

MScMM

MESTRADO EM
**MICROBIOLOGIA
MÉDICA**

MARIA ISABEL SARGENTO GODINHO
BSc em Biologia Celular e Molecular

**Diversidade genética e resistência
primária aos inibidores da integrase no
vírus da imunodeficiência humana do tipo
1 (HIV-1) na província de Benguela
(Angola)**

Abril, 2024



Universidade Nova de Lisboa

**Instituto de Tecnologia Química e Biológica António
Xavier**

**Diversidade genética e resistência primária aos
inibidores da integrase no vírus da
imunodeficiência humana tipo 1 (HIV-1) na
província de Benguela (Angola)**

Autora: Maria Isabel Sargento Godinho

Orientador: Prof. Doutor João Piedade, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa

Co-Orientadora: Prof.^a Doutora Filomena Pereira, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do grau de
Mestre em Microbiologia Médica

Abril, 2024



INSTITUTO DE HIGIENE E
MEDICINA TROPICAL
DESDE 1902



Elementos bibliográficos

- Godinho I., Quitéria R., Yefimenko L., Martins Pereira F., Piedade J. Resistência primária aos inibidores da integrase no vírus da imunodeficiência humana do tipo 1 (HIV-1) na província de Benguela (Angola). 6º Congresso Nacional de Medicina Tropical. 20 e 21 de abril de 2023. Lisboa, Portugal

(Comunicação em póster; Anexo 1)

Agradecimentos

Gostaria de expressar os meus sinceros agradecimentos a todas as pessoas que contribuíram para a conclusão desta dissertação.

Ao meu orientador, Professor Doutor João Piedade, pela orientação e apoio. O seu conhecimento, incentivo, paciência e dedicação ao longo deste processo, não só aprimoraram a minha compreensão sobre o tema, mas também fortaleceram a minha confiança e segurança ao longo deste trabalho.

À Doutora Lesya Yefimenko e à Professora Doutora Filomena Pereira pela sua generosidade no fornecimento das amostras biológicas e dos dados epidemiológicos necessários para o avanço deste estudo. Obrigada pela disponibilidade e simpatia demonstradas.

Aos meus amigos pelo seu encorajamento, paciência e compreensão. Um especial obrigado à Joana, Elena, Inês e Pedro que, de alguma forma, se tornaram fundamentais para superar os desafios que surgiram ao longo do caminho e pela partilha de conhecimento. Gostaria de expressar, com sincera gratidão e de forma especial, um agradecimento à Margarida, pela sua total disponibilidade e momentos de partilha.

Não posso deixar de agradecer o amor e apoio essencial de toda a minha família durante esta etapa, em especial aos meus pais, Isabel e Bento e à minha irmã Sara. Sem eles não estaria onde me encontro hoje.

Este trabalho não teria sido possível sem o apoio e contribuições de todos vocês.

Obrigada!

Resumo

O HIV-1 exibe uma notável diversidade genética, especialmente acentuada em África, classificando-se em quatro grupos filogenéticos, entre os quais, o grupo M, responsável pela pandemia, compreende diversos subtipos (A-D, F-H e J-L). A terapia antirretroviral combinada tem reduzido a mortalidade, morbidade e transmissibilidade viral, mas a capacidade do HIV-1 desenvolver resistência aos antirretrovirais é uma preocupação, incluindo aos inibidores de transferência de cadeia da integrase (INSTI). Em 2019, a Organização Mundial da Saúde recomendou o Dolutegravir, um INSTI, como tratamento de primeira e segunda linhas, com introdução em Angola em 2021. Neste estudo, realizado no Hospital Geral de Benguela, em Angola, em 2016/2017, foram analisadas 226 amostras de sangue de indivíduos com HIV-1, colhidas em cartões FTA® (GE Healthcare). Partindo do DNA proviral, foi realizado um protocolo de *nested* PCR para amplificação da região codificante da integrase no gene *pol* do HIV-1, tendo-se obtido 97 amostras positivas (42,9%). Estas deram origem a 92 sequências pelo método de sequenciação de Sanger, submetidas a genotipagem por análise filogenética e identificação de mutações associadas a uma diminuição da suscetibilidade aos INSTI e polimorfismos genéticos. Os resultados indicaram predominância do sub-subtipo F1 (28,3%), seguido pelos subtipos A (20,7%) e C (19,6%), encontrando-se ainda os subtipos D, G e H, e CRF02_AG, CRF18_cpx, CRF19_cpx e CRF27_cpx. 91,6% das sequências mostraram total suscetibilidade aos INSTI. Foi encontrada a mutação principal E92G, relevante para vigilância epidemiológica, associada a níveis de resistência intermédio, baixo e potencial de nível baixo, bem como as mutações acessórias L74M, Q95K, T97A e E157Q, associadas a resistência potencial de nível baixo, num total de 10 sequências. Adicionalmente, foram identificados numerosos polimorfismos genéticos, alguns como assinaturas genéticas dos subtipos A, C e F1. Os resultados obtidos contribuem para a vigilância epidemiológica molecular do HIV-1 em Angola, incluindo a monitorização da resistência primária aos INSTI.

Palavras-chave: HIV-1, Angola, diversidade genética, inibidores da integrase, resistência primária

Abstract

HIV-1 demonstrates a notable genetic diversity, particularly pronounced in Africa, being classified into four phylogenetic groups, among which group M, responsible for the pandemic, encompasses various subtypes (A-D, F-H, and J-L). Highly active antiretroviral therapy has reduced mortality, morbidity, and viral transmissibility, but HIV-1 ability to develop resistance to antiretrovirals, including integrase strand transfer inhibitors (INSTIs), is a big concern. In 2019, the World Health Organization recommended Dolutegravir, an INSTI, as first and second-line treatment, being introduced in Angola in 2021. In this study, conducted at the Benguela General Hospital in Angola during 2016/2017, 226 blood samples from individuals infected with HIV-1, collected on FTA[®] cards (GE Healthcare, United Kingdom), were analyzed. Starting from proviral DNA, a nested PCR protocol was set up to amplify the integrase partial coding region in the *pol* gene of HIV-1, yielding 97 positive samples (42.9%). These samples generated 92 nucleotide sequences using the Sanger sequencing method, which were subjected to genotyping through phylogenetic analysis and identification of mutations associated with decreased susceptibility to INSTIs and genetic polymorphisms. The results indicated a predominance of sub-subtype F1 (28.3%), followed by subtypes A (20.7%) and C (19.6%), with subtypes D, G, and H, as well as CRF02_AG, CRF18_cpx, CRF19_cpx, and CRF27_cpx also present. 91.6% of the sequences showed total susceptibility to INSTIs. The major mutation E92G, relevant for epidemiological surveillance, associated with intermediate, low, and potential low-level resistance, along with accessory mutations L74M, Q95K, T97A, and E157Q, associated with potential low-level resistance, were found in a total of 10 sequences. Additionally, numerous genetic polymorphisms were identified, some acting as genetic signatures of subtypes A, C, and F1. The obtained results contribute to the molecular epidemiological surveillance of HIV-1 in Angola, including monitoring of primary resistance to INSTIs.

Keywords: HIV-1, Angola, genetic diversity, integrase inhibitors, primary resistance

Índice

Elementos bibliográficos	i
Agradecimentos.....	ii
Resumo.....	iii
Abstract.....	iv
Índice de Figuras.....	viii
Índice de Tabelas	xi
Lista de Siglas, Acrónimos e Abreviaturas.....	xii
1. Introdução	1
1.1 Contextualização histórica	1
1.1.1 Origem do HIV-1, transmissão zoonótica e disseminação pelo mundo	1
1.2 Caracterização do vírus da imunodeficiência humana do tipo 1 (HIV-1)	3
1.2.1 Taxonomia	3
1.2.2 Morfologia da partícula viral	3
1.2.3 Organização estrutural do genoma	4
1.3 Diversidade genética e epidemiologia	6
1.3.1 Classificação filogenética	6
1.3.1.1 Impacto epidemiológico dos diferentes subtipos do HIV-1	7
1.3.1.2 Mecanismos geradores de diversidade genética	8
1.3.2 Distribuição geográfica	9
1.3.2.1 Diversidade genética e prevalência do HIV-1 em África	12
1.3.2.2 Diversidade genética e prevalência do HIV-1 em Angola	14
1.4 Caracterização da integrase, respetiva região codificante e função no ciclo replicativo.....	16
1.5 Antirretrovirais inibidores de transferência de cadeia da integrase (INSTI) ..	19
1.5.1 Esquemas terapêuticos	20
1.6 Mecanismo de inibição dos INSTI	22
1.7 Tipos de mutações de resistência aos antirretrovirais no HIV-1	22
1.7.1 Mutações de resistência principais e acessórias	23
1.7.1.1 Resistência aos antirretrovirais transmitida (TDR)	24
1.8 Caracterização das mutações de resistência aos INSTI	25

1.9	Monitorização das resistências e sua importância	28
1.10	Objetivo do trabalho	29
2.	Material e métodos	31
2.1	Caracterização geral da amostra populacional em estudo	31
2.1.1	Colheita das amostras e rastreio da infecção pelo HIV-1	31
2.2	Preparação de DNA proviral do HIV-1 a partir de amostras de sangue seco imobilizadas em cartões FTA [®]	32
2.3	Amplificação de DNA proviral do HIV-1 (região codificante da integrase) por <i>nested</i> PCR	34
2.3.1	Caracterização dos <i>primers</i> utilizados	34
2.3.2	Planificação do protocolo de amplificação	35
2.3.2.1	Reações de PCR	37
2.4	Visualização dos produtos de PCR	38
2.5	Purificação dos produtos de PCR	39
2.6	Sequenciação nucleotídica dos produtos de PCR nas duas cadeias pelo método de Sanger	40
2.7	Análise bioinformática das sequências nucleotídicas	40
2.7.1	Edição e confirmação de identidade	40
2.7.2	Genotipagem automatizada	41
2.7.3	Análise filogenética	41
2.8	Análise bioinformática das sequências aminoacídicas	42
2.8.1	Pesquisa de mutações associadas a resistência aos INSTI	42
2.9	Análise estatística	43
3.	Resultados e discussão	45
3.1	Caracterização da amostra populacional em estudo	45
3.2	Amplificação da região codificante da integrase por <i>nested</i> PCR	47
3.3	Análise bioinformática das sequências nucleotídicas	50
3.3.1	Edição e confirmação de identidade	50
3.3.2	Genotipagem automatizada	50
3.3.3	Análise filogenética	51
3.4	Análise bioinformática das sequências aminoacídicas e pesquisa de mutações de resistência associadas aos inibidores da integrase	61
3.4.1	Identificação de mutações associadas a resistência aos INSTI e outros polimorfismos genéticos	63

4. Conclusões	70
5. Referências bibliográficas.....	73
6. Anexos.....	90
Anexo 1. Póster apresentado no 6.º Congresso Nacional de Medicina Tropical	90
Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo.....	91
Anexo 3. Protocolo para preparação de cartões <i>FTA</i> ®	104
Anexo 4. Nomenclatura dos nucleótidos e aminoácidos	105
Anexo 5. Representação do genoma mosaico das formas recombinantes circulantes....	107
Anexo 6. Listagem dos polimorfismos genéticos identificados nas 92 sequências aminoacídicas putativas em estudo e associação com significado estatístico a subtipos ($p < 0,01$) (assinaturas genéticas de subtipo), recorrendo ao teste exato de Fisher bilateral ou χ^2	112
Anexo 7. Composição dos tampões utilizados na realização do presente estudo.....	133

Índice de Figuras

- Figura 1.** Árvores filogenéticas demonstrando as relações filogenéticas entre sequências do gene *pol* de **a)** diferentes lentivírus de mamíferos, incluindo dos primatas; **b)** diferentes estirpes do SIV_{cpzPts}, SIV_{gor}, SIV_{cpzPtt} e dos vários grupos do HIV-1 (5). 1
- Figura 2.** Representação geográfica da região da África central-ocidental, destacando as diferentes áreas de distribuição das subespécies de chimpanzés e das espécies de gorilas, cada uma identificada por cores distintas. O círculo amarelo indica a localização específica, no sudeste dos Camarões, onde foram encontradas estirpes do SIV_{cpzPtt} relacionadas filogeneticamente com o grupo M do HIV-1 (6). 2
- Figura 3.** Morfologia e organização estrutural da partícula viral imatura e virião do HIV-1, respectivamente (adaptado de ViralZone. https://viralzone.expasy.org/7?outline=all_by_species). 3
- Figura 4.** Representação esquemática do genoma proviral do HIV-1, com indicação das diferentes regiões codificantes. Genes estruturais: *gag*, *pol* e *env*. MA: proteína da matriz; CA: proteína da cápside; NC: proteína da nucleocápside. PR: protease; RT: transcriptase reversa; IN: integrase. SU: glicoproteína de superfície; TM: glicoproteína transmembranar. Genes reguladores: *rev* e *tat*. Genes acessórios: *vpr*, *vif*, *vpu* e *nef*. Nas extremidades 5' e 3' estão representadas as LTRs, com estrutura U3-R-U5. (Adaptado de Biorender. <https://www.biorender.com/>). 5
- Figura 5.** Distribuição geográfica mundial dos diferentes subtipos e formas recombinantes do HIV-1 (32). 10
- Figura 6.** Prevalência dos diferentes subtipos, sub-subtipos e CRFs do HIV-1 na África ocidental (**a**) e oriental (**b**) (36). 13
- Figura 7.** Organização dos domínios da IN. Os números dos aminoácidos (a.a.) delimitam o início e o fim de cada domínio: o domínio N-terminal (NTD; a vermelho), o domínio do núcleo catalítico (CCD; a azul) e o domínio C-terminal (CTD; a verde) (81). Os tamanhos (em a.a.) dos *linkers* entre os domínios estão indicados por baixo das linhas a preto. O íon Zn²⁺ no NTD é apresentado como uma esfera de cor azul-clara; um dos dois íons Mg²⁺, ligado ao centro ativo do CCD, é representado por uma esfera de cor verde-amarelada (81). 18

