24):4963–72. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/11812826>
113. Ghosh M, Jacques PS, Rodgers DW, Ottman M, Darlix J-L, Le Grice SFJ. Alterations to the Primer Grip of p66 HIV-1 Reverse Transcriptase and Their Consequences for Template-Primer Utilization. *Biochemistry.* 1996;35(26):8553–62. Disponível em:  
<http://pubs.acs.org/doi/pdf/10.1021/bi952773j>
114. Mulky A, Vu C, Conway JA, Hughes SH, Kappes JC. Analysis of Amino Acids in the  $\beta 7$ – $\beta 8$  Loop of Human Immunodeficiency Virus Type 1 Reverse Transcriptase for their Role in Virus Replication. *J Mol Biol.* 2 de Fevereiro de 2007;365(5):1368–78. Disponível em:  
<http://www.sciencedirect.com/science/article/pii/S0022283606015178?via%3Dih>

ub

115. Roquebert B, Marcelin A-G. The involvement of HIV-1 RNase H in resistance to nucleoside analogues. *J Antimicrob Chemother.* 6 de Março de 2008;61(5):973–5. Disponível em: <https://academic.oup.com/jac/article-lookup/doi/10.1093/jac/dkn060>
  
116. Mulky A, Sarafianos SG, Jia Y, Arnold E, Kappes JC. Identification of Amino Acid Residues in the Human Immunodeficiency Virus Type-1 Reverse Transcriptase Tryptophan-repeat Motif that are Required for Subunit Interaction Using Infectious Virions. *J Mol Biol.* 17 de Junho de 2005;349(4):673–84. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0022283605003487?via%3Dihub>
  
117. Tachedjian G, Aronson H-EG, de los Santos M, Seehra J, McCoy JM, Goff SP. Role of residues in the tryptophan repeat motif for HIV-1 reverse transcriptase dimerization. *J Mol Biol.* 14 de Fevereiro de 2003;326(2):381–96. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/12559908>
  
118. Tachedjian G, Radzio J, Sluis-Cremer N. Relationship between enzyme activity and dimeric structure of recombinant HIV-1 reverse transcriptase. *Proteins Struct Funct Bioinforma.* 25 de Abril de 2005;60(1):5–13. Disponível em: <http://doi.wiley.com/10.1002/prot.20480>
  
119. Wapling J, Moore KL, Sonza S, Mak J, Tachedjian G. Mutations that abrogate human immunodeficiency virus type 1 reverse transcriptase dimerization affect maturation of the reverse transcriptase heterodimer. *J Virol.* Agosto de 2005;79(16):10247–57. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/16051818>
  
120. Arts EJ, Hazuda DJ. HIV-1 antiretroviral drug therapy. *Cold Spring Harb Perspect Med.* Abril de 2012;2(4):a007161. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/22474613>

121. Kulpa DA, Chomont N. HIV persistence in the setting of antiretroviral therapy: when, where and how does HIV hide? *J virus Erad.* Abril de 2015;1(2):59–66. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/26448966>
122. Franco RA, Saag MS. When to start antiretroviral therapy: as soon as possible. *BMC Med.* 14 de Dezembro de 2013;11(1):147. Disponível em: <http://bmcmmedicine.biomedcentral.com/articles/10.1186/1741-7015-11-147>
123. World Health Organization. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. 2015. Disponível em: [https://apps.who.int/iris/bitstream/handle/10665/186275/9789241509565\\_eng.pdf;jsessionid=3F2275A5EA52ACD1D24F4D7213E1B203?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/186275/9789241509565_eng.pdf;jsessionid=3F2275A5EA52ACD1D24F4D7213E1B203?sequence=1)
124. Department of Health and Human Services. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the Use of Antiretroviral Agents in Adults and Adolescents with HIV. 2017. Disponível em: <http://aidsinfo.nih.gov/contentfiles/lvguidelines/AdultandAdolescentGL.pdf>
125. Meadows DC, Gervay-Hague J. Current Developments in HIV Chemotherapy. *ChemMedChem.* 16 de Janeiro de 2006;1(1):16–29. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/16892329>
126. Cortez KJ, Maldarelli F. Clinical management of HIV drug resistance. *Viruses.* 2011;3(4):347–78. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21994737>
127. Götte M. The distinct contributions of fitness and genetic barrier to the development of antiviral drug resistance. *Curr Opin Virol.* 1 de Outubro de 2012;2(5):644–50. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1879625712001307?via%3DiHub#bib0175>
128. Clutter DS, Jordan MR, Bertagnolio S, Shafer RW. HIV-1 drug resistance and resistance testing. *Infect Genet Evol.* Dezembro de 2016;46:292–307. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S1567134816303690>

129. Skar H, Hedskog C, Albert J. HIV-1 evolution in relation to molecular epidemiology and antiretroviral resistance. *Ann N Y Acad Sci*. 2011;1230(1):108–18. Disponível em: <https://nyaspubs.onlinelibrary.wiley.com/doi/epdf/10.1111/j.1749-6632.2011.06128.x>
130. Wallis CL, Godfrey C, Fitzgibbon JE, Mellors JW. Key Factors Influencing the Emergence of Human Immunodeficiency Virus Drug Resistance in Low- and Middle-Income Countries. *J Infect Dis*. 2017;216(9):S851–6. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/29207000>
131. Stella-Ascariz N, Arribas JR, Paredes R, Li JZ. The Role of HIV-1 Drug-Resistant Minority Variants in Treatment Failure. *J Infect Dis*. 2017;216(9):S847–50. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/29207001>
132. Geretti AM. Antiretroviral Resistance in Clinical Practice. *Antiretroviral Resistance in Clinical Practice*. Mediscript; 2006. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21249765>
133. Broder S. The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic. *Antiviral Res*. Janeiro de 2010;85(1):1–18. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/20018391>
134. Shafer RW. Genotypic testing for human immunodeficiency virus type 1 drug resistance. *Clin Microbiol Rev*. Abril de 2002;15(2):247–77. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/11932232>
135. Liu TF, Shafer RW. Web resources for HIV type 1 genotypic-resistance test interpretation. *Clin Infect Dis*. 1 de Junho de 2006;42(11):1608–18. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/16652319>
136. Menéndez-Arias L. Molecular basis of human immunodeficiency virus drug resistance: An update. *Antiviral Res*. Janeiro de 2010;85(1):210–31. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19616029>

137. De Clercq E. Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV. *Int J Antimicrob Agents*. Abril de 2009;33(4):307–20. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19108994>
138. Tapp Alter K. The protease inhibitors. *Prim Care Update Ob Gyns*. Março de 2001;8(2):59–64. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/11246029>
139. U.S. Food & Drug Administration. HIV/AIDS Treatment - Antiretroviral drugs used in the treatment of HIV infection. Disponível em: <https://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm>
140. Hull MW, Montaner JSG. Ritonavir-boosted protease inhibitors in HIV therapy. *Ann Med*. 18 de Agosto de 2011;43(5):375–88. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21501034>
141. Sahali S, Chaix M-L, Delfraissy J-F, Ghosn J. Ritonavir-boosted protease inhibitor monotherapy for the treatment of HIV-1 infection. *AIDS Rev*. 10(1):4–14. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/18385776>
142. Ghosh AK, Osswald HL, Prato G. Recent Progress in the Development of HIV-1 Protease Inhibitors for the Treatment of HIV/AIDS. *J Med Chem*. 9 de Junho de 2016;59(11):5172–208. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/26799988>
143. U.S. Food & Drug Administration. Drugs@FDA: FDA Approved Drug Products - Doravirine. Disponível em: <https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&ApplNo=210806>
144. Deeks ED. Doravirine: First Global Approval. *Drugs*. 19 de Outubro de 2018;78(15):1643–50. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/30341683>
145. Stanford University. Stanford HIVdb FAQ: Questions and Answers. 2008. Disponível em: [https://hivdb.stanford.edu/pages/FAQ/FAQ\\_answers.html](https://hivdb.stanford.edu/pages/FAQ/FAQ_answers.html)

146. Weber IT, Kneller DW, Wong-Sam A. Highly resistant HIV-1 proteases and strategies for their inhibition. *Future Med Chem.* 2015;7(8):1023–38. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/26062399>
147. Shafer RW, Schapiro JM. HIV-1 drug resistance mutations: an updated framework for the second decade of HAART. *AIDS Rev.* 2008;10(2):67–84. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/18615118>
148. Weber IT, Harrison RW. Tackling the problem of HIV drug resistance. *Postepy Biochem.* 62(3):273–9. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/28132481>
149. Wensing AM, Calvez V, Günthard HF, Johnson VA, Paredes R, Pillay D, et al. Special Contribution - 2017 Update of the Drug Resistance Mutations in HIV-1. *Top Antivir Med.* 2017;24(4). Disponível em: <https://www.iasusa.org/wp-content/uploads/2018/01/2017-drug-resistance-mutations-figures-and-article.pdf>
150. Sluis-Cremer N, Wainberg MA, Schinazi RF. Resistance to reverse transcriptase inhibitors used in the treatment and prevention of HIV-1 infection. *Future Microbiol.* 2015;10(11):1773–82. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/26517190>
151. The World Bank. Angola | Data. Disponível em: <https://data.worldbank.org/country/angola>
152. Pineda-Peña A-C, Varanda J, Sousa JD, Theys K, Bártolo I, Leitner T, et al. On the contribution of Angola to the initial spread of HIV-1. *Infect Genet Evol.* Dezembro de 2016;46:219–22. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/27521160>
153. Afonso JM, Morgado MG, Bello G. Evidence of multiple introductions of HIV-1 subtype C in Angola. *Infect Genet Evol.* Outubro de 2012;12(7):1458–65. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/22634597>
154. Bártolo I, Rocha C, Bartolomeu J, Gama A, Marcelino R, Fonseca M, et al. Highly divergent subtypes and new recombinant forms prevail in the HIV/AIDS

- epidemic in Angola: new insights into the origins of the AIDS pandemic. *Infect Genet Evol.* Julho de 2009;9(4):672–82. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/18562253>
155. Castelbranco EPAF, Souza E da S, Cavalcanti AMS, Martins AN, Alencar LCA de, Tanuri A. Frequency of Primary Resistance to Antiretroviral Drugs and Genetic Variability of HIV-1 Among Infected Pregnant Women Recently Diagnosed in Luanda-Angola. *AIDS Res Hum Retroviruses.* 2010;26(12):1313–6. Disponível em:  
[https://hivdb.stanford.edu/surveillance/refs/Castelbranco\\_2010\\_AIDS Res Hum Retroviruses.pdf](https://hivdb.stanford.edu/surveillance/refs/Castelbranco_2010_AIDS Res Hum Retroviruses.pdf)
156. UNAIDS. Country Factsheets Angola 2017. 2018. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/28715460>
157. Bártolo I, Epalanga M, Bartolomeu J, Fonseca M, Mendes A, Gama A, et al. High Genetic Diversity of Human Immunodeficiency Virus Type 1 in Angola. *AIDS Res Hum Retroviruses.* Abril de 2005;21(4):306–10. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/15943573>
158. Abecasis A, Paraskevis D, Epalanga M, Fonseca M, Burity F, Bartolomeu J, et al. HIV-1 genetic variants circulation in the North of Angola. *Infect Genet Evol.* 2005;5:231–7. Disponível em:  
<http://www.sciencedirect.com/science/article/pii/S1567134804001182>
159. Garrido C, Zahonero N, Fernández D, Serrano D, Silva AR, Ferraria N, et al. Subtype variability, virological response and drug resistance assessed on dried blood spots collected from HIV patients on antiretroviral therapy in Angola. *J Antimicrob Chemother.* 2008;61(3):694–8. Disponível em:  
<https://academic.oup.com/jac/article-lookup/doi/10.1093/jac/dkm515>
160. Bártolo I, Rocha C, Bartolomeu J, Gama A, Fonseca M, Mendes A, et al. Antiretroviral drug resistance surveillance among treatment-naïve human immunodeficiency virus type 1-infected individuals in Angola: evidence for low level of transmitted drug resistance. *Antimicrob Agents Chemother.* Julho de