Figura 8. Fases do ciclo replicativo do HIV-1 correspondentes à transcrição reversa e integração do genoma viral no cromossoma do hospedeiro. A verde está representado o genoma viral de RNA, a azul o DNA viral retrotranscrito e a cinzento o genoma do hospedeiro (adaptado de Andrade <i>et al.</i> , 2015) (81).....	19
Figura 9. Mutações na região codificante da IN associadas a resistência aos INSTI (adaptado de Wensing <i>et al.</i> , 2022) (105).....	27
Figura 10. Posição relativa dos <i>primers</i> diretos (INPS1 e INPS3) e reversos (INPR9 e INPR8) na sequência nucleotídica da região codificante da IN (retirado de <i>Snapgene</i> . https://www.snapgene.com/).	35
Figura 11. Representação gráfica (em percentagem) dos dados demográficos da amostra populacional em estudo, relativos ao sexo, idade, estado civil e nível de escolaridade. 45	
Figura 12. Representação gráfica (em percentagem) dos dados demográficos da amostra populacional em estudo relativos à área de residência, ao local de residência e outras infeções.	46
Figura 13. Representação gráfica (em percentagem) dos resultados obtidos após amplificação por nested PCR da região codificante da IN.	47
Figura 14. Visualização dos padrões de migração dos produtos de amplificação por <i>nested</i> PCR da região codificante da IN do HIV-1, após eletroforese em gel de agarose a 1,5% (p/v), em tampão TAE 0,5X, corado com <i>Green Safe Premium</i> (NZYTech, Portugal). Legenda: <i>NZYTech DNA Ladder V</i> – Marcador de massa molecular (100-1000 pb); 379-388 – Amostras em estudo; H ₂ O e H ₂ O _c – Controlos negativos.....	49
Figura 15. Visualização dos padrões de migração dos produtos de amplificação por <i>nested</i> PCR da região codificante da IN do HIV-1, após eletroforese em gel de agarose a 1,5% (p/v), em tampão TAE 0,5X, corado com <i>Green Safe Premium</i> (NZYTech, Portugal). Legenda: <i>NZYTech DNA Ladder V</i> – Marcador de massa molecular (100-1000 pb); 367-536 – Amostras em estudo; H ₂ O e H ₂ O _c – Controlos negativos.....	49
Figura 16. Representação gráfica (em percentagem) dos subtipos obtidos após genotipagem preliminar automatizada das sequências nucleotídicas em estudo, utilizando o <i>software REGA HIV-1 Subtyping Tool</i> (versão 3.46). NA é a designação para as sequências cujos génotipos não foram atribuídos.....	51

Figura 17. Relações filogenéticas entre 150 sequências nucleotídicas da região codificante da integrase, baseadas num alinhamento múltiplo de 682 nucleótidos. Na árvore filogenética foram incluídas as 92 sequências em estudo (indicadas com o símbolo), 27 sequências de referência do grupo M (A1-A4, A6-A8, B, D, F1, F2, G, H, J, K, L, CRF18_cpx e CRF27_cpx) e uma do grupo O (como *outgroup*) e 30 sequências obtidas por pesquisa de semelhança local entre sequências (análise *BLAST*). Com exceção das sequências em estudo, a designação de cada sequência inicia-se com o respetivo genótipo, seguido do código do país de origem (AO – Angola, AU – Austrália, BE – Bélgica, BR – Brasil, BW – Botsuana, CD – República Democrática do Congo, CF – República Centro-Africana, CM – Camarões, CN – China, CU – Cuba, CV – Cabo Verde, ES – Espanha, ET – Etiópia, FI – Finlândia, FR – França, GB – Reino Unido, GH – Gana, IN – Índia, MW – Maláui, NG – Nigéria, PT- Portugal, SE – Suécia, SN – Senegal, TH – Tailândia, UA – Ucrânia, UG – Uganda, ZA – África do Sul, ZM – Zâmbia), nome da estirpe/isolado viral e número de acesso no *GenBank*. A inferência filogenética foi realizada com recurso ao *software MEGA* versão 11, 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 2% de distância genética). A robustez da inferência filogenética foi analisada por *bootstrapping* de 1000 amostragens aleatórias, considerando-se significativos valores $\geq 70\%$. As sequências nas quais foi detetada a ocorrência de fenómenos de hipermutação estão indicadas com o símbolo. 57

Figura 18. a) Representação esquemática de todas as mutações encontradas associadas a resistência aos INSTI, na sequência codificante da IN. As letras superiores representam o aminoácido (código de uma letra) da sequência consenso do subtipo B (suscetibilidade), o número refere-se à sua posição na sequência e as letras inferiores indicam o aminoácido associado à mutação que confere resistência aos INSTI. A negrito está destacada a mutação principal, enquanto as restantes são mutações acessórias. **b)** Representação gráfica da prevalência das mutações associadas a resistência aos INSTI identificadas na região codificante da IN das sequências em estudo..... 64

Figura 19. Representação gráfica da distribuição relativa dos níveis de resistência fenotípicos putativos aos cinco INSTI: Elvitegravir (EVG), Raltegravir (RAL), Bictegravir (BIC), Dolutegravir (DTG) e Cabotegravir (CAB); associados às sequências da região codificante da IN analisadas. 68

Índice de Tabelas

Tabela 1. Exemplos de diferenças significativas em diferentes formas genéticas do HIV-1 relativamente a diversas características epidemiológicas.	8
Tabela 2. Lista de mutações de resistência transmitida aos INSTI, utilizadas na vigilância epidemiológica da resistência primária (baseado na <i>HIV Drug Resistance Database, Stanford University, 2020</i>) (111).	28
Tabela 3. Descrição dos <i>primers</i> externos e internos da reação de <i>nested</i> PCR. Posições nucleotídicas retiradas da base de dados <i>Los Alamos National Laboratory</i> , ferramenta <i>Sequence locator</i> . https://www.hiv.lanl.gov/content/sequence/LOCATE/locate.html ..	34
Tabela 4. Protocolo de amplificação para a 1. ^a reação da <i>nested</i> PCR, com a utilização dos <i>primers</i> externos INPS1 e INPR8.	37
Tabela 5. Protocolo de amplificação para a 2. ^a reação da <i>nested</i> PCR, com a utilização dos <i>primers</i> internos INPS3 e INPR9.	37
Tabela 6. Resultados da genotipagem para a região codificante da IN das 92 sequências em estudo após construção da árvore filogenética (NA= genótipo não atribuído).....	54
Tabela 7. Assinaturas genéticas identificadas nas sequências da região codificante da IN em estudo pertencentes aos subtipos A, C e F1 ($p < 0,01$). Os polimorfismos assinalados a negrito indicam uma associação altamente significativa ao respetivo genótipo ($p < 0,00001$). O código internacional IUPAC de uma letra foi utilizado para representar os aminoácidos em cada posição (Anexo 4).....	69

Lista de Siglas, Acrónimos e Abreviaturas

ADR – Resistência aos antirretrovirais adquirida (do inglês *acquired drug resistance*)

ANRS – *Agence Nationale de Recherches sur le Sida et les Hépatites Virales*

APOBEC – *Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like*

ARVs – Antirretrovirais

BIC – Bictegravir

BLAST – *Basic Local Alignment Search Tool*

CAB – Cabotegravir

CCD – Domínio do núcleo catalítico (do inglês *catalytic core domain*)

CDC – Centros de Controle e Prevenção de Doenças (do inglês *Centers for Disease Control and Prevention*)

CRF – Forma recombinante circulante (do inglês, *circulating recombinant form*)

CTD – Domínio C-terminal (do inglês *C-terminal domain*)

ddNTP – Didesoxinucleótidos (do inglês *dideoxynucleotides*)

DNA – Ácido desoxirribonucleico (do inglês *deoxyribonucleic acid*)

DRM – Mutações de resistência aos antirretrovirais (do inglês *drug resistance mutations*)

DTG – Dolutegravir

EMA – *European Medicines Agency*

EVG – Elvitegravir

FDA – *Food and Drug Administration*

FV – Falência virológica

HAART – Terapia antirretroviral de alta eficácia (do inglês *highly active antiretroviral therapy*)

HIVdb – *Genotypic Resistance Interpretation Algorithm da Stanford University HIV Drug Resistance Database*

HIVDR – resistência a fármacos contra o HIV (do inglês *HIV drug resistance*)

HIV-1 – Vírus da imunodeficiência humana do tipo 1 (do inglês *human immunodeficiency virus type 1*)

HIV-2 – Vírus da imunodeficiência humana do tipo 2 (do inglês *human immunodeficiency virus type 2*)

HSH – Homens que fazem sexo com outros homens

IAS-USA – *International Antiviral Society-USA*

IHMT – Instituto de Higiene e Medicina Tropical

IN – Integrase (do inglês *integrase*)

INLS – Instituto Nacional de Luta contra a SIDA (de Angola)

INSTI – Inibidores de transferência de cadeia da integrase (do inglês *integrase strand transfer inhibitors*)

IST – Infecção sexualmente transmissível

IUPAC – *International Union of Pure and Applied Chemistry*

LEDGF/p75 – *Lens Epithelium-derived Growth Factor/p75*

LTR – *Long Terminal Repeat*

mRNA – RNA mensageiro

NNRTI – Inibidores da transcriptase reversa não análogos a nucleósidos (do inglês *non-nucleoside reverse transcriptase inhibitors*)

NRTI – Inibidores da transcriptase reversa análogos a nucleósidos (do inglês *nucleoside reverse transcriptase inhibitors*)

NTD – Domínio N-terminal (do inglês *N-terminal domain*)

nts – Nucleótidos

OMS – Organização Mundial da Saúde

PALOPs – Países Africanos de Língua Oficial Portuguesa

PDR – Resistência aos antirretrovirais pré-tratamento (do inglês *pre-treatment drug resistance*)

PEPFAR – Plano de Emergência do Presidente (dos EUA) para o Alívio da SIDA (do inglês *President's Emergency Plan for AIDS Relief*)

PFV – *Prototype foamy virus*

PI – Inibidores da protease (do inglês *protease inhibitors*)

PIC – Complexo de pré-integração (do inglês *pre-integration complex*)

PR – Protease (do inglês *protease*)

PrEP – Profilaxia de pré-exposição (do inglês *pre-exposure prophylaxis*)

RAL – Raltegravir

RC – República do Congo

RDC – República Democrática do Congo

RNA – Ácido ribonucleico (do inglês *ribonucleic acid*)

RT – Transcriptase reversa (do inglês *reverse transcriptase*)

SDRM – *Surveillance drug resistance mutations*

SIDA – Síndrome da Imunodeficiência Adquirida

SIV – Vírus da imunodeficiência símia (do inglês *simian immunodeficiency virus*)

SIV_{cpzPtt} – SIV do chimpanzé-central (subespécie *Pan troglodytes troglodytes*)

SIV_{gor} – SIV do gorila-ocidental-das-terras-baixas (subespécie *Gorilla gorilla gorilla*)

SIV_{smm} – SIV do mangabei-fuliginoso (espécie *Cercocebus atys*)

SU – Glicoproteína de superfície (do inglês *surface glycoprotein*)

TA – Temperatura ambiente

TDF – Tenofovir

TDR – Resistência aos antirretrovirais transmitida (do inglês *transmitted drug resistance*)

TM – Glicoproteína transmembranar (do inglês *transmembrane glycoprotein*)

UNAIDS – *Joint United Nations Programme on HIV/AIDS*

URF – Forma recombinante única (do inglês *unique recombinant form*)

3TC – Lamiduvina

1. Introdução

1.1 Contextualização histórica

1.1.1 Origem do HIV-1, transmissão zoonótica e disseminação pelo mundo

O vírus da imunodeficiência humana do tipo 1 (HIV-1, do inglês *human immunodeficiency virus type 1*) foi isolado pela primeira vez associado à síndrome da imunodeficiência adquirida (SIDA) em 1983, por Françoise Barré-Sinoussi e colaboradores (1–3). Este vírus resultou de múltiplos momentos de transmissão de carácter zoonótico de vírus da imunodeficiência dos símios (SIV, do inglês *simian immunodeficiency virus*) (Figura 1a) de primatas não-humanos para humanos, em África (4). Momentos estes que, distintos e independentes, deram origem a várias linhagens genéticas do HIV-1, classificadas como diferentes grupos, nomeadamente M, N, O e P (2,4,5). Os HIV-1 dos grupos M (pandémico) e N são originários de um mesmo vírus, o SIV_{cpzPtt}, (Figura 1b), que infeta uma das subespécies de chimpanzés, *Pan troglodytes troglodytes*, que se encontra no sudeste e centro-sul dos Camarões (4–6). Os restantes dois grupos (O e P), pensa-se terem derivado de transmissões independentes de estirpes do SIV_{gor} (Figura 1b), de gorilas para humanos, nas planícies ocidentais dos Camarões (4–6). Coloca-se ainda a questão, do próprio SIV_{gor} ter sido derivado de um ancestral do SIV_{cpzPtt}, de uma linhagem divergente (4). Atualmente, não é claro se os gorilas foram a fonte imediata do HIV-1 do grupo O para os humanos, ou se, pelo contrário, os vírus dos chimpanzés foram transmitidos paralelamente aos gorilas e aos humanos (4,6).

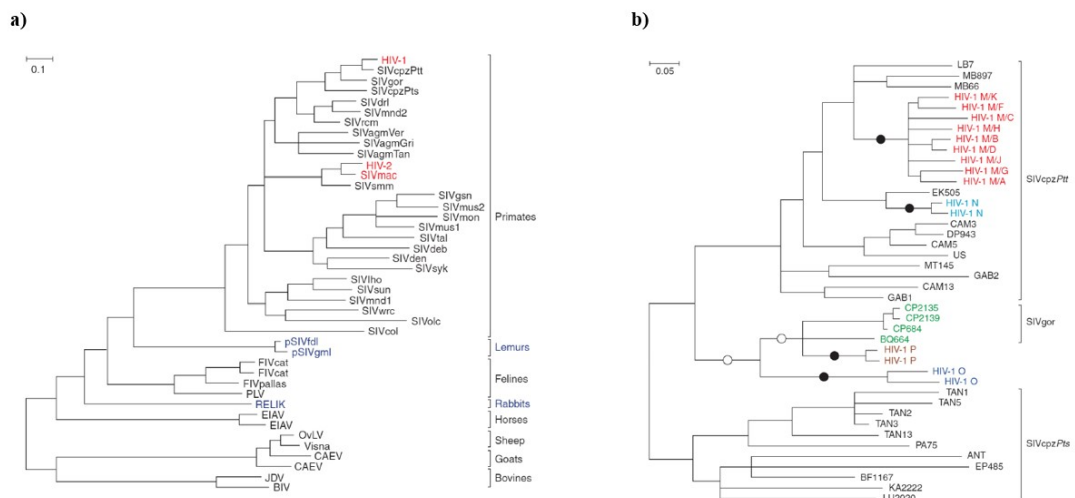


Figura 1. Árvores filogenéticas demonstrando as relações filogenéticas entre sequências do gene *pol* de a) diferentes lentivírus de mamíferos, incluindo dos primatas; b) diferentes estirpes do SIV_{cpzPts}, SIV_{gor}, SIV_{cpzPt} e dos vários grupos do HIV-1 (5).

Após a transmissão para os humanos, por volta do início do século XX (4), o HIV-1 do grupo M registou um aumento do número de infeções de forma lenta, na primeira metade deste século, tendo sido documentada a origem posterior da dispersão pandémica do grupo M do HIV-1 na cidade de Kinshasa, na República Democrática do Congo (RDC) (4–7). Este facto explica a razão pela qual esta cidade apresenta a grande diversidade genética atual (2,8), comparativamente com outros locais.

A transmissão zoonótica dos chimpanzés para os humanos terá ocorrido no sudeste dos Camarões (2,4–6), local onde os chimpanzés infetados com estirpes de SIV_{cpzPtt} filogeneticamente mais próximas do grupo M foram identificados (Figura 2) (2,4,6). Esta transmissão provavelmente esteve relacionada com a caça e manuseamento de primatas (4) que ocorria nos Camarões (4–6). Em meados da década de 1980, do século XX, os eventos de dispersão foram iniciados a partir de locais secundários, estabelecendo-se assim a distribuição geograficamente heterogénea deste vírus observada em toda a África (2).