- 2009;53(7):3156–8. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/19433560>
161. Afonso JM, Bello G, Guimarães ML, Sojka M, Morgado MG. HIV-1 Genetic Diversity and Transmitted Drug Resistance Mutations among Patients from the North, Central and South Regions of Angola. *PLoS One*. 2012;7(8):1–11. Disponível em:  
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0042996>
162. Bártolo I, Zakovic S, Martin F, Palladino C, Carvalho P, Camacho R, et al. HIV-1 diversity, transmission dynamics and primary drug resistance in Angola. *PLoS One*. 2014;9(12):1–17. Disponível em:  
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0113626>
163. Bártolo I, Calado R, Borrego P, Leitner T, Taveira N. Rare HIV-1 Subtype J Genomes and a New H/U/CRF02\_AG Recombinant Genome Suggests an Ancient Origin of HIV-1 in Angola. *AIDS Res Hum Retroviruses*. 2016;32(8):822–8. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/27098898>
164. Fortin M-F, Côte J, Filion F. Fundamentos e Etapas do Processo de Investigação. Lusodidacta; 2009.
165. Healgen Scientific. HIV Rapid Test: Healgen. 2017. Disponível em:  
<http://www.healgen.com/infectious-disease/hiv/>
166. Türklab. Anti-HIV 1/2 TEST | TURKLAB. 2017. Disponível em:  
<https://www.turklab.com.tr/hiv-tests/anti-hiv-1-2-test>
167. GE Healthcare. Whatman Indicating FTA Elute micro cards. 2017. Disponível em: [https://www.gelifesciences.com/shop/forensics/cards-kits-and-collectors/whatman-indicating-fta-elute-micro-cards-\(one-sample-area\)-p-00636](https://www.gelifesciences.com/shop/forensics/cards-kits-and-collectors/whatman-indicating-fta-elute-micro-cards-(one-sample-area)-p-00636)
168. Human Diagnostics. Hexagon HIV. 2017. Disponível em:  
[https://www.human.de/fileadmin/content/flyer/en/981565\\_Hexagon\\_HIV\\_EN.pdf](https://www.human.de/fileadmin/content/flyer/en/981565_Hexagon_HIV_EN.pdf)

169. Human Diagnostics. Hexagon HIV - Dispositivo Médico para Diagnóstico In vitro. 2017.
170. GE Healthcare. Whatman™ FTA Elute™. 2011. p. 4. Disponível em: <https://cdn.gelifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=17044>
171. GE Healthcare. Indicating FTA Elute Micro Card - Frequent Asked Questions. 2017. Disponível em: [https://www.gelifesciences.com/shop/whatman-indicating-fta-elute-micro-cards-\(one-sample-area\)-p-00636/28421907](https://www.gelifesciences.com/shop/whatman-indicating-fta-elute-micro-cards-(one-sample-area)-p-00636/28421907)
172. Videira e Castro S. Diversidade Genética e Resistência aos Anti-Retrovirais Inibidores Enzimáticos de Vírus da Imunodeficiência Humana Tipo 1 (HIV-1) de uma População de Toxicodependentes por Via Endovenosa da Grande Lisboa. Instituto de Higiene e Medicina Tropical; 2010.
173. Steegen K, Demecheleer E, De Cabooter N, Nges D, Temmerman M, Ndumbe P, et al. A sensitive in-house RT-PCR genotyping system for combined detection of plasma HIV-1 and assessment of drug resistance. J Virol Methods. 2006;133(2):137–45.
174. Sérgio Simões CS. Caracterização de polimorfismos no gene pol associados a resistência aos inibidores da protease utilizados em regimes terapêuticos de prevenção da transmissão mãe-filho do vírus da imunodeficiência humana tipo 1. Instituto de Higiene e Medicina Tropical; 2016.
175. NZY Tech. Supreme NZYTaQ II 2 × Green Master Mix. 2017. p. 2. Disponível em: <https://www.nzytech.com/products-services/pcr-master-mixes/mb360/>
176. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. Julho de 1987;4(4):406–25. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/3447015>
177. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. Dezembro de 1980;16(2):111–20. Disponível em:

- <http://www.ncbi.nlm.nih.gov/pubmed/7463489>
178. Günthard HF, Calvez V, Paredes R, Pillay D, Shafer RW, Wensing AM, et al. Human Immunodeficiency Virus Drug Resistance: 2018 Recommendations of the International Antiviral Society-USA Panel. *Clin Infect Dis.* 7 de Janeiro de 2019;68(2):177–87. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/30052811>
179. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H. Drug Resistance Mutations for Surveillance of Transmitted HIV-1 Drug-Resistance: 2009 Update. *PLoS One.* 2009;4(3):1–8. Disponível em: <http://hivdb.stanford.edu/cgi-bin/AgMutPrev.cgi>
180. Shafer RW, Rhee S-Y, Pillay D, Miller V, Sandstrom P, Schapiro JM, et al. HIV-1 protease and reverse transcriptase mutations for drug resistance surveillance. *AIDS.* 11 de Janeiro de 2007;21(2):215–23. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/17197813>
181. Peeters M, Jung M, Ayoub A. The origin and molecular epidemiology of HIV. *Expert Rev Anti Infect Ther.* 10 de Setembro de 2013;11(9):885–96. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/24011334>
182. Zazzi M, Hu H, Prosperi M. The global burden of HIV-1 drug resistance in the past 20 years. *PeerJ.* 2018;6(e4848):1–16. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5971836/>
183. Gupta RK, Gregson J, Parkin N, Haile-Selassie H, Tanuri A, Andrade Forero L, et al. HIV-1 drug resistance before initiation or re-initiation of first-line antiretroviral therapy in low-income and middle-income countries: a systematic review and meta-regression analysis. *Lancet Infect Dis.* Março de 2018;18(3):346–55. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/29198909>
184. Denniff P, Spooner N. Effect of storage conditions on the weight and appearance of dried blood spot samples on various cellulose-based substrates. *Bioanalysis.*

- Novembro de 2010;2(11):1817–22. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/21083489>
185. Hearps AC, Ryan CE, Morris LM, Plate MM, Greengrass V, Crowe SM. Stability of dried blood spots for HIV-1 drug resistance analysis. *Curr HIV Res.* Março de 2010;8(2):134–40. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/20163343>
186. Pineda-Peña A-C, Faria NR, Imbrechts S, Libin P, Abecasis AB, Deforche K, et al. Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: Performance evaluation of the new REGA version 3 and seven other tools. *Infect Genet Evol.* Outubro de 2013;19:337–48. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/23660484>
187. Fabeni L, Berno G, Fokam J, Bertoli A, Alteri C, Gori C, et al. Comparative Evaluation of Subtyping Tools for Surveillance of Newly Emerging HIV-1 Strains. *J Clin Microbiol.* 2017;55(9):2827–37. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/28701420>
188. Pasquier C, Millot N, Njouom R, Sandres K, Cazabat M, Puel J, et al. HIV-1 subtyping using phylogenetic analysis of pol gene sequences. *J Virol Methods.* Maio de 2001;94(1–2):45–54. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/11337039>
189. Bello G, Passaes CP, Guimarães ML, Lorete RS, Matos Almeida SE, Medeiros RM, et al. Origin and evolutionary history of HIV-1 subtype C in Brazil. *AIDS.* 1 de Outubro de 2008;22(15):1993–2000. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/18753928>
190. Xiao P, Li J, Fu G, Zhou Y, Huan X, Yang H. Geographic Distribution and Temporal Trends of HIV-1 Subtypes through Heterosexual Transmission in China: A Systematic Review and Meta-Analysis. *Int J Environ Res Public Health.* 2017;14(7):1–18. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/28737729>

191. Shen C, Craigo J, Ding M, Chen Y, Gupta P. Origin and dynamics of HIV-1 subtype C infection in India. *PLoS One*. 2011;6(10):e25956. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/22016790>
192. Han J, Liu S, Guo W, Bao Z, Wang X, Li L, et al. Development of an HIV-1 Subtype Panel in China: Isolation and Characterization of 30 HIV-1 Primary Strains Circulating in China. Yang R, editor. *PLoS One*. 27 de Maio de 2015;10(5):e0127696. Disponível em: <https://dx.plos.org/10.1371/journal.pone.0127696>
193. Bello G, Afonso JM, Morgado MG. Phylodynamics of HIV-1 subtype F1 in Angola, Brazil and Romania. *Infect Genet Evol*. Julho de 2012;12(5):1079–86. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/22484759>
194. Guimarães ML, Vicente A, Otsuki K, da Silva R, Francisco M, da Silva F, et al. Close phylogenetic relationship between Angolan and Romanian HIV-1 subtype F1 isolates. *Retrovirology*. 22 de Abril de 2009;6(1):39. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19386115>
195. Harris RS, Liddament MT. Retroviral restriction by APOBEC proteins. *Nat Rev Immunol*. Novembro de 2004;4(11):868–77. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/15516966>
196. Abecasis AB, Lemey P, Vidal N, de Oliveira T, Peeters M, Camacho R, et al. Recombination Confounds the Early Evolutionary History of Human Immunodeficiency Virus Type 1: Subtype G Is a Circulating Recombinant Form. *J Virol*. 15 de Agosto de 2007;81(16):8543–51. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/17553886>
197. Derache A, Shin H-S, Balamane M, White E, Israelski D, Klausner JD, et al. HIV drug resistance mutations in proviral DNA from a community treatment program. *PLoS One*. 2015;10(1):e0117430. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/25635815>
198. Brodin J, Zanini F, Thebo L, Lanz C, Bratt G, Neher RA, et al. Establishment

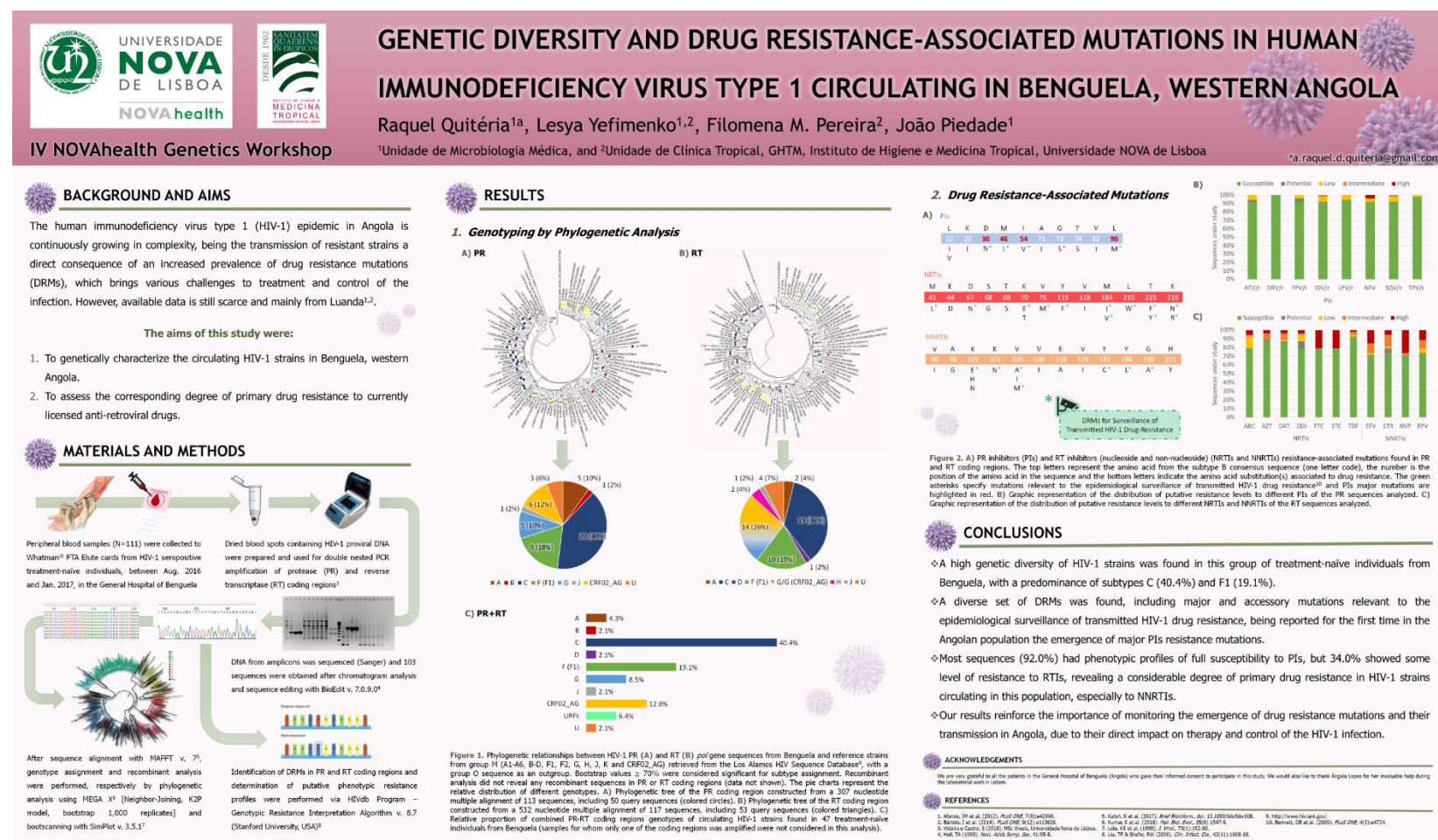
- and stability of the latent HIV-1 DNA reservoir. *Elife*. 2016;5:e18889.  
Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/27855060>
199. Banks L, Gholamin S, White E, Zijenah L, Katzenstein DA. Comparing Peripheral Blood Mononuclear Cell DNA and Circulating Plasma viral RNA pol Genotypes of Subtype C HIV-1. *J AIDS Clin Res*. Fevereiro de 2012;3(2):141–7. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23019537>
200. Novitsky V, Zahralban-Steele M, McLane MF, Moyo S, van Widenfelt E, Gaseitsiwe S, et al. Long-Range HIV Genotyping Using Viral RNA and Proviral DNA for Analysis of HIV Drug Resistance and HIV Clustering. *J Clin Microbiol*. Agosto de 2015;53(8):2581–92. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/26041893>
201. Rhee S-Y, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Res*. 1 de Janeiro de 2003;31(1):298–303. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/12520007>
202. Chang MW, Torbett BE. Accessory mutations maintain stability in drug-resistant HIV-1 protease. *J Mol Biol*. 22 de Julho de 2011;410(4):756–60. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21762813>
203. Masquelier B, Assoumou KL, Descamps D, Bocket L, Cottalorda J, Ruffault A, et al. Clinically validated mutation scores for HIV-1 resistance to fosamprenavir/ritonavir. *J Antimicrob Chemother*. 13 de Março de 2008;61(6):1362–8. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/18390885>
204. Kantor R, Katzenstein DA, Efron B, Carvalho AP, Wynhoven B, Cane P, et al. Impact of HIV-1 Subtype and Antiretroviral Therapy on Protease and Reverse Transcriptase Genotype: Results of a Global Collaboration. Ho DD, editor. *PLoS Med*. 26 de Abril de 2005;2(4):e112. Disponível em: <http://dx.plos.org/10.1371/journal.pmed.0020112>

205. Paraschiv S, Otelea D, Dinu M, Maxim D, Tinischi M. Polymorphisms and resistance mutations in the protease and reverse transcriptase genes of HIV-1 F subtype Romanian strains. *Int J Infect Dis.* 1 de Março de 2007;11(2):123–8. Disponível em:  
<https://www.sciencedirect.com/science/article/pii/S1201971206000567>
206. Dumans AT, Barreto CC, Santos AF, Arruda M, Sousa TM, Machado ES, et al. Distinct resistance mutation and polymorphism acquisition in HIV-1 protease of subtypes B and F1 from children and adult patients under virological failure. *Infect Genet Evol.* Janeiro de 2009;9(1):62–70. Disponível em:  
<http://www.ncbi.nlm.nih.gov/pubmed/18992847>
207. Gallant JE. The M184V mutation: what it does, how to prevent it, and what to do with it when it's there. *AIDS Read.* Outubro de 2006;16(10):556–9. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/17096474>
208. Whitcomb JM, Parkin NT, Chappey C, Hellmann NS, Petropoulos CJ. Broad Nucleoside Reverse-Transcriptase Inhibitor Cross-Resistance in Human Immunodeficiency Virus Type 1 Clinical Isolates. *J Infect Dis.* 1 de Outubro de 2003;188(7):992–1000. Disponível em: <https://academic.oup.com/jid/article-lookup/doi/10.1086/378281>
209. Stanford University. NNRTI Resistance Notes - HIV Drug Resistance Database. 2018 [citado 16 de Dezembro de 2018]. Disponível em:  
<https://hivdb.stanford.edu/dr-summary/resistance-notes/NNRTI/#4Rimsky2012>
210. Boyer PL, Das K, Arnold E, Hughes SH. Analysis of the Zidovudine Resistance Mutations T215Y, M41L, and L210W in HIV-1 Reverse Transcriptase. *Antimicrob Agents Chemother.* Dezembro de 2015;59(12):7184–96. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/26324274>
211. Rhee S-Y, Liu T, Ravela J, Gonzales MJ, Shafer RW. Distribution of human immunodeficiency virus type 1 protease and reverse transcriptase mutation patterns in 4,183 persons undergoing genotypic resistance testing. *Antimicrob Agents Chemother.* Agosto de 2004;48(8):3122–6. Disponível em:

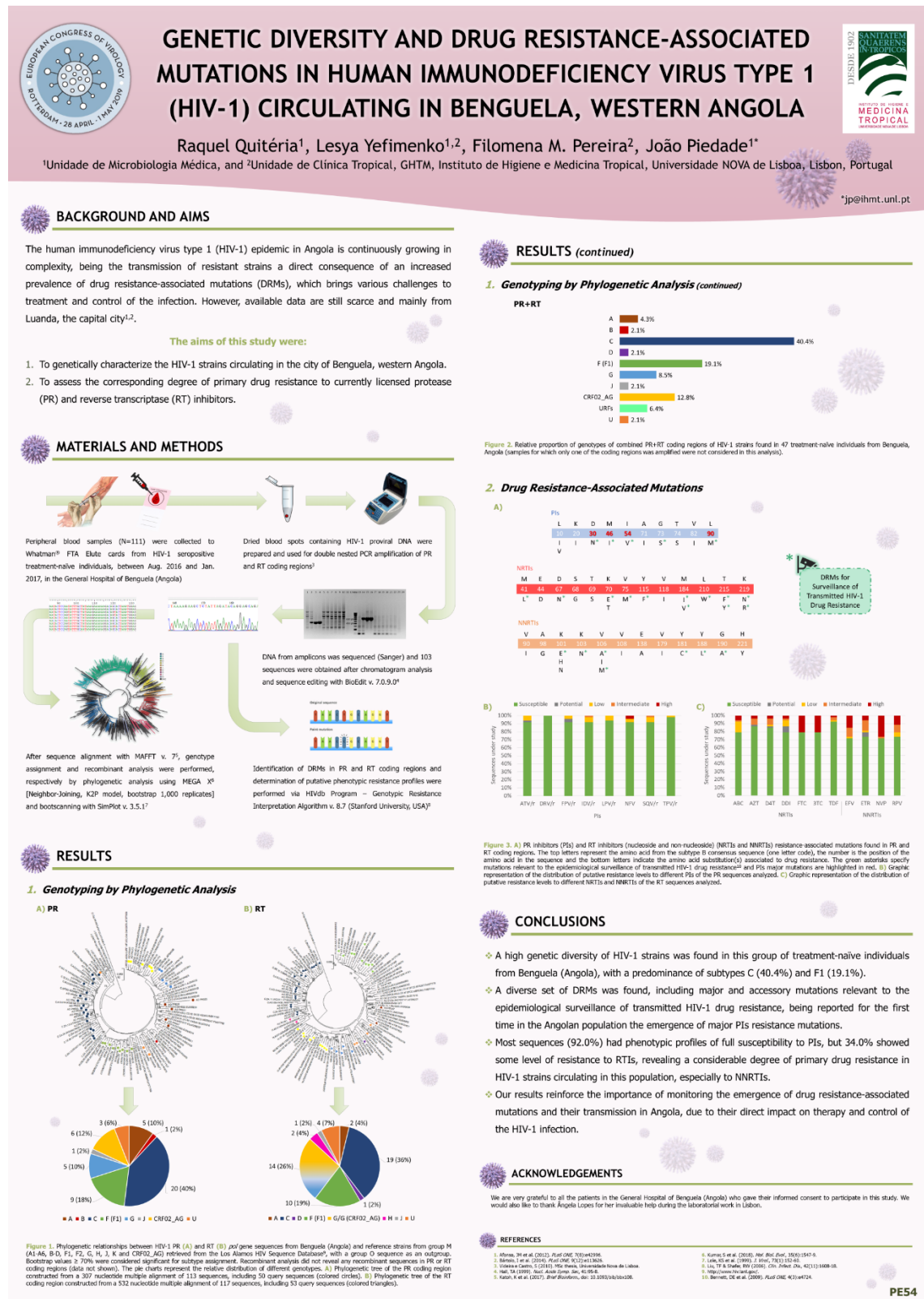
- <http://www.ncbi.nlm.nih.gov/pubmed/15273130>
212. Melikian GL, Rhee S-Y, Varghese V, Porter D, White K, Taylor J, et al. Non-nucleoside reverse transcriptase inhibitor (NNRTI) cross-resistance: implications for preclinical evaluation of novel NNRTIs and clinical genotypic resistance testing. *J Antimicrob Chemother.* Janeiro de 2014;69(1):12–20. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23934770>
213. Guo W, Li H, Zhuang D, Jiao L, Liu S, Li L, et al. Impact of Y181C and/or H221Y mutation patterns of HIV-1 reverse transcriptase on phenotypic resistance to available non-nucleoside and nucleoside inhibitors in China. *BMC Infect Dis.* 5 de Maio de 2014;14:237. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/24885612>
214. Ceccherini-Silberstein F, Svicher V, Sing T, Artese A, Santoro MM, Forbici F, et al. Characterization and structural analysis of novel mutations in human immunodeficiency virus type 1 reverse transcriptase involved in the regulation of resistance to nonnucleoside inhibitors. *J Virol.* Outubro de 2007;81(20):11507–19. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/17686836>
215. Serrano D, Furtado ML. Normas de Tratamento Antirretroviral. Luanda; 2015. Disponível em: [https://aidsfree.usaid.gov/sites/default/files/ao\\_normastratamentoarv.pdf](https://aidsfree.usaid.gov/sites/default/files/ao_normastratamentoarv.pdf)