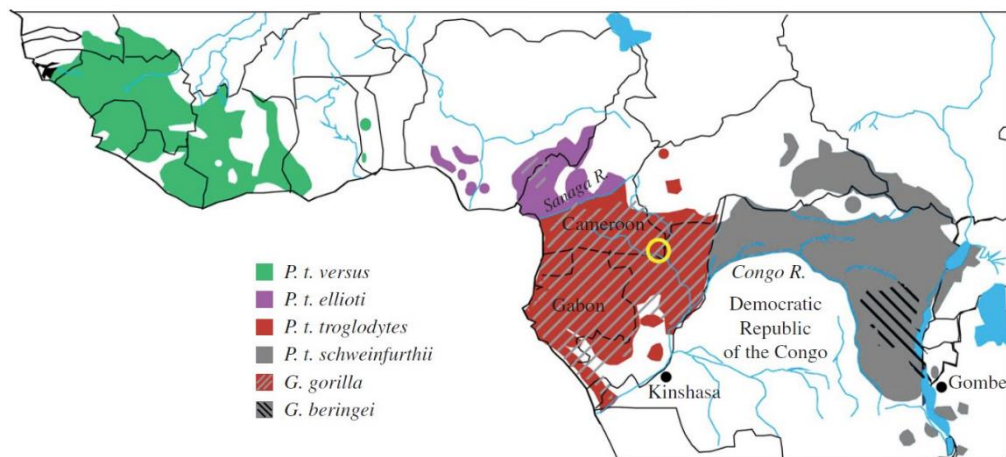


Figura 2. Representação geográfica da região da África central-ocidental, destacando as diferentes áreas de distribuição das subespécies de chimpanzés e das espécies de gorilas, cada uma identificada por cores distintas. O círculo amarelo indica a localização específica, no sudeste dos Camarões, onde foram encontradas estirpes do SIV_{cpzPtt} relacionadas filogeneticamente com o grupo M do HIV-1 (6).

Dados do *Joint United Nations Programme on HIV/AIDS* (UNAIDS) estimam que, aproximadamente, 39 milhões de pessoas viviam com HIV no final de 2022, com cerca de 1,3 milhões de novos infetados nesse ano e um número estimado de mortes de 630 mil.

1.2 Caracterização do vírus da imunodeficiência humana do tipo 1 (HIV-1)

1.2.1 Taxonomia

O HIV-1 foi classificado no domínio *Riboviria*, reino *Pararnavirae*, filo *Artverviricota*, classe *Revtraviricetes*, ordem *Ortervirales* e família *Retroviridae* (9,10). A família *Retroviridae* inclui duas subfamílias *Orthoretrovirinae* e *Spumaretrovirinae*, sendo que o HIV-1 pertence à subfamília *Orthoretrovirinae*, género *Lentivirus* (9–11).

1.2.2 Morfologia da partícula viral

O HIV-1 é um vírus com invólucro, de forma esférica a pleomórfica com cerca de 80-120 nm de diâmetro e simetria complexa (12).

A partícula viral pode apresentar-se sob a forma de madura (virião) ou imatura (Figura 3) (12), diferindo na organização estrutural das proteínas virais (12), sendo que a última corresponde a uma forma não infecciosa. A passagem para viriões (infecciosos) coincide com a saída das células hospedeiras ou ocorre já no exterior (13).

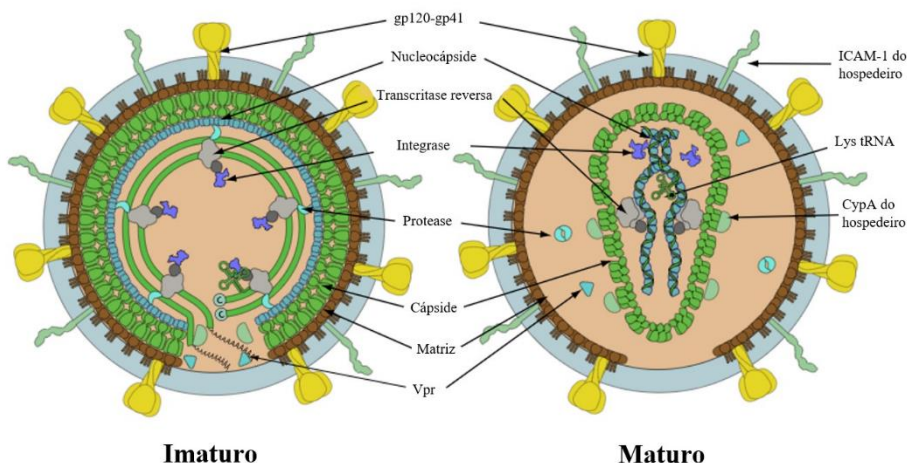


Figura 3. Morfologia e organização estrutural da partícula viral imatura e virião do HIV-1, respetivamente (adaptado de *ViralZone*. https://viralzone.expasy.org/7?outline=all_by_species).

O HIV-1 apresenta um invólucro lipídico, derivado da membrana citoplasmática das células hospedeiras e ao qual estão associadas as glicoproteínas do invólucro (gp120 e gp41), que permitem a entrada do vírus nas células hospedeiras (12). A segunda camada, subjacente ao invólucro, é constituída pela proteína da matriz e a terceira é composta pela proteína da cápside (12). Por fim, no vírus imaturo, existe uma última camada, mais

interna, constituída pela proteína da nucleocápside, enquanto no virião, esta proteína apresenta-se associada às duas moléculas de ácido ribonucleico (RNA, do inglês *ribonucleic acid*) que vão constituir o genoma deste vírus (Figura 3) (12). As glicoproteínas do invólucro e as proteínas da matriz, cápside e nucleocápside são designadas de proteínas estruturais, fazendo ainda parte da constituição da partícula viral, enzimas e algumas proteínas acessórias e reguladoras.

1.2.3 Organização estrutural do genoma

O HIV-1 tem genoma de RNA de cadeia simples, com polaridade positiva. Cada partícula viral contém duas cópias do RNA genómico normalmente idênticas (14). Cada molécula de RNA codifica a totalidade da informação genética necessária para a replicação viral (14). As duas moléculas estão entrelaçadas, sob a forma de um complexo ribonucleoproteico, juntamente com as proteínas da nucleocápside (12,14). No ciclo replicativo existe um passo de transcrição reversa, no qual é garantida a conversão da informação do RNA numa molécula de ácido desoxirribonucleico (DNA, do inglês *deoxyribonucleic acid*), de cadeia dupla, que é posteriormente integrada no genoma do hospedeiro (11), tomando então o nome de provírus.

O genoma do HIV-1 tem cerca de 9,2 kb (15) e é constituído por 3 genes principais (genes estruturais): o gene *gag* (com 1,5 kb), o gene *pol* (com 2,8 kb) e o gene *env* (com 2,6 kb), como se pode observar na Figura 4 (7,16). O gene *gag* codifica uma proteína precursora de cerca de 55 kDa (13), cuja clivagem pela protease viral é estritamente necessária para a maturação do vírus (13). Esta clivagem dá origem às proteínas estruturais, matriz, cápside e nucleocápside (17). O gene *pol* codifica para uma poliproteína precursora de cerca 160 kDa (11,13), cuja clivagem pela protease viral (13,18) dá origem a três enzimas essenciais: protease (PR, do inglês *protease*), que é necessária para a maturação do vírus; transcriptase reversa (RT, do inglês *reverse transcriptase*), que permite a transcrição de RNA genómico de cadeia simples a DNA de cadeia dupla; e integrase (IN, do inglês *integrase*), que permite a integração do DNA de cadeia dupla no genoma do hospedeiro, passando a DNA proviral (13,19). O gene *gag* e *pol* apresentam uma sobreposição de cerca de 200 nucleótidos (nts) (Gag-Pol) (20). Desta forma, a síntese das proteínas neles

codificadas necessita de dois passos distintos. Síntese de RNA mensageiro (mRNA) Gag, dando origem à poliproteína Gag, e a síntese de um mRNA Gag/Pol, que irá corresponder ao genoma total, apresentando dupla função: traduzir as enzimas Pol, sob a forma Gag-Pol, através de um *frameshift* ribossomal aquando da tradução e assegurar a encapsidação das mesmas (21). O mecanismo de *frameshift* consiste na alteração do quadro de leitura de um mRNA, retrocedendo um nucleótido durante a tradução, passando a evitar um codão *stop*, neste caso, no final do gene *gag* (21), e sintetizando, assim, Gag-Pol. Por fim, o gene *env* codifica uma poliproteína glicosilada de cerca de 160 kDa (gp160), que é subsequentemente clivada por uma protease da célula hospedeira, dando origem às duas glicoproteínas maduras do invólucro viral, sendo elas gp120, glicoproteína de superfície (SU, do inglês *surface glycoprotein*) e gp41, glicoproteína transmembranar (TM, do inglês *transmembrane glycoprotein*) (22). Para além dos genes estruturais, existem também genes reguladores, como *tat* e *rev*, que codificam para proteínas reguladoras (Tat e Rev) e genes acessórios *vpr*, *vif*, *vpu* e *nef*, que codificam para proteínas acessórias (Vpr, Vif, Vpu e Nef) (12,16).

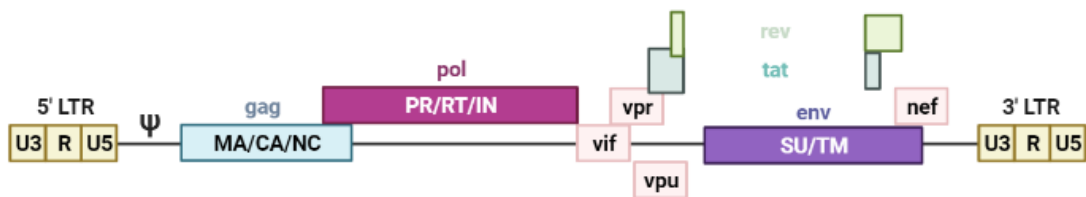


Figura 4. Representação esquemática do genoma proviral do HIV-1, com indicação das diferentes regiões codificantes. Genes estruturais: *gag*, *pol* e *env*. MA: proteína da matriz; CA: proteína da cápside; NC: proteína da nucleocápside. PR: protease; RT: transcriptase reversa; IN: integrase. SU: glicoproteína de superfície; TM: glicoproteína transmembranar. Genes reguladores: *rev* e *tat*. Genes acessórios: *vpr*, *vif*, *vpu* e *nef*. Nas extremidades 5' e 3' estão representadas as LTRs, com estrutura U3-R-U5. (Adaptado de Biorender. <https://www.biorender.com/>).

Uma outra particularidade deste genoma é que, na sua forma proviral, apresenta regiões reguladoras da expressão génica, localizadas em ambos os extremos e designadas de *Long Terminal Repeats* (LTRs) (12) (Figura 4). As LTRs apresentam a estrutura “U3-R-U5” e, apesar de idênticas, em 5' e em 3', têm funções distintas: a LTR-5' funciona como promotor para a transcrição dos genes virais e a LTR-3' permite a poliadenilação dos RNAs virais, incluindo o genoma (12).

1.3 Diversidade genética e epidemiologia

1.3.1 Classificação filogenética

Existem dois lentivírus geneticamente distintos, tendo sido transmitidos aos humanos através de eventos de transmissão de carácter zoonótico diferentes. Os diferentes SIV, nomeadamente SIV_{cpzPtt}, SIV_{gor} e o SIV do mangabei-fuliginoso, que infeta a espécie *Cercocebus atys* (SIV_{smm}), ultrapassaram a barreira das espécies em múltiplas ocasiões e, dessa forma, deram origem a dois tipos diferentes de HIV, os dois primeiros ao HIV-1 e o último ao vírus da imunodeficiência humana do tipo 2 (HIV-2, do inglês *human immunodeficiency virus type 2*) (4,6). Dos grupos pertencentes ao HIV-1, já mencionados, o grupo M é o único considerado pandémico e representa mais de 95% de todas as estirpes virais isoladas (23). Os grupos O, N e P contribuem menos significativamente para essa mesma pandemia (representando aproximadamente 1-2% das infeções) (23). Contudo, as infeções por HIV-1 do grupo O demonstram-se de maior preocupação comparativamente aos outros dois grupos, uma vez que já foram descritas em diferentes países da África central-ocidental, bem como na Europa e Estados Unidos da América (EUA), ainda que sempre relacionados com casos provenientes dos Camarões (24). Este grupo encontra-se maioritariamente nos Camarões, representando 0,6 a 1% de todas as infeções por HIV neste país (24). Até à data, só foram descritos 22 casos de infeção pelo HIV-1 do grupo N, todos nos Camarões, exceto um, diagnosticado em França (cujo doente era de origem togolesa) e foram notificados dois casos de infeção por HIV-1 do grupo P em doentes de origem camaronesa (24).

Atualmente, a literatura descreve o grupo M como subdivido em 10 subtipos diferentes (A-D, F-H, J-L) sendo que os subtipos A e F, devido à sua elevada diversidade genética, ainda são divididos em sub-subtipos (A1-A8 e F1-F2) (23,25–29). A classificação em subtipos é baseada em análises da sequência genómica completa dos isolados (existirem pelo menos 3) ou nas sequências nucleotídicas derivadas de regiões subgenómicas do gene *gag*, *pol* e *env* (pelo menos 2 sequências com tamanho superior a 8 kb, juntamente com outras regiões subgenómicas) (30). A diversidade genética dentro de um mesmo subtipo varia habitualmente entre 8-17% (máximo de 30%), enquanto que entre subtipos diferentes tem uma amplitude entre 17-35% (máximo de 42%) (31).

Para além desta classificação em subtipos, existem ainda as chamadas formas recombinantes circulantes (CRFs, do inglês *circulating recombinant forms*) e únicas (URFs, do inglês *unique recombinant forms*). Estas formas surgem de fenómenos de recombinação genética, após a coinfeção/superinfecção de um mesmo indivíduo por, pelo menos, dois vírus geneticamente distintos (23,32). Uma CRF resulta da recombinação entre genomas virais provenientes de diferentes subtipos, dando origem a um genoma mosaico, composto por várias regiões genómicas de cada subtipo (32). As CRFs são definidas como sequências completas, ou quase completas, com os mesmos padrões de recombinação, de estirpes virais encontradas em, pelo menos, três indivíduos não relacionados epidemiologicamente (23,32). As CRFs são nomeadas pela ordem em que são reportadas e pelos subtipos recombinantes que compõem o seu genoma (4,23). A designação de “cpx” indica uma forma recombinante circulante complexa, o que significa que possuem mais do que dois subtipos diferentes no mesmo genoma mosaico (4,23,32). As CRFs com maior prevalência a nível mundial são a CRF01_AE e CRF02_AG. Existem, até à data, cerca de 157 CRFs descritas (33). As URFs representam um grupo heterogéneo de estirpes e o seu aparecimento baseia-se, também, na recombinação genética de subtipos. Os seus genomas mosaico mostram pontos de recombinação únicos e, ao contrário das CRFs, são apenas encontradas uma vez, em apenas um indivíduo ou em grupos de indivíduos epidemiologicamente relacionados (32).