## 6 Anexos

## Anexo 1: Póster apresentado no IV NOVAhealth Genetics Workshop



## Anexo 2: Póster apresentado no 7<sup>th</sup> European Congress of Virology



### Anexo 3: Dados demográficos e clínicos da população em estudo

Dados demográficos e clínicos dos 111 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. Na tabela é indicada a identificação (ID) das amostras, idade dos indivíduos (em anos, relativamente ao ano de 2016), género, estado civil, nível de escolaridade (“sem escolaridade”=analfabetismo ou ciclo básico de escolaridade incompleto), local de residência (município da província de Benguela) e área de residência (rural ou urbana). O símbolo “X” representa infecção com sífilis (activa ou passada) e/ou hepatite B activa (presença de AgHBs).

| ID     | Idade (anos) | Género    | Estado civil | Nível de escolaridade | Local de residência | Área de residência | Outras infecções |            |
|--------|--------------|-----------|--------------|-----------------------|---------------------|--------------------|------------------|------------|
|        |              |           |              |                       |                     |                    | Sífilis          | Hepatite B |
| AO_6   | Não sabe     | Feminino  | Viúvo        | Sem escolaridade      | Benguela            | Rural              | X                |            |
| AO_7   | 58           | Masculino | Solteiro     | Ensino secundário     | Benguela            | Rural              |                  |            |
| AO_25  | 38           | Feminino  | Casado       | Ensino básico         | Cubal               | Urbana             | X                |            |
| AO_26  | 48           | Masculino | Casado       | Ensino básico         | Baía Farta          | Urbana             | X                |            |
| AO_51  | 52           | Feminino  | Solteiro     | Ensino secundário     | Cubal               | Urbana             |                  | X          |
| AO_80  | 22           | Masculino | Solteiro     | Ensino básico         | Benguela            | Urbana             |                  |            |
| AO_84  | 52           | Feminino  | Viúvo        | Sem escolaridade      | Benguela            | Rural              |                  |            |
| AO_92  | 30           | Masculino | Casado       | Ensino superior       | Chongoroi           | Urbana             |                  |            |
| AO_119 | 26           | Feminino  | Solteiro     | Ensino secundário     | Benguela            | Urbana             |                  |            |
| AO_141 | 53           | Feminino  | Divorciado   | Ensino secundário     | Balombo             | Urbana             |                  |            |
| AO_153 | 31           | Masculino | Casado       | Ensino superior       | Ganda               | Urbana             |                  |            |
| AO_202 | 46           | Masculino | Divorciado   | Ensino básico         | Benguela            | Urbana             | X                |            |
| AO_330 | 24           | Masculino | Solteiro     | Ensino secundário     | Lobito              | Urbana             |                  |            |
| AO_380 | 10           | Feminino  | Solteiro     | Sem escolaridade      | Benguela            | Rural              |                  |            |

| ID     | Idade (anos) | Género    | Estado civil | Nível de escolaridade | Local de residência | Área de residência | Outras infecções |            |
|--------|--------------|-----------|--------------|-----------------------|---------------------|--------------------|------------------|------------|
|        |              |           |              |                       |                     |                    | Sífilis          | Hepatite B |
| AO_381 | 51           | Masculino | Casado       | Ensino secundário     | Cubal               | Urbana             |                  |            |
| AO_386 | 55           | Feminino  | Divorciado   | Ensino secundário     | Chongoroi           | Rural              |                  |            |
| AO_440 | 52           | Feminino  | Casado       | Ensino básico         | Cubal               | Rural              |                  | X          |
| AO_441 | 0            | Masculino | Solteiro     | Sem escolaridade      | Benguela            | Rural              |                  |            |
| AO_442 | 15           | Feminino  | Solteiro     | Sem escolaridade      | Benguela            | Rural              |                  |            |
| AO_443 | 3            | Masculino | Solteiro     | Sem escolaridade      | Benguela            | Urbana             |                  |            |
| AO_444 | 32           | Masculino | Divorciado   | Ensino básico         | Benguela            | Urbana             |                  |            |
| AO_445 | 61           | Masculino | Casado       | Ensino secundário     | Benguela            | Urbana             |                  |            |
| AO_446 | 78           | Feminino  | Viúvo        | Sem escolaridade      | Cubal               | Urbana             |                  |            |
| AO_447 | 41           | Feminino  | Casado       | Ensino superior       | Benguela            | Urbana             |                  |            |
| AO_448 | 50           | Masculino | Solteiro     | Ensino básico         | Ganda               | Urbana             |                  |            |
| AO_449 | 52           | Masculino | Divorciado   | Ensino básico         | Lobito              | Urbana             |                  |            |
| AO_450 | 2            | Masculino | Solteiro     | Sem escolaridade      | Benguela            | Rural              |                  |            |
| AO_451 | 53           | Feminino  | Casado       | Ensino básico         | Benguela            | Urbana             |                  |            |
| AO_452 | 16           | Feminino  | Solteiro     | Sem escolaridade      | Lobito              | Urbana             |                  |            |
| AO_453 | 5            | Masculino | Solteiro     | Sem escolaridade      | Lobito              | Urbana             |                  |            |
| AO_454 | 24           | Masculino | Solteiro     | Ensino básico         | Benguela            | Urbana             |                  |            |
| AO_455 | 72           | Feminino  | Casado       | Ensino básico         | Catumbela           | Urbana             |                  |            |
| AO_456 | 67           | Masculino | Casado       | Sem escolaridade      | Catumbela           | Rural              |                  |            |
| AO_457 | 18           | Masculino | Solteiro     | Sem escolaridade      | Benguela            | Rural              |                  |            |

| ID     | Idade (anos) | Género    | Estado civil | Nível de escolaridade | Local de residência | Área de residência | Outras infecções |            |
|--------|--------------|-----------|--------------|-----------------------|---------------------|--------------------|------------------|------------|
|        |              |           |              |                       |                     |                    | Sífilis          | Hepatite B |
| AO_458 | 1            | Masculino | Solteiro     | Sem escolaridade      | Bocoio              | Rural              |                  |            |
| AO_459 | 23           | Feminino  | Solteiro     | Ensino superior       | Benguela            | Urbana             |                  |            |
| AO_460 | 59           | Feminino  | Casado       | Ensino secundário     | Benguela            | Urbana             |                  |            |
| AO_461 | 41           | Masculino | Divorciado   | Ensino básico         | Balombo             | Rural              |                  |            |
| AO_462 | 14           | Masculino | Solteiro     | Sem escolaridade      | Ganda               | Rural              |                  |            |
| AO_463 | 66           | Feminino  | Divorciado   | Ensino básico         | Benguela            | Urbana             |                  |            |
| AO_466 | 26           | Masculino | Solteiro     | Ensino básico         | Catumbela           | Urbana             |                  |            |
| AO_467 | 64           | Feminino  | Casado       | Sem escolaridade      | Benguela            | Urbana             |                  |            |
| AO_468 | 6            | Masculino | Solteiro     | Sem escolaridade      | Baía Farta          | Urbana             |                  |            |
| AO_469 | 27           | Feminino  | Solteiro     | Ensino secundário     | Baía Farta          | Urbana             |                  |            |
| AO_470 | 12           | Masculino | Solteiro     | Sem escolaridade      | Lobito              | Urbana             |                  |            |
| AO_471 | 53           | Feminino  | Solteiro     | Ensino secundário     | Cubal               | Urbana             |                  |            |
| AO_472 | 1            | Masculino | Solteiro     | Sem escolaridade      | Benguela            | Rural              |                  |            |
| AO_473 | 57           | Feminino  | Casado       | Sem escolaridade      | Benguela            | Rural              |                  |            |
| AO_474 | 22           | Feminino  | Solteiro     | Sem escolaridade      | Benguela            | Rural              |                  |            |
| AO_475 | 32           | Feminino  | Divorciado   | Ensino secundário     | Benguela            | Urbana             |                  |            |
| AO_476 | 20           | Masculino | Solteiro     | Ensino básico         | Bocoio              | Rural              |                  |            |
| AO_478 | 6            | Masculino | Solteiro     | Sem escolaridade      | Chongoroi           | Rural              |                  |            |
| AO_479 | 27           | Feminino  | Solteiro     | Ensino secundário     | Bocoio              | Rural              |                  |            |
| AO_480 | 31           | Masculino | Casado       | Ensino básico         | Ganda               | Urbana             |                  |            |

| ID     | Idade (anos) | Género    | Estado civil | Nível de escolaridade | Local de residência | Área de residência | Outras infecções |            |
|--------|--------------|-----------|--------------|-----------------------|---------------------|--------------------|------------------|------------|
|        |              |           |              |                       |                     |                    | Sífilis          | Hepatite B |
| AO_481 | 50           | Feminino  | Casado       | Ensino básico         | Benguela            | Urbana             |                  |            |
| AO_482 | 9            | Feminino  | Solteiro     | Sem escolaridade      | Caimbambo           | Urbana             | X                |            |
| AO_484 | 40           | Masculino | Solteiro     | Ensino básico         | Benguela            | Urbana             |                  |            |
| AO_485 | 36           | Masculino | Casado       | Ensino básico         | Benguela            | Rural              |                  |            |
| AO_486 | 26           | Feminino  | Solteiro     | Ensino básico         | Catumbela           | Urbana             |                  |            |
| AO_488 | 60           | Feminino  | Viúvo        | Ensino secundário     | Benguela            | Rural              |                  |            |
| AO_489 | 13           | Feminino  | Solteiro     | Sem escolaridade      | Catumbela           | Rural              | X                |            |
| AO_490 | 54           | Masculino | Casado       | Ensino secundário     | Lobito              | Urbana             |                  | X          |
| AO_491 | 33           | Feminino  | Casado       | Ensino básico         | Chongoroi           | Urbana             |                  |            |
| AO_492 | 40           | Feminino  | Divorciado   | Ensino básico         | Bocoio              | Rural              |                  | X          |
| AO_493 | 46           | Feminino  | Casado       | Ensino secundário     | Benguela            | Urbana             |                  |            |
| AO_494 | 11           | Masculino | Solteiro     | Sem escolaridade      | Catumbela           | Urbana             | X                |            |
| AO_495 | 42           | Masculino | Divorciado   | Ensino básico         | Benguela            | Rural              |                  |            |
| AO_496 | 6            | Feminino  | Solteiro     | Sem escolaridade      | Chongoroi           | Rural              |                  |            |
| AO_497 | 35           | Feminino  | Casado       | Ensino básico         | Ganda               | Rural              |                  | X          |
| AO_499 | 54           | Feminino  | Casado       | Ensino secundário     | Baía Farta          | Rural              |                  |            |
| AO_500 | 48           | Feminino  | Casado       | Ensino secundário     | Benguela            | Urbana             |                  |            |
| AO_501 | 42           | Feminino  | Casado       | Ensino secundário     | Cubal               | Urbana             |                  |            |
| AO_502 | 50           | Masculino | Casado       | Ensino básico         | Balombo             | Urbana             |                  |            |
| AO_503 | 52           | Feminino  | Divorciado   | Ensino secundário     | Ganda               | Urbana             |                  |            |

| ID     | Idade (anos) | Género    | Estado civil | Nível de escolaridade | Local de residência | Área de residência | Outras infecções |            |
|--------|--------------|-----------|--------------|-----------------------|---------------------|--------------------|------------------|------------|
|        |              |           |              |                       |                     |                    | Sífilis          | Hepatite B |
| AO_505 | 15           | Feminino  | Solteiro     | Sem escolaridade      | Benguela            | Rural              | X                |            |
| AO_506 | 71           | Masculino | Viúvo        | Sem escolaridade      | Benguela            | Urbana             |                  |            |
| AO_507 | 0            | Masculino | Solteiro     | Sem escolaridade      | Benguela            | Rural              | X                |            |
| AO_508 | 17           | Feminino  | Solteiro     | Sem escolaridade      | Benguela            | Urbana             | X                |            |
| AO_510 | 23           | Masculino | Solteiro     | Ensino básico         | Bocoio              | Rural              |                  |            |
| AO_511 | 39           | Feminino  | Casado       | Ensino básico         | Benguela            | Rural              |                  |            |
| AO_512 | 24           | Feminino  | Solteiro     | Sem escolaridade      | Benguela            | Urbana             |                  |            |
| AO_513 | 73           | Feminino  | Viúvo        | Sem escolaridade      | Catumbela           | Rural              |                  |            |
| AO_515 | 57           | Feminino  | Casado       | Sem escolaridade      | Cubal               | Rural              |                  |            |
| AO_516 | 0            | Feminino  | Solteiro     | Sem escolaridade      | Ganda               | Rural              |                  |            |
| AO_518 | 57           | Masculino | Divorciado   | Ensino básico         | Benguela            | Urbana             |                  |            |
| AO_519 | 58           | Masculino | Solteiro     | Ensino secundário     | Lobito              | Urbana             |                  |            |
| AO_520 | 5            | Masculino | Solteiro     | Sem escolaridade      | Benguela            | Urbana             |                  |            |
| AO_521 | 42           | Feminino  | Divorciado   | Ensino básico         | Cubal               | Urbana             |                  |            |
| AO_522 | 35           | Feminino  | Divorciado   | Ensino superior       | Benguela            | Urbana             |                  |            |
| AO_523 | 4            | Feminino  | Solteiro     | Sem escolaridade      | Chongoroi           | Rural              |                  |            |
| AO_524 | 14           | Feminino  | Solteiro     | Sem escolaridade      | Caimbambo           | Urbana             |                  |            |
| AO_525 | 17           | Feminino  | Solteiro     | Sem escolaridade      | Caimbambo           | Rural              |                  |            |
| AO_526 | 19           | Masculino | Solteiro     | Ensino básico         | Benguela            | Urbana             |                  |            |
| AO_527 | 7            | Masculino | Solteiro     | Sem escolaridade      | Catumbela           | Rural              | X                | X          |

| ID     | Idade (anos) | Género    | Estado civil | Nível de escolaridade | Local de residência | Área de residência | Outras infeções |            |
|--------|--------------|-----------|--------------|-----------------------|---------------------|--------------------|-----------------|------------|
|        |              |           |              |                       |                     |                    | Sífilis         | Hepatite B |
| AO_528 | 52           | Feminino  | Solteiro     | Sem escolaridade      | Benguela            | Urbana             |                 |            |
| AO_529 | 33           | Masculino | Solteiro     | Sem escolaridade      | Cubal               | Rural              |                 |            |
| AO_530 | 49           | Feminino  | Divorciado   | Ensino básico         | Balombo             | Urbana             |                 |            |
| AO_531 | 36           | Feminino  | Casado       | Desconhecido          | Baía Farta          | Rural              |                 |            |
| AO_532 | 63           | Feminino  | Viúvo        | Sem escolaridade      | Baía Farta          | Rural              |                 |            |
| AO_533 | 29           | Feminino  | Divorciado   | Ensino básico         | Lobito              | Urbana             |                 |            |
| AO_534 | 59           | Feminino  | Casado       | Ensino secundário     | Benguela            | Urbana             |                 |            |
| AO_535 | 47           | Feminino  | Divorciado   | Ensino secundário     | Benguela            | Urbana             |                 |            |
| AO_536 | 18           | Masculino | Solteiro     | Ensino básico         | Benguela            | Urbana             |                 |            |
| AO_537 | 7            | Masculino | Solteiro     | Sem escolaridade      | Benguela            | Urbana             |                 |            |
| AO_538 | 51           | Feminino  | Casado       | Ensino básico         | Cubal               | Desconhecido       |                 |            |
| AO_539 | 35           | Feminino  | Casado       | Ensino básico         | Benguela            | Urbana             |                 |            |
| AO_540 | 52           | Feminino  | Divorciado   | Ensino básico         | Lobito              | Urbana             |                 |            |
| AO_541 | 25           | Feminino  | Casado       | Ensino básico         | Benguela            | Urbana             |                 |            |
| AO_542 | 50           | Feminino  | Casado       | Ensino básico         | Benguela            | Urbana             |                 |            |
| AO_543 | 32           | Feminino  | Casado       | Ensino básico         | Caimbambo           | Urbana             |                 |            |
| AO_545 | 58           | Masculino | Divorciado   | Ensino básico         | Benguela            | Urbana             |                 |            |

**Anexo 4: Composição dos tampões utilizados**

Composição dos tampões utilizados na realização do presente estudo.

| Meios e soluções utilizados | Composição                                                                                                                                                                                         |
|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Tampão TAE 50× (1000 mL)    | <ul style="list-style-type: none"> <li>❖ 242 g Tris-base</li> <li>❖ 57,1 mL ácido acético glacial</li> <li>❖ 100 mL 0,5 M EDTA, pH 8,0</li> <li>❖ Adicionar água destilada para 1000 mL</li> </ul> |
| Tampão TE                   | <ul style="list-style-type: none"> <li>❖ 10 mM Tris-HCl</li> <li>❖ 0,1 mM EDTA, pH 8,0</li> </ul>                                                                                                  |
| Tampão de aplicação         | <ul style="list-style-type: none"> <li>❖ 0,25% (p/v) azul de bromofenol</li> <li>❖ 40% (p/v) sacarose</li> </ul>                                                                                   |

**Anexo 5: Protocolo *Whatman™ FTA® Elute™***

Protocolo recomendado para análises que requerem grandes quantidades de DNA:

1. Cortar três discos de 3,0 mm do centro da mancha de sangue utilizando um perfurador estéril, para um microtubo Eppendorf estéril de 1,5 mL;
  - a. Evitar contaminações cruzadas cortando três vezes entre amostras um pedaço de papel *Whatman™* 3MM com o perfurador.
2. Lavar adicionando 500 µL de água esterilizada e vortexar brevemente três vezes durante 5 s;
  - a. Se o lavado ficar de cor rosa, significa que a mancha de sangue não se encontrava completamente seca; neste caso, deve-se descartar os discos e secar completamente a mancha de sangue, antes de cortar novos discos e repetir o passo 2.
3. Remover o lavado e pressionar gentilmente os discos na parede do microtubo para remover o máximo de líquido possível;
4. Adicionar 100 µL de água esterilizada e fazer um *spin down* 5 s para garantir que os discos estão completamente imersos;
5. Transferir o microtubo para um bloco de aquecimento a 95–100 °C durante 15–30 min;
6. Vortexar brevemente, aproximadamente 60 vezes;
7. Efectuar *spin down* (aproximadamente 30 s) para separar os discos do eluente contendo o DNA extraído;
8. Remover cuidadosamente os discos com uma ponta de pipeta esterilizada, descartando-os;
9. Armazenar o DNA eluído a -20 °C.

**Anexo 6: Nomenclatura dos nucleótidos e aminoácidos**

Código da IUPAC para a nomenclatura dos nucleótidos.

| <b>Código IUPAC</b> | <b>Base orgânica</b>   |
|---------------------|------------------------|
| A                   | Adenina                |
| C                   | Citosina               |
| G                   | Guanina                |
| T (ou U)            | Timina (ou Uracilo)    |
| R                   | A ou G                 |
| Y                   | C ou T                 |
| S                   | C ou G                 |
| W                   | A ou T                 |
| K                   | G ou T                 |
| M                   | A ou C                 |
| B                   | C ou G ou T            |
| D                   | A ou G ou T            |
| H                   | A ou C ou T            |
| V                   | A ou C ou G            |
| N                   | Qualquer base orgânica |

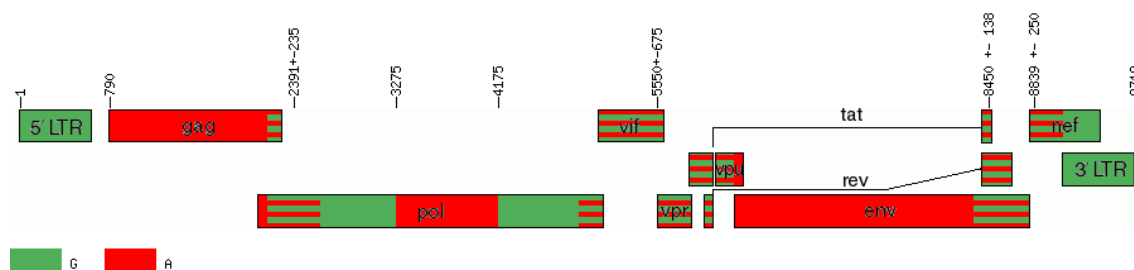
Código de três letras e de uma letra da IUPAC para a nomenclatura dos aminoácidos.

| Aminoácido      | Código de três letras | Código IUPAC |
|-----------------|-----------------------|--------------|
| Ácido aspártico | Asp                   | D            |
| Ácido glutâmico | Glu                   | E            |
| Alanina         | Ala                   | A            |
| Arginina        | Arg                   | R            |
| Asparagina      | Asn                   | N            |
| Cisteína        | Cys                   | C            |
| Fenilalanina    | Phe                   | F            |
| Glicina         | Gly                   | G            |
| Glutamina       | Gln                   | Q            |
| Histidina       | His                   | H            |
| Isoleucina      | Ile                   | I            |
| Leucina         | Leu                   | L            |
| Lisina          | Lys                   | K            |
| Metionina       | Met                   | M            |
| Prolina         | Pro                   | P            |
| Serina          | Ser                   | S            |
| Tirosina        | Tyr                   | Y            |
| Treonina        | Thr                   | T            |
| Triptofano      | Trp                   | W            |
| Valina          | Val                   | V            |

## Anexo 7: Mapas genômicos de formas recombinantes circulantes seleccionadas

Representação esquemática dos genomas mosaico das formas recombinantes circulantes CRF02\_AG, CRF26\_A5U e CRF49\_cpx (disponíveis em <https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/breakpoints.html>).