1.3.1.1 Impacto epidemiológico dos diferentes subtipos do HIV-1

Vários estudos feitos até à data têm demonstrado que os diferentes subtipos do HIV-1 do grupo M podem ter diferenças significativas relativamente a certas características, como a transmissibilidade, sintomatologia, progressão para a doença, carga viral, virulência (4,34–36), transmissão vertical (37), etc. Em 1999, já eram conhecidas diferenças na transmissibilidade e virulência do HIV-1 em relação ao HIV-2, apresentando potenciais patogénicos distintos (38). Concluiu-se que o HIV-2 é menos transmissível e a progressão para SIDA é mais lenta, comparativamente ao HIV-1 (38). Assim, colocou-se a hipótese de, também entre subtipos diferentes do HIV-1, poder haver diferenças significativas nas suas características epidemiológicas (38). Na Tabela 1 encontram-se descritas algumas dessas características, relativamente aos diferentes subtipos virais, embora algumas das associações sugeridas permaneçam controversas.

Tabela 1. Exemplos de diferenças significativas em diferentes formas genéticas do HIV-1 relativamente a diversas características epidemiológicas.

Subtipo A	<ul style="list-style-type: none"> • Tem menor propensão para atingir o estado de SIDA comparativamente com subtipos não-A (39); • Tem maiores taxas de transmissibilidade que o subtipo D (39); • A falha no tratamento antirretroviral, bem como as taxas de resistência aos fármacos, são menos frequentes do que em indivíduos infetados com o subtipo D (35).
Subtipo B	<ul style="list-style-type: none"> • A atividade replicativa da RT é superior relativamente ao subtipo C (7,40); • Poderá estar associado a uma progressão para a doença superior à da CRF07_BC, mas inferior à da CRF01_AE (41).
Subtipo C	<ul style="list-style-type: none"> • A transmissibilidade é superior à de vírus dos subtipos A ou D (7); • Tem uma capacidade replicativa inferior à dos outros subtipos (7,39), tornando-o menos virulento, comparativamente aos subtipos A e D (39); • Está associado com taxas de declínio de células T CD4+ significativamente menores, bem como menores frequências de progressão para a doença, comparativamente com os outros subtipos (35,39).
Subtipo D	<ul style="list-style-type: none"> • Associado à diminuição da sobrevivência em crianças infetadas por transmissão vertical (37); • A progressão para a doença é mais rápida comparativamente ao subtipo A, independentemente da carga viral (7,34,35,39,42); • Frequência mais elevada na formação de sincícios (células multinucleadas) e uma taxa replicativa superior à de isolados de outros subtipos (35); • Associado a um mais rápido declínio na contagem de células T CD4+, comparativamente aos subtipos A e B, e a uma mortalidade maior comparativamente ao subtipo B (35).
Subtipo G	<ul style="list-style-type: none"> • Associados à recorrência e maior gravidade das síndromes neurológicas causadas pelo HIV-1 [incluindo um desempenho cognitivo significativamente pior, comparativamente, por exemplo, com infeções causadas pela CRF02_AG (35,43)]; • Mais frequente em indivíduos envolvidos no uso de drogas injetáveis (44).
CRF07_BC	<ul style="list-style-type: none"> • Progressão para a doença mais lenta do que no subtipo B ou CRF01_AE (45).
CRF02_AG	<ul style="list-style-type: none"> • Tem um <i>fitness</i> replicativo maior do que os subtipos progenitores A e G, o que resulta numa carga viral mais elevada e maior transmissibilidade (43).
Recombinantes inter-subtipos A, C e D	<ul style="list-style-type: none"> • São, aparentemente, mais transmissíveis durante a amamentação do que os subtipos puros seus progenitores (37).

1.3.1.2 Mecanismos geradores de diversidade genética

Uma das características do HIV-1 é apresentar uma elevada diversidade genética. Esta diversidade é visível, não só entre indivíduos infetados, como no próprio indivíduo, materializando-se na forma de *quasispecies* (4). Uma *quasispecies* é um conjunto de variantes virais distintas, mas geneticamente relacionadas, presentes num mesmo

indivíduo infetado (46). Neste caso, a seleção natural atua, não sobre um genoma viral singular, mas sim sobre uma grande quantidade de formas genômicas filogeneticamente aparentadas, circulantes no mesmo indivíduo (a diversidade genética pode ultrapassar valores superiores a 10%) (4).

A elevada diversidade genética característica do HIV-1 é consequência de um conjunto de fatores, como: *a*) um *turnover* viral elevado, ou seja, uma rápida replicação viral (estima-se que sejam produzidas até 10^{10} partículas virais/dia/indivíduo infetado) (4,47,48); *b*) uma elevada taxa de erros produzidos pela RT durante a retrotranscrição do genoma viral [tendo em conta que o tamanho do genoma do HIV-1 é aproximadamente 10 000 nts e que uma mutação é introduzida em cada 1 000 – 10 000 nts, isto significa que 1 – 10 mutações poderão ser geradas em cada ciclo de replicação do genoma viral] (4,7,37,47,48); *c*) hipermutação mediada pela APOBEC [hipermutação guanina (G) → adenina (A)] – introduzindo mutações no genoma do vírus, esta proteína do hospedeiro bloqueia a replicação viral, mas se essas mutações não levarem à completa paragem do ciclo replicativo, podem funcionar como geradoras de diversidade genética (48); *d*) elevada taxa de recombinação, promovida pela RT, que tem uma baixa afinidade para os seus substratos (4,7,37,48); *e*) fixação de mutações selecionadas pela pressão exercida pelo sistema imunitário do hospedeiro ou pela terapia antirretroviral de alta eficácia (HAART, do inglês *highly active antiretroviral therapy*) (48,49); e, por fim, *f*) elevada plasticidade das proteínas virais (as sequências codificantes no genoma viral podem acumular muitas mutações e isso não tem, necessariamente, de ser deletério para o vírus) (23,48).

1.3.2 Distribuição geográfica

As formas genéticas do HIV-1 apresentam um padrão de distribuição geográfica heterogêneo, ou seja, existe uma presença predominante de determinados subtipos/formas recombinantes em áreas específicas. É de notar que, devido ao peso significativo das infeções na região do continente africano, cerca de metade das infeções por HIV-1 à escala global correspondem ao subtipo C (50), uma vez que este subtipo é altamente prevalente na África meridional.

Europa e Oceânia

O subtipo B é prevalente no continente europeu e Oceânia (Figura 5). Os subtipos A e C são também encontrados na Europa, conseguindo-se observar um aumento da sua prevalência ao longo dos anos, nomeadamente na zona este deste continente (25). Outras formas genéticas, como F1 e a CRF02_AG, foram introduzidos ao longo dos anos (32) especialmente na Europa. Em Portugal, o subtipo B tem sido o mais prevalente desde 1980, principalmente na população heterossexual (32,44). Mais recentemente, tem estado mais associado à população que usa drogas injetáveis (44). O subtipo G e CRF14_BG estão a tornar-se mais prevalentes em Portugal, especialmente entre pessoas que usam drogas injetáveis (44). De acordo com a literatura foi possível observar um padrão epidemiológico relativamente original, relacionado com uma diferença significativa da prevalência de diferentes subtipos entre indivíduos portugueses e indivíduos imigrantes (originários de África) (51). O subtipo B observou-se mais prevalente em portugueses (66,9%) do que nos indivíduos imigrantes (22,6%), enquanto as estirpes não-B foram encontradas maioritariamente entre os africanos (77,4%) (51). Noutro estudo realizado recentemente, conseguiu-se observar esta diferença evidente de distribuição de subtipos B vs. não-B, entre a população de origem portuguesa vs. população imigrante, neste caso entre indivíduos de nacionalidade portuguesa, brasileira e africana (Países Africanos de Língua Oficial Portuguesa, PALOPs) (52).

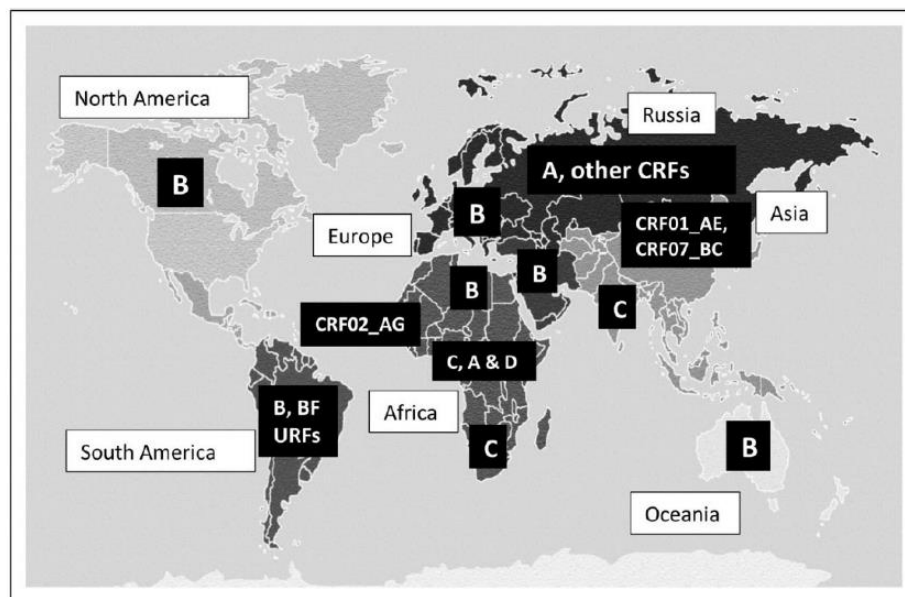


Figura 5. Distribuição geográfica mundial dos diferentes subtipos e formas recombinantes do HIV-1 (32).

Américas

O subtipo B é o subtipo geograficamente predominante em todo o mundo (Figura 5), nomeadamente no continente americano (32,53), a América do Norte e do Sul são um exemplo desta predominância (32). Recombinantes inter-subtipo têm também sido reportados (32), como se pode observar na América do Sul, verificando-se a presença de vários recombinantes BF (32,54). A prevalência do subtipo C no sul do Brasil tem vindo a aumentar, existindo ainda uma pequena contribuição do sub-subtipo F1, bem como outros subtipos e recombinantes inter-subtipo na região norte (32,53).

Médio Oriente e norte de África

Os países do Médio Oriente e do norte de África apresentam em comum o facto de o subtipo B ser predominante (32). A epidemia em Israel inicialmente dominada pelo subtipo B entre HSH, passou a ter a presença dos subtipos A e C, incluindo novos sub-subtipos, como A1, A6-A8, devido à migração massiva da Etiópia e dos países da União Económica Eurasiática (55).

África subsariana

A maior diversidade genética do HIV-1 tem sido reportada no continente africano, especialmente na África central-ocidental, onde circulam praticamente todos os subtipos conhecidos. O subtipo C demonstra uma elevada prevalência na África meridional, enquanto na África oriental o subtipo A predomina, estando também presentes os subtipos D e C (32). A CRF02_AG tem sido bastante descrita na África ocidental/central-ocidental (32). Os subtipos A e G coexistem nalguns países, como a Nigéria. Também se tem observado um aumento do sub-subtipo F2 e outras formas recombinantes nos Camarões (32,43).

Ásia

O continente asiático tem sido descrito como um grande foco de formas recombinantes de HIV-1 (32). O subtipo A predomina na Rússia, bem como algumas CRFs (32). A China é o país que apresenta um maior número de CRFs em circulação. A prevalência da CRF07_BC, por exemplo, tem vindo a aumentar muito significativamente, revelando-se de grande preocupação (45). Os subtipos B, C e as CRF01_AE, CRF07_BC, CRF55_01B

e CRF08_BC são os predominantes na China (41). No Sudeste Asiático, a CRF01_AE é a mais frequente, enquanto o subtipo C predomina na Índia, com um aumento de URFs reportadas no nordeste do país (32).

1.3.2.1 Diversidade genética e prevalência do HIV-1 em África

África é o continente que apresenta maior diversidade genética do HIV-1, nomeadamente a África central-ocidental, o que está relacionado com o facto de a origem desta pandemia ter sido nesta região (23,32). A epidemia do HIV-1 em África é muito heterogénea, com diferenças regionais significativas na distribuição e prevalência dos múltiplos subtipos (36). Ainda que a África central-ocidental apresente uma maior diversidade genética, os países que a constituem têm taxas de prevalência de infeção mais baixas, quando comparadas com as dos países da África oriental e, culminando, nos países da África meridional, com taxas significativamente mais elevadas (36). No sul de África, o subtipo C é o predominante, enquanto no leste é o subtipo A, ainda que os vírus dos subtipos D e C também circulem (23,36). Nos últimos tempos tem havido um aumento no número de vírus recombinantes em África, sendo este facto também descrito no resto do mundo (23). Desde 2010 que se tem verificado uma redução do número de infeções de cerca de 23%, contudo, a África subsariana continua a ser a região mundial mais afetada por esta pandemia, tendo apresentado, por si só, em 2022, cerca de 25,6 milhões de pessoas a viverem com HIV-1 e 760 mil novas infeções (56).

A África setentrional apresenta dominância do subtipo B, ainda que com baixa prevalência (32,57). Num estudo feito em 2021, no Egito, observou-se, no entanto, uma maior prevalência de subtipos não-B, revelando um grande aumento da presença da CRF02_AG, que passou a ser a forma genética mais comum em circulação, seguida do subtipo B (57).

Na África ocidental observa-se uma epidemia geneticamente heterogénea (36). Como se pode observar na Figura 6a (36), existe uma grande diversidade genética nesta região, onde a CRF02_AG predomina, tanto no oeste, como na África central-ocidental, segundo dados de 2021 (23,32,36). Encontram-se também presentes nesta região, com prevalências significativas, o subtipo G e o sub-subtipo A3, enquanto os restantes subtipos e CRFs apresentam uma prevalência mais baixa (36).

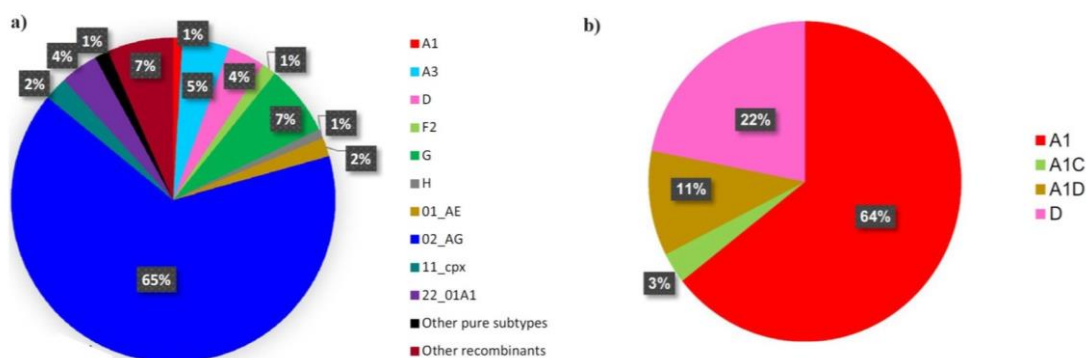


Figura 6. Prevalência dos diferentes subtipos, sub-subtipos e CRFs do HIV-1 na África ocidental (a) e oriental (b) (36).

Na África oriental estão largamente representados os subtipos A1 e D, e seus recombinantes, com o sub-subtipo A1 a predominar (Figura 6b) (36). Num estudo feito em 2019, verificou-se que o subtipo A era prevalente, embora também coexistissem vírus dos subtipos D e C (32). Foram ainda identificados recombinantes inter-subtipo entre os subtipos A1 e C ou A1 e D (36). Uma análise mais recente indica que a circulação das formas puras do subtipo A tem vindo a diminuir, acompanhada de um aparecimento crescente de estirpes recombinantes suas derivadas (36).

Na África central, tal como na África ocidental, circula maioritariamente a CRF02_AG, estando também presentes inúmeras outras formas recombinantes (23). A África central apresenta igualmente uma grande diversidade de formas genéticas virais em circulação, o que está relacionado com a origem desta pandemia ter ocorrido nos Camarões e na RDC. As características replicativas e de recombinação do HIV-1, bem como o fator espaço-temporal, podem ter influenciado a geração da elevada diversidade genética nesta região. Nos Camarões, tem sido relatada a circulação de todos os subtipos do HIV-1 do grupo M, bem como várias CRFs e URFs (30). Ao longo do tempo, a elevada transmissão do HIV-1 entre indivíduos nos Camarões tem levado a um aumento de estirpes recombinantes virais e diminuição dos subtipos puros (30), com exceção do sub-subtipo F2 (23). Como referido anteriormente, Kinshasa, capital da RDC, tem sido descrita como a origem da dispersão desta pandemia (5). Num estudo feito entre 2016-2018, os principais subtipos descritos nesta cidade foram A, G e C (58), com os recombinantes a representar um terço das infeções, identificando-se a CRF47_BF pela primeira vez em África (58). Um aumento significativo dos sub-subtipos A4, A5 e A6 e do sub-subtipo F1

e uma redução significativa do sub-subtipo A1 e do subtipo D, foram observados, durante o período de 2016-2018, comparativamente com as variantes que circularam na cidade de 1983-2008 (58). A identificação de casos de infecção pelo subtipo mais recentemente descrito (subtipo L), parece atualmente restrita à RDC (28).

A África meridional é a região mais afetada no continente africano, apresentando, aproximadamente, 25% de novas infecções em 2020 (23). Nesta região, observa-se uma clara predominância do subtipo C, desde 1990 (23,32). A aparente baixa virulência deste subtipo parece estar associada a um período assintomático mais prolongado, o que pode justificar a dominância e rápida expansão que o subtipo C apresenta nesta região (59).

Como se viu, regiões geográficas distintas apresentam padrões heterogêneos de distribuição de subtipos e formas recombinantes do HIV-1 (23).

Existem múltiplos fatores responsáveis por esta heterogeneidade, bem como pelas taxas de progressão para a doença em toda a África subsaariana (36), incluindo fenômenos de efeito fundador, mas também fatores socioculturais, tais como, as práticas de circuncisão masculina, fatores socioeconômicos, coinfeções, microbiota, fatores genéticos do hospedeiro e virais (36). Outros estudos têm argumentado que os fatores sócio-históricos têm sido os principais determinantes da rápida expansão, por exemplo, do subtipo C na África subsaariana (60). Deste modo, justifica-se a manutenção de uma vigilância molecular das estirpes circulantes em cada momento, considerando o aumento significativo do número de CRFs e a mudança dos padrões de distribuição geográfica dos subtipos e formas recombinantes do HIV-1 ao longo do tempo (23).

1.3.2.2 Diversidade genética e prevalência do HIV-1 em Angola

Angola é um país na África meridional, que faz fronteira com a República do Congo (RC), a RDC, a Zâmbia e a Namíbia. Dados epidemiológicos sugerem que a epidemia da SIDA em Angola terá tido o seu início em 1961, tendo como principal causa de expansão a guerra da independência, ou “Luta Armada de Libertação Nacional” (61).

Em 1986, a taxa de infecções por HIV-1 entre mulheres grávidas em Luanda, era de 0,3% (62). Em 2001, essa taxa tinha aumentado para 8,6% (62). No mesmo ano, um inquérito realizado em Luanda mostrou uma prevalência de 33% entre trabalhadores do sexo e de 3,2% entre militares (62). Em Cabinda, a taxa de infecção em mulheres grávidas aumentou de 6,8%, em 1992, para 8,5% em 1996 (62).

Em 2004 e 2005, foram feitos estudos de prevalência e diversidade genética em Luanda e Cabinda, envolvendo principalmente mulheres grávidas, dadores de sangue seropositivos para HIV e soldados do exército (62,63). No entanto, estes estudos apresentaram uma amostragem pequena, o que não permitiu conclusões estatisticamente significativas.

Em 2008, Angola apresentava, no contexto africano, uma prevalência notavelmente baixa, na população adulta, com cerca de 3,7% (61), tendo sido publicado o primeiro estudo realizado no país de epidemiologia molecular do HIV-1 num grande número de indivíduos, de várias províncias de Angola (61). Os subtipos A e C foram os mais prevalentes, com sequências recombinantes presentes em 47,1% dos indivíduos (61). Ainda em 2008, outro estudo verificou uma predominância do subtipo C, seguido dos subtipos F e H, sendo a CRF02_AG a forma recombinante mais prevalente de entre os recombinantes (64). Em 2009, um estudo incluindo 196 indivíduos não demonstrou grande variação na diversidade e prevalência de subtipos em relação a 2008 (65). Os subtipos A, C e F1 eram prevalentes, junto com várias formas recombinantes, nomeadamente a CRF02_AG (65). Análises da evolução da diversidade genética do HIV-1 em Luanda, entre 2001 e 2009, mostraram uma diminuição significativa na prevalência do subtipo A, que foi substituído pelo subtipo C como o mais dominante. Para além disso, a percentagem de URFs aumentou mais de duas vezes nesse período (66). Outro estudo analisou 101 sequências nucleotídicas virais e constatou que 51% eram de subtipos puros do grupo M do HIV-1, 42% eram recombinantes e 7% eram sequências de subtipo desconhecido (67). Esta informação anuncia já um ligeiro aumento no número de recombinantes inter-subtipo do HIV-1 no período de 2008-2012 (67). Neste período, os subtipos do grupo M mais frequentes em Angola foram, novamente, o subtipo C, com a maior prevalência, seguido do sub-subtipo F1 (67). Para além disto, demonstrou-se, um vez mais, a predominância da CRF02_AG, de entre as formas recombinantes (67). Observaram-se, ainda, sequências recombinantes não conhecidas, que exibiam a mesma estrutura mosaico UH na região PR/RT (67), sendo a região U uma estrutura mosaico, ainda não descrita, resultante da combinação de vários subtipos e CRFs. O sub-subtipo F1 demonstrou ser um dos mais prevalentes em Angola, revelando um aumento de 8% para 23%, entre os anos de 2008-2012 (68). Em 2014, em Luanda, continuou-se a observar uma maior prevalência do subtipo C, seguida dos subtipos F1 e G (66).

Uma revisão mais recente, publicada em 2020, mostrou que os subtipos C, F1, CRF02_AG e um putativo novo recombinante U/H eram os mais prevalentes em Angola. Durante o período de 2010 a 2019, observou-se um aumento no número de estirpes recombinantes. O novo recombinante U/H, mencionado anteriormente, foi identificado como uma nova CRF, já em 2022, a CRF124_cpx, que é derivada dos subtipos A, G, H e da CRF27_cpx (69). Pensa-se que esta forma genética terá começado a dispersar-se em Angola em meados dos anos 70, do século XX, não tendo sido detetada em mais nenhum país da África Central (69).

Em Luanda, foram realizados estudos em mulheres grávidas em 2019 e 2021, observando-se o mesmo padrão de prevalência anteriormente descrito (70). No geral, verificou-se um aumento na complexidade genética do HIV-1 em Angola ao longo do tempo, com destaque para o aumento de estirpes recombinantes (71). O subtipo C foi o mais prevalente no país, seguido do sub-subtipo F1, sendo a CRF02_AG a mais frequente entre as formas recombinantes (71). Os subtipos J e K também estão presentes em Angola, mas a níveis baixos, desde 1993 (72). Em 2022, Angola apresentava 310 mil indivíduos infetados, 15 mil novas infeções e 13 mil mortes associadas a SIDA (73).

Em 2019, foi desenvolvido um estudo sobre a diversidade genética e resistência aos antirretrovirais no HIV-1 em circulação em Benguela, uma província de Angola (74). Observou-se que Benguela apresentava um padrão de prevalências semelhante aos outros estudos, realizados noutras regiões do país. Desta forma, notou-se, novamente, uma predominância do subtipo C, com uma prevalência de cerca de 40%, seguido do sub-subtipo F1, com aproximadamente 18% (74). A CRF02_AG era novamente a forma recombinante predominante, com cerca de 17% de prevalência (74). Os restantes subtipos encontrados (A, B, D, G, H, J e U), apresentavam-se a níveis abaixo de 10% (74). É de notar que o trabalho desenvolvido na presente dissertação, constituiu um seguimento do trabalho apresentado neste estudo.

1.4 Caracterização da integrase, respetiva região codificante e função no ciclo replicativo

Como mencionado anteriormente, o gene *pol* codifica três enzimas, sendo uma delas a IN, que está presente no virião num número estimado aproximado de 120 cópias (11). No

HIV-1, a região codificante da IN tem um tamanho de cerca de 867 nts, estando localizada na região 3' do gene *pol* (18), mais especificamente, entre as posições 4230 e 5096 da sequência padrão HXB2 (75). A IN é uma enzima com cerca de 32 kDa (76,77), composta por 289 resíduos de aminoácidos (11,76,78) e apresenta-se como um multímero, sendo formada, na sua forma ativa, por dímeros ou tetrâmeros (79,80). A sua forma dimérica parece ser suficiente para realizar a reação de processamento a 3', enquanto o tetrâmero é necessário para a integração das duas extremidades do DNA complementar viral no DNA do hospedeiro (80). A IN é uma enzima especializada em recombinação de DNA (18), que pertence a uma grande superfamília de nucleases e recombinases (18,80,81). Relativamente ao seu modo de ação, a IN é uma nucleotidiltransferase (11), ou seja, é uma enzima que faz a transferência de grupos fosfato, neste caso, dos ácidos nucleicos.

Todas as IN retrovirais partilham três domínios comuns, ligados por *linkers* flexíveis: o domínio N-terminal (NTD, do inglês *N-terminal domain*), o domínio do núcleo catalítico (CCD, do inglês *catalytic core domain*) e o domínio C-terminal (CTD, do inglês *C-terminal domain*) (Figura 7) (11,76–78,80–82). O NTD dobra-se numa estrutura compacta de três hélices (com uma estrutura secundária de *helix-turn-helix*), estabilizado pela coordenação de um ião Zn^{2+} (76,77,80,81,83). O CCD adota uma conformação do tipo RNase H, no seu centro ativo, apresentando uma estrutura secundária do tipo α/β , sendo coordenado por dois iões Mg^{2+} (11,76,77,81,83). Por fim, o CTD apresenta uma conformação de barril- β do tipo SH3 (11,76,77,80,81,83).

Após a entrada do vírus na célula hospedeira, o RNA viral é retrotranscrito pela RT para uma molécula complementar de DNA de cadeia dupla, que vem a ser integrada no genoma do hospedeiro, por ação da IN, dando origem ao provírus (Figura 8). Esta integração vai permitir uma transcrição eficiente e uma replicação viral produtiva, assegurando igualmente a segregação dos genes virais para as células filhas, durante a divisão celular (18,84). A integração no cromossoma do hospedeiro, permite ao vírus o estado de latência que o caracteriza (81). A IN é libertada dentro das partículas virais durante a maturação proteolítica das partículas virais imaturas por ação da PR (11,80,84).

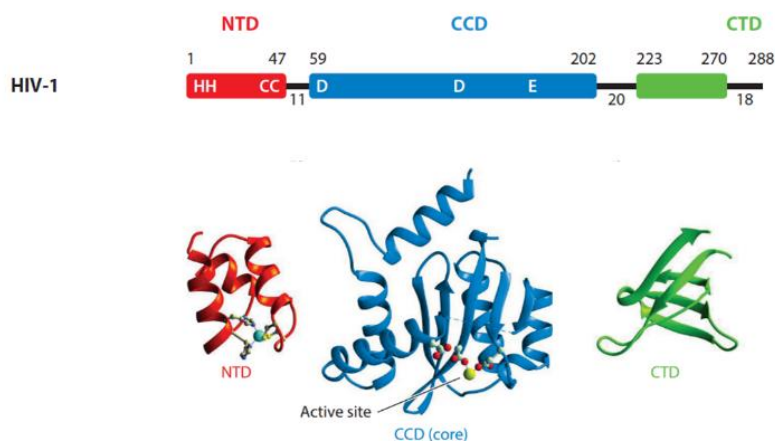


Figura 7. Organização dos domínios da IN. Os números dos aminoácidos (a.a.) delimitam o início e o fim de cada domínio: o domínio N-terminal (NTD; a vermelho), o domínio do núcleo catalítico (CCD; a azul) e o domínio C-terminal (CTD; a verde) (81). Os tamanhos (em a.a.) dos *linkers* entre os domínios estão indicados por baixo das linhas a preto. O íon Zn^{2+} no NTD é apresentado como uma esfera de cor azul-clara; um dos dois íons Mg^{2+} , ligado ao centro ativo do CCD, é representado por uma esfera de cor verde-amarelada (80).

Relativamente à integração, propriamente dita, esta compreende dois passos principais consecutivos: o processamento a 3' e, a transferência de cadeia (Figura 8). A IN reconhece e atua, nas LTRs, sobre as extremidades 3' de ambas as cadeias da molécula linear de cadeia dupla de DNA complementar, que é produzida durante a transcrição reversa (83,84). Na proximidade imediata de cada extremidade 3' do DNA viral está presente um dinucleótido CA, invariável, que marca o local de hidrólise pela IN de um pequeno segmento de dois a três nucleótidos a jusante (11,18,80,82–84). Este processamento ocorre no âmbito da formação de um complexo nucleoproteico altamente estável, composto por um multímetro da IN em interação com as extremidades do DNA viral, denominado complexo de pré-integração (PIC, do inglês *pre-integration complex*) (11,80,85). Para além do DNA viral e da IN, este complexo contém uma série de outros componentes de origem viral, mas também das células hospedeiras (11,80,82,83,85). O PIC permite o transporte do DNA viral processado para o núcleo da célula hospedeira (11,18,81). No núcleo, as extremidades 3'-OH do DNA viral processado são unidas ao DNA da célula hospedeira, durante o evento de transferência de cadeia (11,80–83) (Figura 8). Um dos componentes do PIC, e de extrema importância para o processo de integração, é a proteína LEDGF/p75 (do inglês *Lens Epithelium-derived Growth Factor/p75*), um cofator da célula hospedeira que permite manter o complexo DNA viral-IN ligado ao cromossoma onde irá ocorrer a integração, para que ocorra a transferência de cadeia. (82,85). As extremidades 5' do DNA viral não processado permanecem não

unidas, pelo que a integração resulta, inicialmente, numa forma hemi-integrada do DNA viral, flanqueado por pequenos *gaps* de cadeias simples e pequenas pontas soltas a 5' (11,80,82). O intermediário da integração vai, posteriormente, depender da ação de enzimas da célula hospedeira (até agora pouco caracterizadas) para juntar a extremidade 5' do DNA viral ao cromossoma do hospedeiro (11,18,80,81). Neste mecanismo de reparação do DNA, os *gaps* dão origem a pequenas repetições da sequência de DNA alvo, flanqueando o DNA proviral (82).