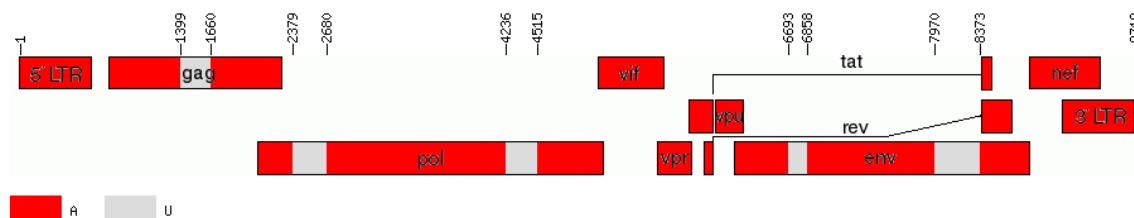
### ❖ CRF02\_AG



Pontos de recombinação em HXB2 (Carr *et al.*, 1998):

1 789 G  
 790 2155 A  
 2156 2625 A/G\*  
 2626 3274 G\*  
 3275 4174 A  
 4175 4874 G  
 4875 6224 A/G  
 6225 8311 A  
 8312 9089 A/G  
 9090 9719 G

### ❖ CRF26\_A5U

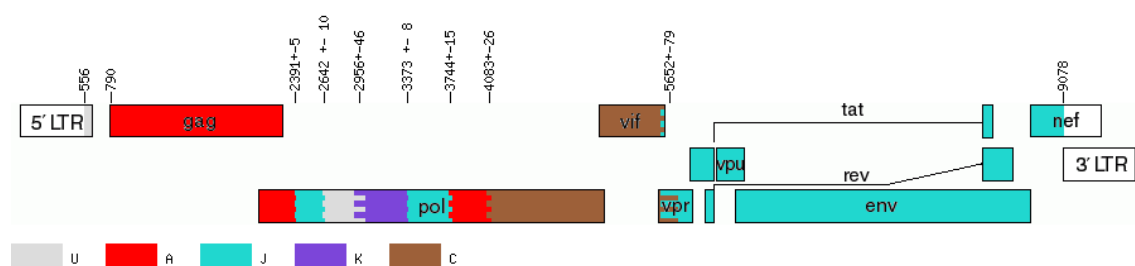


Pontos de recombinação em HXB2 (M. Peeters, comunicação pessoal):

1 1398 A  
 1399 1659 U  
 1660 2378 A

2379 2679 U\*\*  
 2680 4235 A\*  
 4236 4514 U  
 4515 6692 A  
 6693 6857 U  
 6858 7969 A  
 7970 8872 U  
 8373 9719 A

#### ❖ CFR49\_cpx



#### Pontos de recombinação em HXB2 (Silva *et al.*, 2010):

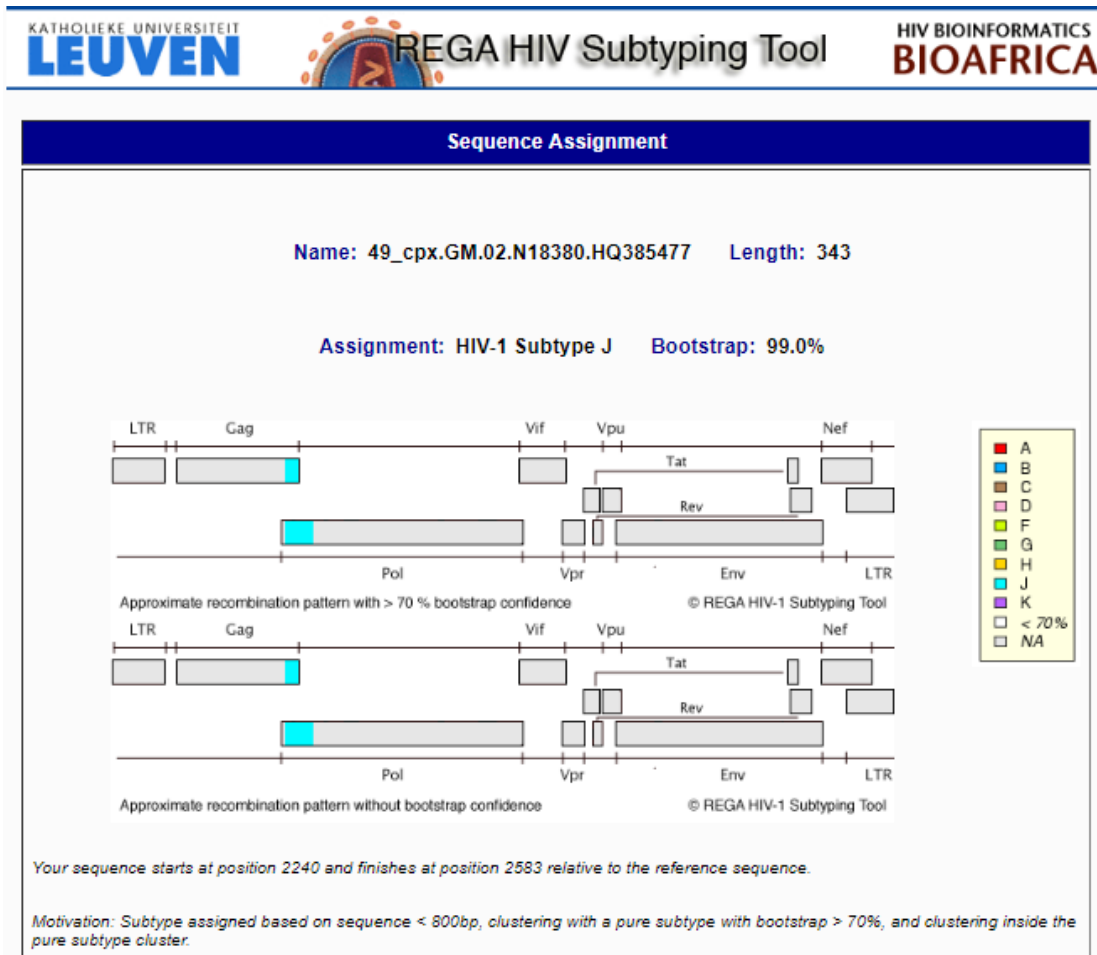
556 789 U  
 790 2385 A  
 2386 2395 A/J  
 2396 2631 J\*  
 2632 2652 J/U  
 2653 2909 U  
 2910 3001 U/K  
 3002 3364 K\*  
 3365 3381 K/J  
 3382 3728 J  
 3729 3758 J/A  
 3759 4056 A  
 4057 4108 A/C  
 4109 5572 C  
 5573 5730 C/J  
 5731 9078 J

\* Segmentos que incluem a região codificante da protease no gene *pol* (posições nucleotídicas 2167 a 2583, em HXB2).

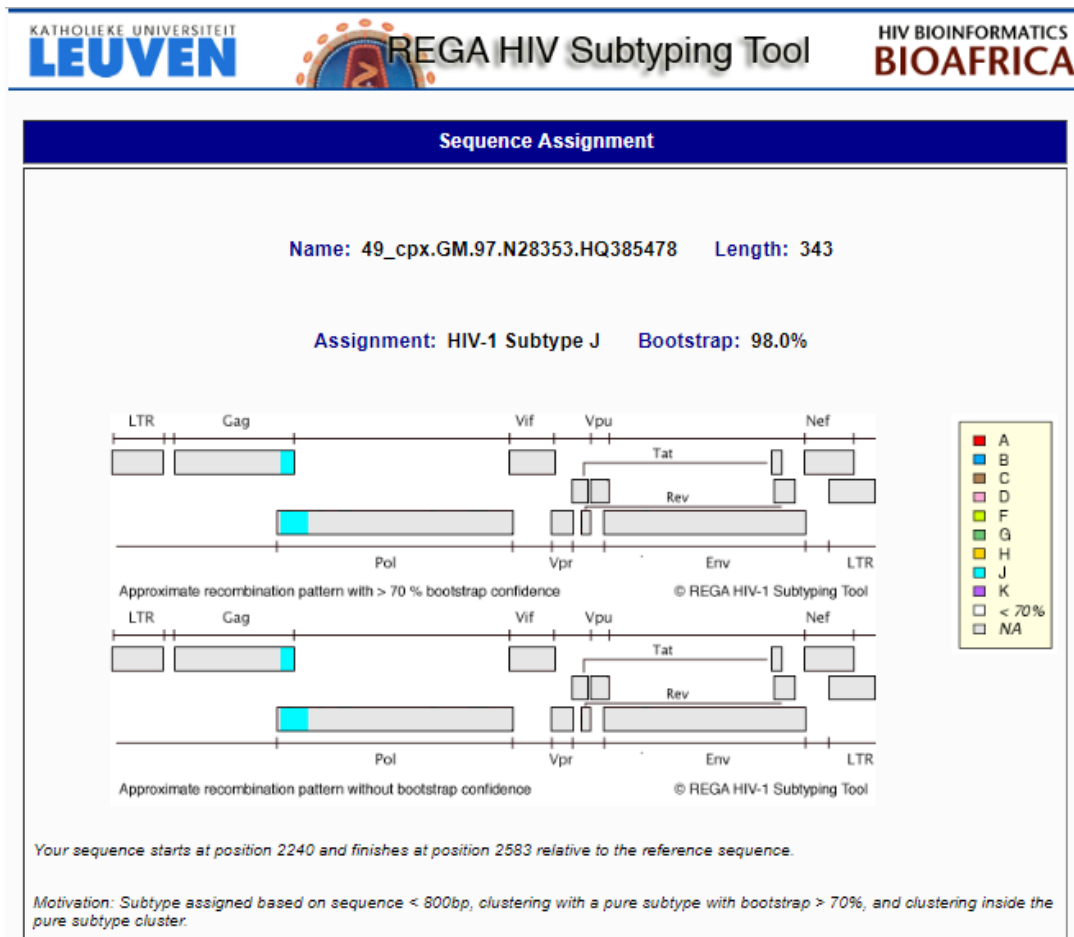
\* Segmentos que incluem a região codificante em estudo da transcriptase reversa no gene *pol* (posições nucleotídicas 2633 a 3230, em HXB2).

## Anexo 8: Resultados da genotipagem rápida automatizada das sequências de referência CRF49\_cpx utilizadas

Resultado da genotipagem rápida automatizada, utilizando o *software REGA HIV-1 Subtyping Tool* (versão 3.0), da sequência CRF49\_cpx.GM.02.N18380.HQ385477 utilizada como referência na construção da árvore filogenética para a região codificante da protease.



Resultado da genotipagem rápida automatizada, utilizando o software *REGA HIV-1 Subtyping Tool* (versão 3.0), da sequência CRF49\_cpx.GM.97.N28353.HQ385478 utilizada como referência na construção da árvore filogenética para a região codificante da protease.