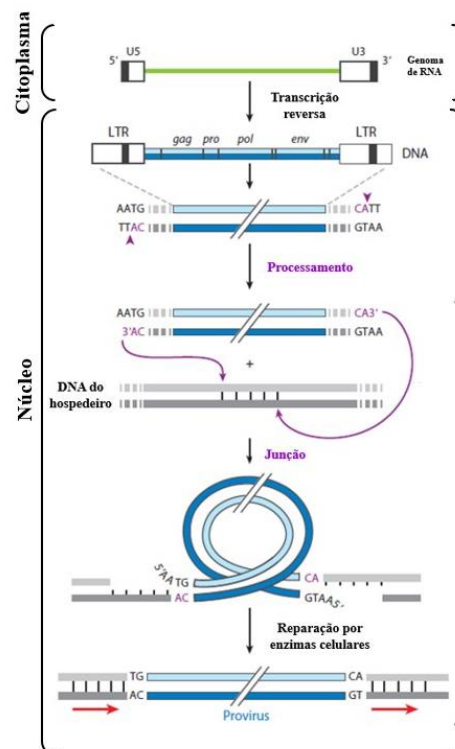


Figura 8. Fases do ciclo replicativo do HIV-1 correspondentes à transcrição reversa e integração do genoma viral no cromossoma do hospedeiro. A verde está representado o genoma viral de RNA, a azul o DNA viral retrotranscrito e a cinzento o genoma do hospedeiro (adaptado de Andrade *et al.*, 2015) (80).

1.5 Antirretrovirais inibidores de transferência de cadeia da integrase (INSTI)

Segundo dados da UNAIDS, no final de dezembro de 2022, 29,8 milhões de pessoas tinham acesso a HAART (cerca de $\frac{3}{4}$ de todas as pessoas infetadas), contra 7,7 milhões em 2010. Neste mesmo ano, 86% de todas as pessoas infetadas com HIV conheciam o seu estado de doença, 89% das quais estavam a aceder ao tratamento e, entre estas, 93% estavam com supressão viral. A HAART permitiu, ao longo dos tempos, a redução da transmissibilidade viral, morbidade e mortalidade associada a SIDA (86).

Uma vez que as células humanas não possuem uma enzima homóloga à IN, esta constituiu um alvo ideal para o desenvolvimento de fármacos antirretrovirais (ARVs). No entanto, ao contrário da primeira geração de ARVs, dirigidos contra a RT e a PR, os inibidores de transferência de cadeia da IN (INSTI, do inglês *integrase strand transfer inhibitors*) tinham de ser desenvolvidos de raiz, dada a ausência de um mecanismo previamente conhecido ou de inibidores naturais conhecidos (11). Contrariamente às outras enzimas virais, que catalisam milhares de reações individuais, o centro ativo da IN realiza um único evento de transferência de cadeia, durante um ciclo replicativo, pelo que um inibidor que interfira com esse mecanismo impede, imediatamente, a única função dessa enzima (11).

Em 2007, foi aprovado pela *Food and Drug Administration* (FDA) e pela *European Medicines Agency* (EMA) o primeiro antirretroviral pertencente à classe dos INSTI, o Raltegravir (RAL) (77). Desde então foram aprovados mais quatro fármacos para esta classe de ARVs: Elvitegravir (EVG), Dolutegravir (DTG), Bictegravir (BIC) e Cabotegravir (CAB) (77,86). Os INSTI são classificados como de primeira ou segunda geração. Os de primeira geração são o RAL e o EVG, enquanto o DTG, o BIC e o CAB pertencem ao grupo dos de segunda geração. O CAB é um fármaco de ação prolongada, isto é, apresenta um longo tempo de semivida, não sendo necessária a administração de forma tão regular como os outros INSTI, o que permite reduzir eventos de fraca adesão por parte dos indivíduos, limitando assim a potencialidade de desenvolver resistências.

Relativamente às restantes classes de ARVs, os INSTI têm como vantagens, o facto de a grande maioria possuir uma alta barreira genética contra resistências, isto é, o vírus tem menos propensão em desenvolver mutações de resistência relativamente aos fármacos em uso pelo indivíduo, terem uma elevada tolerabilidade, interações fármaco-fármaco limitadas e uma excelente taxa de supressão virológica (87), o que leva a uma melhor adesão por parte dos indivíduos e melhores resultados no tratamento.

1.5.1 Esquemas terapêuticos

A HAART é um método padrão de tratamento dos indivíduos infetados com HIV e consiste numa combinação de três ou mais ARVs que atuam em diferentes alvos. A

terapia antirretroviral visa reduzir a carga viral a níveis indetetáveis, geralmente menor do que 50 cópias de genoma viral por mL de plasma, manter o número de linfócitos T CD4+ acima de 200 células por mm³ de sangue (88,89) e impedir a progressão da doença. As recomendações atuais são que a HAART deve ser iniciada no prazo de sete dias após um diagnóstico confirmado e carga viral detetável, independentemente da contagem de linfócitos T CD4+ ou da presença de sintomas clínicos.

Na generalidade, a HAART compreende combinações de três fármacos distintos, especificamente, dois NRTI (do inglês *nucleoside reverse transcriptase inhibitors*) com um NNRTI (do inglês *non-nucleoside reverse transcriptase inhibitors*) ou PI (do inglês *protease inhibitors*) (2 NRTI + NNRTI ou PI) (90). Em setembro de 2019, o Instituto Nacional de Luta Contra a Sida (INLS) de Angola determinou a transição terapêutica para regimes de terapia antirretroviral otimizados, baseados no DTG. Como esquema de primeira linha, as diretrizes atuais para a HAART em Angola recomendam o DTG em dose fixa combinada com Lamiduvina (3TC) e Tenofovir (TDF), para todos os adultos, adolescentes e crianças (com > de 12 anos de idade e peso > 30 kg), bem como para a profilaxia de pré-exposição (PrEP, do inglês *pre-exposure prophylaxis*) com as mesmas características (91). A Organização Mundial da Saúde (OMS), com o apoio do “Plano de Emergência do Presidente dos EUA para o Alívio da SIDA” (PEPFAR, do inglês *President's Emergency Plan for AIDS Relief*), através dos Centros de Controlo e Prevenção de Doenças (CDC, do inglês *Centers for Disease Control and Prevention*), colaboraram com o INLS de Angola para a atualização das diretrizes nacionais sobre a terapia antirretroviral (92,93). O objetivo seria o da utilização de regimes baseados no DTG para iniciar o tratamento dos indivíduos recém-diagnosticados com HIV ou na implementação de terapia de resgate para os indivíduos submetidos a terapia antirretroviral de longa data, com indícios de falência terapêutica, tornando o DTG num medicamento potencial em regimes de salvamento (77,93). Na fase inicial, esta iniciativa foi implementada em três províncias de Angola, nomeadamente, Benguela, Cunene e Huambo (92). Esta alteração no regime terapêutico antirretroviral fora recomendada pela OMS, desde 2019, como tratamento preferencial de primeira linha, no entanto, introduzida em Angola apenas em 2021 (91,93).

1.6 Mecanismo de inibição dos INSTI

Na base da identificação de compostos eficazes inibidores da IN encontra-se a constatação de que a integração do DNA complementar viral no DNA cromossômico do hospedeiro, por ação da IN, é um mecanismo em duas fases (o processamento a 3' e a transferência de cadeia) dependente da formação de um complexo nucleoproteico ativo para a integração, denominado intassoma (94).

O estudo dos intassomas do PFV (do inglês, *prototype foamy virus*) com RAL ou EVG ligados revelou algumas características importantes do processo de integração (94) e também a razão pela qual os INSTI se ligam e inibem os intassomas, mas não a IN livre, embora da sua ligação não resultem grandes alterações conformacionais no primeiro. Entre outros fatores, a longa semivida dos INSTI, torna-os compostos altamente ativos para a inibição do vírus (77,86,94).

Para combater o rápido aparecimento de variantes do HIV resistentes aos INSTI de primeira geração, foram desenvolvidos novos compostos. A mudança fundamental nos INSTI de segunda geração é o facto de apresentarem estruturas químicas de maiores dimensões (anéis tricíclicos, em vez de bicíclicos) responsáveis pela quelação de metais (77,86,94). A caracterização da estrutura do intassoma do PFV ligado ao DTG mostrou que este ocupa um espaço semelhante ao do RAL ou do EVG na região do centro ativo da IN, mas coordena os iões Mg^{2+} em ângulos ligeiramente diferentes e, presumivelmente, mais otimizados (94). Em resultado desta estrutura alargada, os INSTI de segunda geração estabelecem contactos adicionais com resíduos da IN e, conseqüentemente, possuem uma alta barreira genética contra resistências (11). Foi observado que o DTG é mais potente contra uma estirpe resistente do HIV-1 do que um fármaco derivado sem o terceiro anel (INSTI de primeira geração), enquanto ambos os compostos inibem a estirpe selvagem de forma semelhante (94,95). O aumento desta potência é atribuído a uma semivida dissociativa mais longa, uma vez que o análogo sem o terceiro anel se dissocia mais rapidamente em comparação com o DTG ou o BIC (94,95).

1.7 Tipos de mutações de resistência aos antirretrovirais no HIV-1

Com o aparecimento da HAART, o HIV-1 começou a adquirir mecanismos de resistência (HIVDR, do inglês *HIV drug resistance*) associada a estes fármacos, o que representa

uma ameaça potencial para o seu sucesso a longo prazo e está a surgir como uma ameaça, também, para se atingir a meta proposta de pôr fim à pandemia da SIDA até 2030 (96,97). Esta meta baseia-se na Estratégia Global contra a SIDA para 2025, proposta pela UNAIDS, que inclui o algoritmo 95-95-95: pretende-se que um mínimo de 95% das pessoas infetadas com HIV conheçam o seu estado de infeção, com 95% destas a receber HAART e 95% das pessoas sob terapia a atingir uma supressão virológica duradoura (98).

1.7.1 Mutações de resistência principais e acessórias

Normalmente, as infeções por HIV-1 começam com apenas uma variante viral, responsável pela infeção. Contudo, após algumas semanas, começam a surgir, no indivíduo infetado, inúmeras variantes com origem na primeira, designando-se no seu conjunto, tal como referido anteriormente, de *quasispecies* (99). Nessas *quasispecies* vão surgindo variantes com diferenças nucleotídicas nos respetivos genomas, que levam à aquisição, ou não, de resistências a determinados ARVs. A seleção de variantes resistentes aos fármacos depende da manutenção de uma replicação viral contínua, durante uma terapia de supressão incompleta, da facilidade de aquisição de determinadas mutações de resistência (DRM, do inglês *drug resistance mutations*) e do seu efeito na suscetibilidade aos fármacos e replicação viral (99). As variantes virais com DRM surgem sempre na *quasispecies* em indivíduos não tratados, embora não permaneçam normalmente na população devido ao seu baixo *fitness* replicativo, comparativamente às variantes suscetíveis, na ausência da pressão seletiva do fármaco (99). Na presença da pressão seletiva dos fármacos, estas mutações tornam-se, no entanto, mais vantajosas para os vírus que as possuem. Desta forma, todas as DRM clinicamente significativas, surgem como resultado da pressão imposta pela terapia (99).

Para alguns ARVs, uma DRM é suficiente para reduzir a suscetibilidade do vírus, enquanto para outros são necessárias várias DRM. O número de DRM, bem como o seu efeito no *fitness* replicativo do vírus, contribuem para a barreira genética dos ARVs à resistência (99). Uma supressão viral incompleta favorece o aparecimento de resistência aos fármacos, aumentando a probabilidade de falência virológica (FV), isto é, o retorno

de um indivíduo suprimido pela HAART à carga viral acima do limiar da detecção, o que pode comprometer o tratamento subsequente (100).

As DRM são denominadas principais (*major*) quando concorrem diretamente para um decréscimo da suscetibilidade a um determinado antirretroviral (99) e acessórias (*minor*) quando compensam apenas a perda de *fitness* de variantes virais que apresentam resistências principais e/ou contribuem adicionalmente para a diminuição da suscetibilidade (99).

1.7.1.1 Resistência aos antirretrovirais transmitida (TDR)

A resistência aos ARVs transmitida (TDR, do inglês *transmitted drug resistance*) é um tipo de resistência que ocorre quando uma nova infecção é causada por um vírus portador de uma ou mais mutações de resistência, mas também pode ocorrer com uma exposição subsequente, referida como superinfecção (99–101). A TDR pode surgir de duas formas distintas: a partir de indivíduos com baixa adesão terapêutica, tendo desenvolvido resistência aos ARVs adquirida (ADR, do inglês *acquired drug resistance*) ou a partir de um indivíduo submetido pela primeira vez à HAART, mas que já possui uma TDR (99). Neste último, este tipo de resistência designa-se, também, de resistência de pré-tratamento aos ARVs (PDR, do inglês *pre-treatment drug resistance*) (99,101). Quando uma TDR é detetada num indivíduo que está a receber pela primeira vez HAART, é necessária uma seleção criteriosa inicial do regime terapêutico antirretroviral (101). Relativamente à PrEP, a resistência associada tem risco reduzido de surgir, pois um indivíduo que esteja submetido a PrEP tem menor risco de ser infetado (99), mas aqueles que acabam infetados, têm um risco teórico muito maior de desenvolver resistências (99).

A TDR é um problema emergente em países de médio-baixo rendimento. A introdução recente da HAART nessas regiões teve impacto positivo na saúde, mas os dados de vigilância epidemiológica são limitados, em comparação com regiões de alto rendimento (101). É comum encontrar uma menor prevalência de TDR em regiões menos desenvolvidas, devido à introdução tardia da HAART (101). Contudo após o ano de 2009, onde ocorreu um aumento da implementação dos programas de HAART, observou-se um aumento da prevalência de TDR, nomeadamente PDR, nos países da África subsariana, incluindo Angola, entre os períodos de 2009-2013 a 2014-2019 (99,102). Em Angola,

estudos indicam também um aumento da frequência de TDR entre 2001 (65) e 2008/2010 (67), de 1,6% para 16,3%.