### Anexo 9: Resultados finais comparados da genotipagem das sequências em estudo para as regiões codificantes da protease e da transcriptase reversa

Comparação dos resultados da genotipagem das sequências codificantes da protease e da transcriptase reversa com base na respectiva análise filogenética e genótipo combinado. Legenda: ID – identificação das amostras; \* – genótipos atribuídos após análise de bootscanning das respectivas sequências nucleotídicas.

| ID     | Protease (n=50) | Transcriptase reversa (n=53) | Combinado (n=47) |
|--------|-----------------|------------------------------|------------------|
| AO_6   | C*              | C*                           | C                |
| AO_7   | –               | CRF02_AG                     | –                |
| AO_80  | F (F1)          | F (F1)                       | F (F1)           |
| AO_92  | C               | C                            | C                |
| AO_119 | CRF02_AG        | CRF02_AG                     | CRF02_AG         |
| AO_141 | C               | C                            | C                |
| AO_330 | C               | C                            | C                |
| AO_380 | C               | C*                           | C                |
| AO_442 | C               | C                            | C                |
| AO_446 | –               | U*                           | –                |
| AO_452 | F (F1)          | F (F1)                       | F (F1)           |
| AO_455 | A*              | A                            | A                |
| AO_458 | C               | C                            | C                |
| AO_461 | A*              | H                            | A/H              |
| AO_466 | CRF02_AG        | CRF02_AG                     | CRF02_AG         |
| AO_467 | A*              | A                            | A                |
| AO_469 | C               | C                            | C                |
| AO_471 | C               | C                            | C                |
| AO_472 | F (F1)          | F (F1)                       | F (F1)           |
| AO_473 | CRF02_AG        | CRF02_AG                     | CRF02_AG         |
| AO_474 | U*              | –                            | –                |
| AO_476 | CRF02_AG        | CRF02_AG                     | CRF02_AG         |
| AO_478 | C               | C*                           | C                |
| AO_479 | –               | G                            | –                |
| AO_480 | U*              | U*                           | U                |
| AO_484 | C               | C                            | C                |
| AO_485 | –               | G                            | –                |
| AO_486 | –               | CRF02_AG                     | –                |

| ID     | Protease (n=50) | Transcriptase reversa (n=53) | Combinado (n=47) |
|--------|-----------------|------------------------------|------------------|
| AO_488 | U*              | D*                           | U/D              |
| AO_490 | C               | C                            | C                |
| AO_491 | –               | F (F1)*                      | –                |
| AO_493 | C               | C                            | C                |
| AO_494 | G*              | G*                           | G                |
| AO_495 | F (F1)          | F (F1)                       | F (F1)           |
| AO_497 | F (F1)          | F (F1)*                      | F (F1)           |
| AO_503 | C               | –                            | –                |
| AO_506 | J*              | J                            | J                |
| AO_507 | F (F1)          | F (F1)*                      | F (F1)           |
| AO_510 | CRF02_AG        | CRF02_AG                     | CRF02_AG         |
| AO_512 | C               | C*                           | C                |
| AO_515 | F (F1)*         | F (F1)                       | F (F1)           |
| AO_516 | C               | C                            | C                |
| AO_519 | F (F1)          | F (F1)*                      | F (F1)           |
| AO_520 | G               | G                            | G                |
| AO_522 | C               | C                            | C                |
| AO_524 | C*              | C                            | C                |
| AO_525 | A*              | –                            | –                |
| AO_526 | C               | C                            | C                |
| AO_527 | G               | U*                           | G/U              |
| AO_532 | G*              | CRF02_AG                     | G/CRF02_AG       |
| AO_533 | B               | U*                           | B/U              |
| AO_534 | F (F1)          | F (F1)*                      | F (F1)           |
| AO_538 | A*              | H                            | A/H              |
| AO_540 | CRF02_AG        | CRF02_AG                     | CRF02_AG         |
| AO_542 | G               | G                            | G                |
| AO_545 | C*              | C                            | C                |

### Anexo 10: Resultados da pesquisa de mutações de resistência aos ARVs

Resultados da pesquisa de mutações associadas a resistência aos inibidores da protease e transcriptase reversa nas sequências das amostras em estudo utilizando o programa *HIVdb* (versão 8.5). Na tabela, os blocos cinza reflectem a ausência de uma sequência codificante para a protease ou transcriptase reversa da(s) respectiva(s) amostra(s), estão destacadas a negrito e vermelho as mutações com importância para a vigilância epidemiológica da resistência aos ARVs transmitida do HIV-1, os códons de terminação estão indicados com “\*” (*e.g.* W42\*) e as mutações destacadas em itálico a azul correspondem a mutações extremamente incomuns, localizadas em posições de mutações de resistência, que não têm impacto na resistência aos NNRTIs. Legenda: *Major* – mutação principal; *Minor* – mutação acessória.

| ID     | Mutações de resistência aos PIs |              | Mutações de resistência aos NRTIs | Mutações de resistência aos NNRTIs         | Outras mutações associadas a resistência aos: |       | Mutações APOBEC presentes nas regiões codificantes da: |                                                             |
|--------|---------------------------------|--------------|-----------------------------------|--------------------------------------------|-----------------------------------------------|-------|--------------------------------------------------------|-------------------------------------------------------------|
|        | <i>Major</i>                    | <i>Minor</i> |                                   |                                            | PIs                                           | RTIs  | PR                                                     | RT                                                          |
| AO_6   | –                               | –            | <b>M184I</b>                      | E138A                                      | –                                             | –     | G40E, W42*, G68K, G78R, G86R, G94R                     | G51R, W88*, G93R, G112E, W153*, <i>G190R</i> , W212*, G213R |
| AO_7   |                                 |              | –                                 | –                                          |                                               | –     |                                                        | –                                                           |
| AO_80  | –                               | –            | –                                 | <b>Y181C</b>                               | L10V                                          | –     | –                                                      | –                                                           |
| AO_92  | –                               | –            | <b>K70E, Y115F, M184V</b>         | V108I, <b>Y181C</b> , <b>G190A</b> , H221Y |                                               | K101N | –                                                      | –                                                           |
| AO_119 | –                               | –            | –                                 | –                                          | L10I, K20I                                    | –     | –                                                      | –                                                           |
| AO_141 | –                               | –            | –                                 | <b>V106M, Y181C</b>                        | –                                             | –     | –                                                      | –                                                           |
| AO_330 | –                               | –            | –                                 | A98G, <b>Y181C</b>                         | –                                             | –     | –                                                      | –                                                           |
| AO_380 | –                               | –            | –                                 | –                                          | T74S, V82I                                    | –     | –                                                      | –                                                           |
| AO_442 | –                               | –            | –                                 | –                                          | V82I                                          | –     | –                                                      | –                                                           |
| AO_446 |                                 |              | –                                 | –                                          |                                               | –     |                                                        | –                                                           |

| ID     | Mutações de resistência aos PIs |             | Mutações de resistência aos NRTIs | Mutações de resistência aos NNRTIs | Outras mutações associadas a resistência aos: |       | Mutações APOBEC presentes nas regiões codificantes da: |    |
|--------|---------------------------------|-------------|-----------------------------------|------------------------------------|-----------------------------------------------|-------|--------------------------------------------------------|----|
|        | Major                           | Minor       |                                   |                                    | PIs                                           | RTIs  | PR                                                     | RT |
| AO_452 | —                               | —           | —                                 | —                                  | L10I                                          | —     | —                                                      | —  |
| AO_455 | —                               | —           | —                                 | —                                  | —                                             | —     | —                                                      | —  |
| AO_458 | —                               | —           | <b>M184V</b>                      | V108I, <b>Y181C</b>                | —                                             | —     | —                                                      | —  |
| AO_461 | —                               | —           | —                                 | <b>Y188L, G190A</b>                | L10I, K20I                                    | V90I  | —                                                      | —  |
| AO_466 | —                               | —           | —                                 | —                                  | L10I, K20I                                    | V179I | —                                                      | —  |
| AO_467 | —                               | —           | —                                 | —                                  | —                                             | —     | —                                                      | —  |
| AO_469 | —                               | —           | —                                 | —                                  | —                                             | —     | —                                                      | —  |
| AO_471 | —                               | —           | —                                 | —                                  | —                                             | —     | —                                                      | —  |
| AO_472 | —                               | —           | —                                 | —                                  | L10V                                          | —     | —                                                      | —  |
| AO_473 | —                               | —           | —                                 | —                                  | K20I                                          | S68G  | —                                                      | —  |
| AO_474 | —                               | <b>G73S</b> |                                   |                                    | —                                             |       | W42*, G51R, G86K                                       |    |
| AO_476 | —                               | —           | —                                 | —                                  | L10V, K20I                                    | —     | —                                                      | —  |
| AO_478 | —                               | —           | <b>M184V, T215Y</b>               | <b>V106M, Y181C</b>                | —                                             | V179I | —                                                      | —  |
| AO_479 |                                 |             | —                                 | —                                  |                                               | —     |                                                        | —  |
| AO_480 | —                               | —           | —                                 | <b>K103N</b>                       | —                                             | —     | —                                                      | —  |
| AO_484 | —                               | —           | —                                 | —                                  | —                                             | —     | W42*, G49E, G51R, G52S, G68R, G94R                     | —  |
| AO_485 |                                 |             | —                                 | —                                  |                                               | V106I |                                                        | —  |
| AO_486 |                                 |             | —                                 | —                                  |                                               | T69S  |                                                        | —  |
| AO_488 | —                               | —           | <b>M184V, T215Y</b>               | <b>K103N</b>                       | L10V                                          | —     | —                                                      | —  |

| ID     | Mutações de resistência aos PIs |       | Mutações de resistência aos NRTIs                                   | Mutações de resistência aos NNRTIs                        | Outras mutações associadas a resistência aos: |             | Mutações APOBEC presentes nas regiões codificantes da: |                                              |
|--------|---------------------------------|-------|---------------------------------------------------------------------|-----------------------------------------------------------|-----------------------------------------------|-------------|--------------------------------------------------------|----------------------------------------------|
|        | Major                           | Minor |                                                                     |                                                           | PIs                                           | RTIs        | PR                                                     | RT                                           |
| AO_490 | —                               | —     | —                                                                   | —                                                         | —                                             | —           | —                                                      | —                                            |
| AO_491 |                                 |       | —                                                                   | —                                                         |                                               | V106I       |                                                        | —                                            |
| AO_493 | —                               | —     | —                                                                   | —                                                         | —                                             | —           | —                                                      | —                                            |
| AO_494 | <b>L90M</b>                     | —     | —                                                                   | <b>K101E</b> , V108I, <b>Y181C</b> , <b>G190A</b> , H221Y | K20I, V82I                                    | —           | —                                                      | —                                            |
| AO_495 | —                               | —     | —                                                                   | —                                                         | L10V                                          | V118I       | —                                                      | —                                            |
| AO_497 | —                               | —     | <b>M41L</b> , E44D, <b>D67N</b> , K70T, <b>M184V</b> , <b>T215Y</b> | A98G, <b>Y188M</b>                                        | L10V                                          | S68G, V106I | W42*, G94S                                             | W71*, W88*, G99R, G112E, W153*, L210*, W212* |
| AO_503 | —                               | —     |                                                                     |                                                           | —                                             | —           | G27R, G48R, G49R, G86R                                 |                                              |
| AO_506 | —                               | —     | —                                                                   | —                                                         | —                                             | —           | —                                                      | —                                            |
| AO_507 | —                               | —     | <b>M184V</b>                                                        | <b>V106A</b>                                              | L10V                                          | V118I       | —                                                      | —                                            |
| AO_510 | —                               | —     | —                                                                   | —                                                         | K20I                                          | —           | —                                                      | —                                            |
| AO_512 | —                               | —     | —                                                                   | —                                                         | T74S                                          | —           | —                                                      | —                                            |
| AO_515 | —                               | —     | —                                                                   | —                                                         | L10V                                          | —           | —                                                      | —                                            |
| AO_516 | —                               | —     | —                                                                   | —                                                         | —                                             | —           | —                                                      | —                                            |
| AO_519 | —                               | —     | —                                                                   | —                                                         | L10V                                          | —           | —                                                      | —                                            |
| AO_520 | —                               | —     | <b>M41L</b> , <b>D67N</b> ,                                         | V108I, <b>Y181C</b> , H221Y                               | L10I, K20I, V82I                              | —           | —                                                      | —                                            |