Dada a prevalência da TDR, é altamente recomendada a realização de testes genotípicos de resistência, antes do início da implementação da terapia antirretroviral ou, até mesmo, logo no momento do diagnóstico (101). A TDR é um problema de saúde pública, menos significativo nos países de alto rendimento, uma vez que os testes genotípicos de resistência pré-terapia estão normalmente disponíveis (102), o que não acontece na maioria dos países de médio-baixo rendimento.

1.8 Caracterização das mutações de resistência aos INSTI

Os INSTI ligam-se ao núcleo catalítico da IN e inibem o processo de transferência de cadeia (86). O núcleo catalítico contém os resíduos Asp64, Asp116 e Glu152 no centro ativo, que coordenam os cátions divalentes Mg^{2+} , necessários para o processamento a 3' e o mecanismo de transferência de cadeia (90). A utilização clínica de RAL e EVG levou à seleção, antecipada, de mutações associadas a resistência a estes fármacos. Foi demonstrado que a maioria dos indivíduos que apresentam FV utilizando um regime de tratamento contendo RAL se encontram infetados com vírus com as mutações principais F121Y, Y143R/H/C, Q148H/K/R, N155H e R263K na região codificante da IN (103) (Figura 9). Para além destas, mutações acessórias, como L74M, E92Q, T97A, E138A/K ou G140A/S (103), podem estar presentes na ocorrência de FV com regimes contendo RAL, mas contribuindo para níveis mais baixos de resistência quando não combinadas com as mutações principais (77,103). Estas mutações acessórias parecem reduzir a capacidade de replicação viral, ainda que a níveis menores. No que diz respeito ao EVG, estudos reportaram as mutações principais T66I, E92Q, F121Y, S147G, Q148H/K/R, N155H e R263K, bem como mutações acessórias, tais como T66A/K, E92G e T97A, durante o tratamento com recurso ao EVG (103) (Figura 9). É de salientar que algumas destas mutações exibem padrões de resistência comparáveis aos observados em regimes de tratamento utilizando o RAL, suportando, assim, resistência cruzada (77).

Algumas das vias de resistência mais comuns envolvem os resíduos G140, Y143, Q148 e N155 e R263 da IN. Utilizando o intassoma do PFV como um sistema modelo, foi possível o estudo de alguns mecanismos de resistência. O RAL faz uma interação com o

resíduo Y143 da IN (94). Assim, uma mutação nessa mesma posição resulta na perda da interação do resíduo aminoacídico com o fármaco (94).

As mutações que causam resistência aos INSTI reduzem, em geral, a capacidade da IN de desempenhar as suas funções normais. Embora muitas destas mutações afetem a ligação da IN aos INSTI, algumas, como a G140S, são mutações acessórias que, como descrito anteriormente, permitem compensar a perda de *fitness* das variantes virais que possuem mutações principais e/ou contribuir para a diminuição da suscetibilidade aos ARVs (104). A mutação G140S compensa a perda de atividade da IN das variantes virais que possuem mutações na posição 148: Q148H/K/R (105). Estudos realizados com estruturas cristalinas de intassomas do PFV demonstraram que a presença de mutações nos resíduos G140, Q148 e N155 induzem rearranjos ligeiros, mas significativos, na região do centro ativo da IN, que têm como consequência uma redução na afinidade de ligação do INSTI e/ou um aumento da barreira energética para a ligação de alta afinidade entre os INSTI e a IN (94).

Não obstante se ter verificado uma notável melhoria em termos de potência frente às variantes virais geradas como resposta aos fármacos de primeira geração, os compostos de segunda geração não ficaram isentos de enfrentar o fenómeno de escape viral. A prova mais clara deste facto é o fracasso da monoterapia com DTG (106). Alterações nas posições 50, 51, 66, 74, 92, 97, 118, 138, 140/148, 143, 147, 149, 151, 153, 155, 157, 230 e/ou 263 são comumente observadas como resposta à terapia baseada em DTG, apesar da alta barreira genética para resistências apresentada pelo fármaco, e foram minuciosamente documentadas num artigo de revisão de 2019 (107).

Conforme descrito na literatura (77), as mutações principais correlacionadas com a resistência ao DTG compreendem G118R, G140R, Q148H/K/R e R263K, enquanto as mutações acessórias incluem E138A/K/T, G140A/C/S, S153F/Y e N155H (103) (Figura 9). Como em geral, de entre os fatores de risco associados à resistência ao DTG destacam-se a má adesão ao tratamento, baixa contagem de linfócitos T CD4+ ou elevada carga viral no momento da instauração dos regimes terapêuticos contendo DTG, bem como interações farmacológicas decorrentes da coadministração de medicamentos (77).

A mutação principal de resistência ao BIC é a R263K, sendo as mutações acessórias G118R, E138A/K/T, G140A/C/R/S, Q148H/K/R e S153F/Y (103) (Figura 9). Importa

salientar que as mutações Q148H e R263K foram documentadas como causadoras de resistência cruzada entre INSTI de primeira e segunda geração (77).

Os fármacos de ação prolongada, como o CAB, ajudam a contornar casos de fraca adesão à terapia, limitando assim o potencial desenvolvimento de resistência farmacológica. No entanto, já foram descritas na literatura mutações que conferem resistência ao CAB. Foram registradas como mutações principais G118R, G140R Q148H/K/R, N155H e R263K, e como mutações acessórias T66K, T97A, E138A/K/T, G140A/C/S e S153F/Y (103) (Figura 9). As mutações principais de resistência ao CAB são semelhantes às que causam resistência ao DTG, o que levanta preocupações sobre a sua utilidade aquando de falhas na HAART baseada em DTG (77).

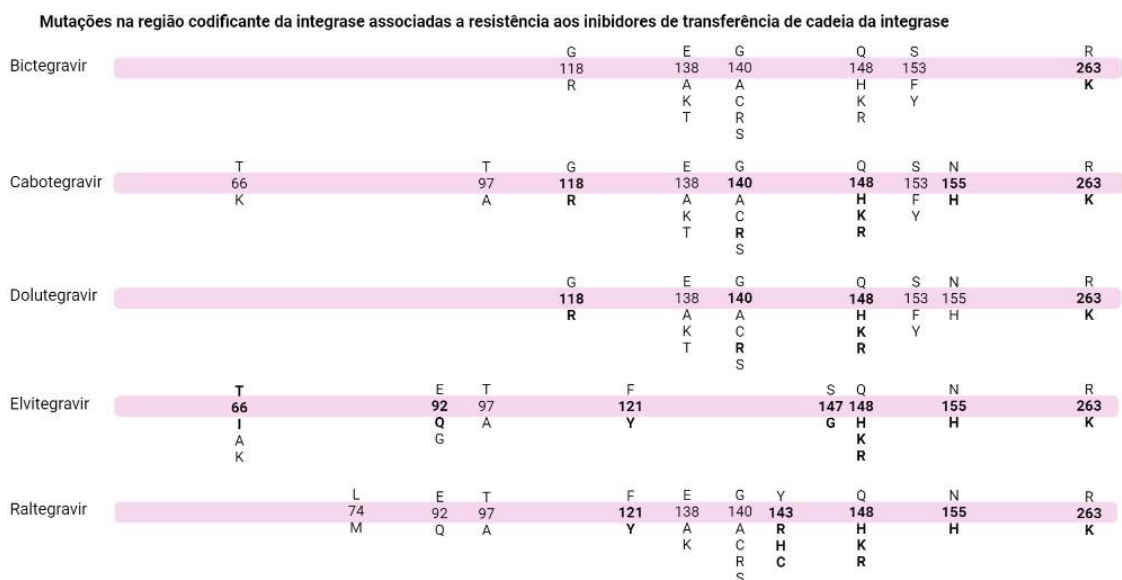


Figura 9. Mutações na região codificante da IN associadas a resistência aos INSTI (adaptado de Wensing *et al.*, 2022) (105).

A IAS-USA (do inglês *International Antiviral Society-USA*) fornece periodicamente a lista de mutações associadas a diminuição de suscetibilidade aos diferentes ARVs. Na Figura 9, encontram-se representadas as mutações principais referentes aos INSTI, assinaladas a negrito, bem como as mutações acessórias (103,108). É importante frisar que existem várias listagens distintas, fundamentadas em critérios nem sempre uniformes, e, conseqüentemente, não apresentam total concordância em todos os casos. Existem também mutações principais designadas de SDRM (do inglês *Surveillance drug resistance mutations*), que podem ser, ainda, classificadas como relevantes para a vigilância epidemiológica da resistência primária (109). Estas encontram-se

representadas na Tabela 2. As SDRM servem para comparar com precisão as taxas de resistência transmitida aos fármacos, entre regiões geográficas e períodos de tempo distintos (110).

Tabela 2. Lista de mutações de resistência transmitida aos INSTI, utilizadas na vigilância epidemiológica da resistência primária (baseado na *HIV Drug Resistance Database, Stanford University, 2020*) (109).

Posição aminoacídica	Mutação
T66	A, I, K
E92	G, Q
G118	R
F121	Y
E138	A, K, T
G140	A, C, S
Y143	C, H, R, S
S147	G
Q148	H, R, K
N155	H
S230	R
R263	K

1.9 Monitorização das resistências e sua importância

A OMS, em 2016, recomendou que todos os indivíduos infetados com HIV devem iniciar imediatamente os regimes de HAART, assim que o diagnóstico é realizado, e também que os indivíduos com risco elevado de infeção pelo HIV devem ter acesso à PrEP (99). Desde 2016, as diretrizes da OMS para a HAART incluem opções terapêuticas alternativas, com melhor tolerabilidade, maior eficácia e menores taxas de descontinuação do tratamento (111). Uma vez que a implementação terapêutica está frequentemente associada ao aparecimento de resistência aos fármacos, torna-se mandatória a sua monitorização, até para que antes de se dar início ao tratamento se tenha o conhecimento de quais os ARVs mais adequados a administrar a cada indivíduo. Contudo, a implementação desta monitorização tem custos e constrangimentos técnicos associados, pelo que, muitas vezes, não é possível um acompanhamento como desejado, principalmente nos países de baixo rendimento.

Relativamente ao continente africano, a OMS, através do “Plano de Ação Regional Africano sobre a Resistência Medicamentosa contra o HIV, 2019-2023”, recomendou a realização de inquéritos regulares a nível nacional, de forma a monitorizar o surgimento e transmissão de DRM representativas (112). É também necessário a produção e o fornecimento de dados fidedignos para fundamentar a seleção dos regimes ideais de HAART, por forma a minimizar o surgimento e a transmissão das DRM (112). Finalmente, realça-se a importância da implementação atempada de métodos de monitorização através da realização de testes genotípicos ou fenotípicos, de forma contínua e o mais cedo possível, antes da implementação da HAART (112). Para tal, a OMS fornece orientação padronizada, conjuntos de ferramentas operacionais e uma base de dados de apoio aos laboratórios de genotipagem de mutações de resistência dos países de baixo rendimento (112), na tentativa de assegurar a garantia da qualidade dos dados obtidos. A OMS incentiva os países a monitorizarem as DRM em indivíduos que estejam a iniciar o tratamento com ARVs, de modo a prevenir a PDR, recomendando ainda que quando o nível de PDR de um país atinge 10%, o esquema terapêutico de primeira linha administrado seja urgentemente substituído por um esquema contendo DTG (111).

1.10 Objetivo do trabalho

Este trabalho insere-se num estudo transversal realizado no Hospital Geral de Benguela (HGB), em Angola, que engloba várias infeções sexualmente transmissíveis (sífilis, infeção pelo HIV e hepatite B) e inclui 226 amostras de sangue total de indivíduos seropositivos para o HIV, testados por métodos imunocromatográficos.

Dada a ausência de dados acerca das mutações que conferem resistência aos INSTI em Benguela (Angola), devido à introdução tardia desses ARVs na região (apenas em 2021) (91), conforme recomendação da OMS em 2019 (93), surge a necessidade de enfrentar esse obstáculo. Com esse propósito, foram delineados os seguintes objetivos:

- (1) Investigar a presença de polimorfismos genéticos naturais na região codificante da IN, estimar a sua prevalência e correlacioná-los com os genótipos virais identificados;
- (2) Estabelecer relações entre os polimorfismos mencionados e os fenómenos de resistência natural aos INSTI.

Estes objetivos visam preencher a falta de conhecimento científico nesta região e contribuir para uma compreensão mais abrangente da resistência aos fármacos na população estudada, permitindo uma abordagem mais eficaz e personalizada no tratamento do HIV-1 em Benguela.

2. Material e métodos

2.1 Caracterização geral da amostra populacional em estudo

O trabalho em curso insere-se num estudo transversal de maior amplitude, realizado em 2016/2017, cujo objetivo consistiu na caracterização de algumas infeções de transmissão sexual, em indivíduos internados, utentes de consultas externas ou utentes de urgência do HGB. Este estudo recebeu um parecer favorável para a sua realização da Comissão Científica e de Ética do HGB. Neste trabalho, foram incluídas 226 amostras de sangue de indivíduos infetados com HIV-1, colhidas em cartões FTA[®] (GE Healthcare, Reino Unido). Foi aplicado um questionário estruturado, pré-codificado, de preenchimento voluntário, a todos os indivíduos que assinaram um termo de consentimento livre e esclarecido, com o objetivo de obter informação sobre algumas características demográficas e epidemiológicas, estando a informação recolhida resumida no Anexo 2. Os indivíduos incluídos neste estudo apresentavam idades compreendidas entre 0 e 95 anos, sendo 42% do sexo masculino e 58% do sexo feminino. Nenhum dos indivíduos desta amostra tinha sido anteriormente submetido a terapia antirretroviral. O diagnóstico da infeção por HIV-1 foi realizado na mesma data da colheita das amostras, isto é, entre agosto de 2016 e janeiro de 2017.

2.1.1 Colheita das amostras e rastreio da infeção pelo HIV-1

Após recolha do consentimento informado dos indivíduos, as respetivas amostras de sangue foram colhidas por punção venosa e testadas de imediato para a presença de anticorpos anti-HIV. Para tal, recorreu-se a dois testes rápidos imunocromatográficos. Como primeiro teste rápido, utilizou-se um teste de terceira geração, *VIH 1/2* (Healgen Scientific, EUA), para pesquisa de anticorpos anti-HIV-1 e -2 em soro, plasma ou sangue total, não sendo possível distinguir entre os dois tipos virais (113). O segundo teste, *Anti-HIV 1/2 Test* (Türklab, Turquia), permitia a deteção de anticorpos anti-HIV-1 e -2, incluindo contra as estirpes virais do grupo O do HIV-1, utilizando os antígenos virais gp120, gp41, gp36 e p24 (114). Seguidamente, para as amostras positivas, em, pelo menos, um dos testes realizados, foram armazenadas algumas gotas de sangue total em cartões *Whatman Indicating FTA[®] Elute Micro Cards* (GE Healthcare, Reino Unido), seguindo as especificações do fabricante. Os cartões deste tipo possibilitam a preservação das amostras à temperatura ambiente (TA) sem sofrerem degradação considerável. Após

a lise das células, devido ao contacto com os cartões, as proteínas sofrem desnaturação, enquanto os ácidos nucleicos são imediatamente estabilizados, garantindo a sua proteção e conservação (115). É relevante mencionar que, devido a questões logísticas, algumas amostras foram armazenadas num tipo diferente de cartões, nomeadamente, *Whatman Non-Indicating FTA[®] Micro Cards* (GE Healthcare, Reino Unido), que possuem uma tecnologia bastante similar à dos cartões mencionados anteriormente. As amostras de sangue positivas foram sujeitas a centrifugação sob condições padrão para separação do soro, que foi armazenado a -20 °C. Além disso, já no Instituto de Higiene e Medicina Tropical (IHMT), foi realizado um terceiro teste imunocromatográfico rápido para confirmação dos resultados obtidos em Angola. Esse teste, também de terceira geração, designa-se *Hexagon HIV-1/2* (Human Diagnostics, Alemanha) e deteta anticorpos anti-HIV-1 e -2 em soro, plasma ou sangue total (116,117). Permite a diferenciação entre anticorpos produzidos contra cada um dos tipos virais, por meio de duas linhas de teste individualizadas, além de detetar anticorpos anti-HIV-1 do grupo O. Todos os testes foram conduzidos em conformidade com as diretrizes dos respetivos fabricantes. Esta etapa do estudo transversal foi realizada, quer em Benguela, quer em Lisboa, pela Dr.^a Lesya Yefimenko.