| ID     | Mutações de resistência aos PIs |             | Mutações de resistência aos NRTIs       | Mutações de resistência aos NNRTIs              | Outras mutações associadas a resistência aos: |             | Mutações APOBEC presentes nas regiões codificantes da: |      |
|--------|---------------------------------|-------------|-----------------------------------------|-------------------------------------------------|-----------------------------------------------|-------------|--------------------------------------------------------|------|
|        | Major                           | Minor       |                                         |                                                 | PIs                                           | RTIs        | PR                                                     | RT   |
|        |                                 |             | <b>V75M, M184V, L210W, T215Y, K219N</b> |                                                 |                                               |             |                                                        |      |
| AO_522 | –                               | –           | –                                       | –                                               | –                                             | –           | –                                                      | –    |
| AO_524 | –                               | –           | –                                       | –                                               | L10V                                          | –           | –                                                      | –    |
| AO_525 | <b>D30N, M46I</b>               | <b>G73S</b> |                                         |                                                 | K20I                                          |             | G27K, W42*, G48E, G49R, G51K, G68K, G78R, G86K, G94S   |      |
| AO_526 | –                               | –           | –                                       | –                                               | –                                             | –           | –                                                      | –    |
| AO_527 | –                               | –           | –                                       | –                                               | K20I, V82I                                    | –           | –                                                      | –    |
| AO_532 | –                               | –           | –                                       | –                                               | L10I, K20I                                    | –           | –                                                      | –    |
| AO_533 | –                               | –           | <b>M184I</b>                            | E138A                                           | A71I                                          | V179I       | G94S                                                   | G93R |
| AO_534 | –                               | –           | <b>M184V, T215F</b>                     | V108I, <b>Y181C</b>                             | L10V                                          | V90I, V179I | –                                                      | –    |
| AO_538 | –                               | –           | <b>K219R</b>                            | –                                               | K20I                                          | –           | –                                                      | –    |
| AO_540 | –                               | –           | –                                       | –                                               | K20I                                          | –           | –                                                      | –    |
| AO_542 | <b>I54V</b>                     | –           | <b>M41L, M184V, T215Y</b>               | A98G, K101H, <b>K103N, Y181C, G190A</b> , H221Y | L10I, K20I, V82I                              | –           | –                                                      | –    |
| AO_545 | –                               | –           | –                                       | –                                               | –                                             | –           | –                                                      | –    |

### Anexo 11: Perfis fenotípicos putativos de resistência aos ARVs

Perfis fenotípicos putativos de resistência aos diferentes inibidores da protease e transcriptase reversa obtidos para as sequências das amostras em estudo utilizando o programa *HIVdb* (versão 8.5). Na tabela, os blocos cinza reflectem a ausência de uma sequência codificante para a protease ou transcriptase reversa da(s) respectiva(s) amostra(s) e encontram-se destacados a vermelho os perfis de elevada resistência encontrados nas amostras em estudo. Legenda: ATV/r – atazanavir/r; DRV/r – darunavir/r; FPV/r – fosamprenavir/r; IDV/r – indinavir/r; LPV/r – lopinavir/r; NFV – nelfinavir; SQV/r – saquinavir/r; TPV/r – tipranavir/r; r – ritonavir; ABC – abacavir; AZT – zidovudine; D4T – stavudine; DDI – didanosine; FTC – emtricitabine; 3TC – lamivudine; TDF – tenofovir; EFV – efavirenz; ETR – etravirine; NVP – nevirapine; RPV – rilpivirine; S – susceptibilidade; P – resistência potencial de nível baixo; B – resistência de nível baixo; I – resistência de nível intermédio; E – resistência de nível elevado.

| ID     | PIs   |       |       |       |       |     |       |       | NRTIs |     |     |     |     |     |     | NNRTIs |     |     |     |
|--------|-------|-------|-------|-------|-------|-----|-------|-------|-------|-----|-----|-----|-----|-----|-----|--------|-----|-----|-----|
|        | ATV/r | DRV/r | FPV/r | IDV/r | LPV/r | NFV | SQV/r | TPV/r | ABC   | AZT | D4T | DDI | FTC | 3TC | TDF | EFV    | ETR | NVP | RPV |
| AO_6   | S     | S     | S     | S     | S     | S   | S     | S     | B     | S   | S   | P   | E   | E   | S   | S      | P   | S   | B   |
| AO_7   |       |       |       |       |       |     |       |       | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_80  | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | I      | I   | E   | I   |
| AO_92  | S     | S     | S     | S     | S     | S   | S     | S     | E     | S   | B   | B   | E   | E   | I   | E      | E   | E   | E   |
| AO_119 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_141 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | E      | I   | E   | I   |
| AO_330 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | I      | I   | E   | E   |
| AO_380 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_442 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_446 |       |       |       |       |       |     |       |       | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_452 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_455 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |

| ID     | PIs   |       |       |       |       |     |       |       | NRTIs |     |     |     |     |     |     | NNRTIs |     |     |     |
|--------|-------|-------|-------|-------|-------|-----|-------|-------|-------|-----|-----|-----|-----|-----|-----|--------|-----|-----|-----|
|        | ATV/r | DRV/r | FPV/r | IDV/r | LPV/r | NFV | SQV/r | TPV/r | ABC   | AZT | D4T | DDI | FTC | 3TC | TDF | EFV    | ETR | NVP | RPV |
| AO_458 | S     | S     | S     | S     | S     | S   | S     | S     | B     | S   | S   | P   | E   | E   | S   | I      | I   | E   | I   |
| AO_461 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | E      | B   | E   | E   |
| AO_466 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_467 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_469 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_471 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_472 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_473 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_474 | P     | S     | P     | B     | S     | B   | B     | S     |       |     |     |     |     |     |     |        |     |     |     |
| AO_476 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_478 | S     | S     | S     | S     | S     | S   | S     | S     | B     | I   | I   | B   | E   | E   | S   | E      | I   | E   | I   |
| AO_479 |       |       |       |       |       |     |       |       | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_480 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | E      | S   | E   | S   |
| AO_484 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_485 |       |       |       |       |       |     |       |       | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_486 |       |       |       |       |       |     |       |       | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_488 | S     | S     | S     | S     | S     | S   | S     | S     | B     | I   | I   | B   | E   | E   | S   | E      | S   | E   | S   |
| AO_490 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_491 |       |       |       |       |       |     |       |       | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_493 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |

| ID     | PIs   |       |       |       |       |     |       |       | NRTIs |     |     |     |     |     |     | NNRTIs |     |     |     |
|--------|-------|-------|-------|-------|-------|-----|-------|-------|-------|-----|-----|-----|-----|-----|-----|--------|-----|-----|-----|
|        | ATV/r | DRV/r | FPV/r | IDV/r | LPV/r | NFV | SQV/r | TPV/r | ABC   | AZT | D4T | DDI | FTC | 3TC | TDF | EFV    | ETR | NVP | RPV |
| AO_494 | B     | S     | B     | I     | B     | E   | I     | S     | S     | S   | S   | S   | S   | S   | S   | E      | E   | E   | E   |
| AO_495 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_497 | S     | S     | S     | S     | S     | S   | S     | S     | E     | E   | E   | E   | E   | E   | I   | B      | P   | I   | B   |
| AO_503 | S     | S     | S     | S     | S     | S   | S     | S     |       |     |     |     |     |     |     |        |     |     |     |
| AO_506 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_507 | S     | S     | S     | S     | S     | S   | S     | S     | B     | S   | S   | P   | E   | E   | S   | I      | S   | E   | S   |
| AO_510 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_512 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_515 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_516 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_519 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_520 | S     | S     | S     | S     | S     | S   | S     | S     | E     | E   | E   | E   | E   | E   | E   | I      | I   | E   | E   |
| AO_522 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_524 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_525 | B     | S     | B     | B     | B     | E   | B     | S     |       |     |     |     |     |     |     |        |     |     |     |
| AO_526 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_527 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_532 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_533 | S     | S     | S     | S     | S     | S   | S     | S     | B     | S   | S   | P   | E   | E   | S   | S      | P   | S   | B   |
| AO_534 | S     | S     | S     | S     | S     | S   | S     | S     | B     | I   | I   | B   | E   | E   | S   | I      | I   | E   | I   |

| ID     | PIs   |       |       |       |       |     |       |       | NRTIs |     |     |     |     |     |     | NNRTIs |     |     |     |
|--------|-------|-------|-------|-------|-------|-----|-------|-------|-------|-----|-----|-----|-----|-----|-----|--------|-----|-----|-----|
|        | ATV/r | DRV/r | FPV/r | IDV/r | LPV/r | NFV | SQV/r | TPV/r | ABC   | AZT | D4T | DDI | FTC | 3TC | TDF | EFV    | ETR | NVP | RPV |
| AO_538 | S     | S     | S     | S     | S     | S   | S     | S     | S     | P   | P   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_540 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |
| AO_542 | B     | S     | P     | B     | B     | B   | B     | B     | I     | I   | I   | I   | E   | E   | B   | E      | E   | E   | E   |
| AO_545 | S     | S     | S     | S     | S     | S   | S     | S     | S     | S   | S   | S   | S   | S   | S   | S      | S   | S   | S   |

## Anexo 12: Resultados da pesquisa de assinaturas genéticas relativos às mutações de resistência aos ARVs

Resultados da pesquisa de assinaturas genéticas relativos às mutações associadas a resistência aos inibidores da protease para as quais foram identificadas correlações positivas com os genótipos especificados, com recurso ao teste de Fisher bilateral aplicado à análise de tabelas de contingência 2×2 (nível de significância de 1%).

| Mutações PIs | Sub-subtipo F1 (n=9) | Subtipos não-F1 (n=41) | <i>p</i> ( $\alpha=0,01$ ) |
|--------------|----------------------|------------------------|----------------------------|
| L10V         | 8                    | 3                      | <0,001                     |
| Mutações PIs | Subtipo G (n=5)      | Subtipos não-G (n=45)  | <i>p</i> ( $\alpha=0,01$ ) |
| K20I         | 5                    | 9                      | <0,001                     |
| V82I         | 4                    | 2                      | <0,001                     |
| Mutações PIs | CRF02_AG (n=6)       | Não-CRF02_AG (n=44)    | <i>p</i> ( $\alpha=0,01$ ) |
| K20I         | 6                    | 8                      | <0,001                     |

Resultado da pesquisa de assinaturas genéticas relativo à mutação associada a resistência aos inibidores da transcriptase reversa para a qual foi identificada uma correlação positiva com o subtipo G, com recurso ao teste de Fisher bilateral aplicado à análise de tabelas de contingência 2×2 (nível de significância de 1%).

| Mutações NNRTIs | Subtipo G (n=5) | Subtipos não-G (n=48) | <i>p</i> ( $\alpha=0,01$ ) |
|-----------------|-----------------|-----------------------|----------------------------|
| H221Y           | 3               | 1                     | 0,002                      |