2.2 Preparação de DNA proviral do HIV-1 a partir de amostras de sangue seco imobilizadas em cartões FTA[®]

De forma a preparar o DNA proviral das amostras de sangue seco imobilizadas em cartões FTA[®] (GE Healthcare, Reino Unido) para o passo seguinte de amplificação por PCR, procedeu-se à realização de um protocolo otimizado anteriormente no IHMT (Anexo 3), com pequenas alterações.

Num primeiro passo, procedeu-se à limpeza, com etanol a 70% (v/v), do tapete de corte e da extremidade do *puncher* (pressionando o êmbolo), tendo sido feito, de seguida, seis furos em papel *Whatman*. Posteriormente, pressionando o *puncher* sobre a amostra de sangue em cartão FTA[®] (GE Healthcare, Reino Unido), preferivelmente no centro da mancha, retirou-se um disco de 3 mm de diâmetro para um microtubo *eppendorf* estéril de 1,5 mL. Entre cada recolha de discos de amostras diferentes procedeu-se à limpeza,

com etanol a 70% (v/v), da extremidade do *puncher*, seguida de seis furos em papel *Whatman*, para evitar contaminações.

Com o objetivo de remover contaminantes e inibidores da reação de PCR (principalmente a hemoglobina, um inibidor da enzima *Supreme Taq polimerase*, que foi posteriormente utilizada na reação de PCR), adicionou-se 200 µL de *FTA purification reagent* (Whatman, EUA) a cada tubo e incubou-se durante 5 min à TA, com agitação ocasional. De seguida, removeu-se o meio (e a espuma, caso se tenha formado) com auxílio de uma micropipeta com pontas com filtro *MultiGuard Barrier Tips*[®] (*Sorenson, BioScience, Inc.*, EUA). Repetiu-se o passo de lavagem com *FTA purification reagent* (Whatman, EUA) mais três vezes, sendo que o último foi realizado com um tempo de incubação de 10 min.

Para proceder à total remoção do reagente anteriormente adicionado, adicionou-se 200 µL de tampão TE (Anexo 7) a cada tubo e incubou-se à temperatura ambiente, durante 5 min, agitando os tubos, ocasionalmente. Removeu-se o meio (e a espuma) com o auxílio da micropipeta com pontas com filtro. Repetiu-se este passo de lavagem mais uma vez.

Por fim, colocaram-se os tubos abertos num evaporador (*SpeedVac, Eppendorf*[®] *centrifugal vacuum concentrator, Model 5301*), no programa para soluções aquosas, a 30 °C, durante 15 min, para garantir a secagem total dos discos. Estes foram conservados a -20°C até posterior utilização para a reação de PCR.

2.3 Amplificação de DNA proviral do HIV-1 (região codificante da integrase) por *nested* PCR

2.3.1 Caracterização dos *primers* utilizados

Tabela 3. Descrição dos *primers* externos e internos da reação de *nested* PCR. Posições nucleotídicas retiradas da base de dados *Los Alamos National Laboratory*, ferramenta *Sequence locator*. <https://www.hiv.lanl.gov/content/sequence/LOCATE/locate.html>

Reação <i>nested</i> PCR	<i>Primers</i>	Sequência nucleotídica	Posição nucleotídica no gene <i>pol</i> , relativamente ao início da região codificante, em HXB2	Posição nucleotídica na região codificante da integrase, relativamente ao início da região codificante, em HXB2
1. ^a reação	INPS1	5'- TAG TAG CCA GCT GTG ATA AAT GTC-3'	2252 → 2275	107 → 130
	INPR8	5'-TTC CAT GTT CTA ATC CTC ATC CTG-3'	2998 → 3012	853 → 867
2. ^a reação	INPS3	5'- GAA GCC ATG CAT GGA CAA G- 3'	2287→ 2305	142→160
	INPR9	5'-ATC CTC ATC CTG TCT ACT TGC C -3'	2988→3009	843→864

Para a realização do protocolo de amplificação por *nested* PCR da região codificante da integrase (IN), foram selecionados quatro *primers*: dois externos (para a 1.^a reação), INPS1 e INPR8, e dois internos (para a 2.^a reação), INPS3 e INPR9, cujas características estão representadas na Tabela 3. O par de *primers* externos delimita um fragmento de 761 pb, enquanto o par de *primers* internos abrange um fragmento de 723 pb, que corresponde a 84% da sequência codificante da IN, permitindo avaliar todas as mutações de resistência (principais e acessórias) associadas aos inibidores da integrase (INSTI). A respetiva posição de ligação dos *primers* na região codificante da IN, na estirpe do HIV-1 HXB2, está representada na Figura 10.

Estes foram descritos pela *Agence Nationale de Recherche sur le Sida et les Hépatites Virales* (ANRS, França) para a amplificação da sequência codificante da IN (118).

Os *primers* foram sintetizados pela empresa Stab Vida (Caparica, Portugal), sendo fornecidos liofilizados, pelo que se procedeu inicialmente à sua hidratação, com água livre de nucleases (*Promega Corporation*, EUA), para uma concentração final de 100 μM .

A partir dessas soluções, fizeram-se soluções stock para cada *primer*, para uma concentração final de 10 μM (diluição de 1:10), utilizando água ultrapura *Milli-Q[®]* (*Reagent Grade Ultrapure Systems*, Merck Millipore, EUA), a partir das quais se prepararam alíquotas de trabalho de 20 μL de volume, para evitar ciclos de congelação/descongelação sucessivos. As soluções de trabalho foram mantidas a -20°C até utilização.

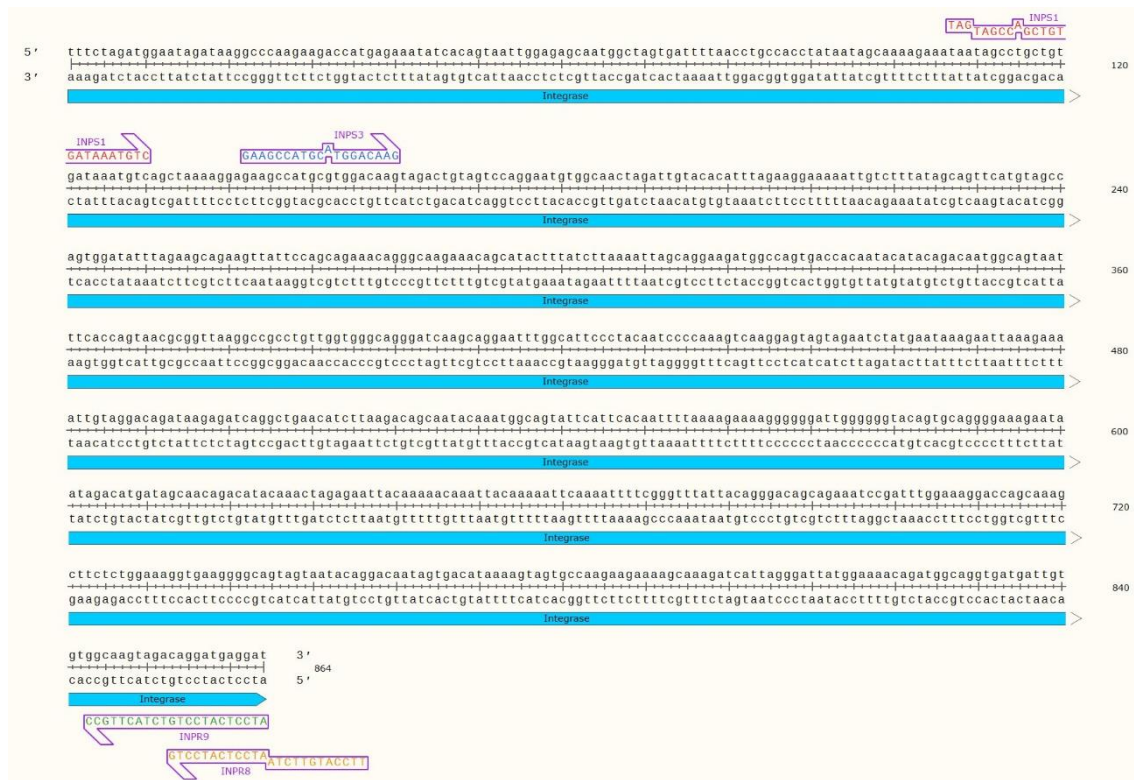


Figura 10. Posição relativa dos *primers* diretos (INPS1 e INPS3) e reversos (INPR9 e INPR8) na sequência nucleotídica da região codificante da IN (retirado de *Snagene*. <https://www.snagene.com/>).

2.3.2 Planificação do protocolo de amplificação

Para a amplificação do DNA proviral do HIV-1, foi planificada a realização de duas reações de polimerização em cadeia sequenciais (*nested PCR*), com recurso à mistura reacional comercial *Supreme NZYTaQ II 2× Green Master Mix* (NZYTech, Lisboa, Portugal). Esta mistura reacional possui a enzima *Supreme NZYTaQ II*, derivada da *Taq polimerase*, que apresenta uma processividade elevada e permite um aumento da sensibilidade da reação de PCR. Esta enzima apresenta, ainda, atividade *hot start*, ou seja, encontra-se inativa à temperatura ambiente, apenas restaurando a sua atividade na primeira etapa de incubação de 5 minutos a 95 °C da PCR evitando-se, desta forma, a extensão por hibridação inespecífica de *primers* ou formação de dímeros de *primers*, fornecendo uma amplificação altamente específica. Utilizaram-se os *primers* externos para a 1.^a reação e os *primers* internos para a 2.^a reação, que se encontram descritos na Tabela 3.

De acordo com as especificações do fabricante, foi preparada, para a 1.^a reação de PCR, uma mistura reacional, para um volume final de 25 µL, contendo 12,5 µL de *Supreme NZYTaQ II 2× Green Master Mix* (concentração da enzima 0,2 U/µL; MgCl₂, para concentração final de 2,5 mM; 200 µM de cada dNTP; tampão de reação e aditivos a concentrações otimizadas), 1 µL de cada *primer* externo (concentração final 0,4 pmol/µL) e perfazendo com 10,5 µL de água ultrapura *Milli-Q*[®] (*Reagent Grade Ultrapure Systems*, Merck Millipore, EUA), em microtubos *ependorf* de 0,5 mL. A cada tubo é adicionado um disco preparado a partir de cartões FTA[®] (GE Healthcare, Reino Unido), como descrito anteriormente, sendo desprezado o respetivo volume ocupado na mistura reacional. Para a realização da 2.^a reação de PCR, foi preparada uma nova mistura reacional, com os *primers* internos (concentração final 0,4 pmol/µL), também para um volume final de 25 µL, tendo em conta a transferência de 2 µL de produto da 1.^a reação. As reações foram realizadas em microtubos de paredes finas de 0,5 mL (*Axygen*, EUA).

A preparação dos *primers* e das misturas reacionais foi realizada num ambiente estéril, com recurso a uma câmara de fluxo laminar (*Laminar Flow Cabinet-class 100*, Flow Laboratories, EUA), enquanto a adição das amostras (DNA matriz) foi realizada num laboratório distinto, de forma a prevenir contaminações. Em todos os passos, foram utilizadas pontas de micropipeta com filtro.

A escolha deste tipo de reação de amplificação (*nested* PCR) permite aumentar a sensibilidade da técnica, considerando a utilização, à partida, de amostras com pequenas quantidades de DNA matriz, conservado, durante um longo período, à TA.

2.3.2.1 Reações de PCR

As condições e o protocolo de amplificação por *nested* PCR foram retirados do documento sobre os procedimentos para sequenciação e PCR para o HIV-1 da ANRS (118), em conjugação com a informação na bula da *Supreme NZYTaQ II 2× Green Master Mix* (NZYTech, Lisboa, Portugal), tendo sido efetuadas algumas alterações (Tabelas 4 e 5). As reações de amplificação foram realizadas no termociclador *PCR Express, Model HBPX110* (Thermo Electron Corporation Hybaid).

Tabela 4. Protocolo de amplificação para a 1.^a reação da *nested* PCR, com a utilização dos *primers* externos INPS1 e INPR8.

<i>Passos do ciclo da 1.^a reação</i>	Temperatura	Tempo	Ciclos
Desnaturação inicial	95 °C	5 min	1
Desnaturação	94 °C	30 s	35
Hibridação	50 °C	30 s	
Extensão	72 °C	30 s	
Extensão final	72 °C	7,5 min	1
Manutenção	4 °C	-	-

Tabela 5. Protocolo de amplificação para a 2.^a reação da *nested* PCR, com a utilização dos *primers* internos INPS3 e INPR9.

<i>Passos do ciclo da 2.^a reação</i>	Temperatura	Tempo	Ciclos
Desnaturação inicial	95 °C	5 min	1
Desnaturação	94 °C	30 s	35
Hibridação	55 °C	30 s	
Extensão	72 °C	30 s	
Extensão final	72 °C	7,5 min	1
Manutenção	4 °C	-	-

No contexto da otimização da técnica de *nested* PCR para maximizar a obtenção de resultados positivos específicos, foram implementados diferentes protocolos de *bi-heminested* PCR. Este procedimento envolveu um total de três reações de PCR, no qual, além das duas reações mencionadas previamente, uma terceira reação foi realizada entre elas, utilizando os pares de *primers* INPS1/INPR9 ou INPS3/INPR8. O objetivo era restringir progressivamente a sequência alvo e aumentar a sensibilidade e especificidade do método. Além das alterações mencionadas, foram também otimizadas as temperaturas de hibridação dos *primers*. Inicialmente, as duas primeiras reações utilizaram uma temperatura de hibridação de 50 °C, enquanto a terceira reação foi conduzida a 55 °C. Posteriormente, a temperatura de hibridação da segunda reação foi gradualmente aumentada até 55 °C, para aumentar a restringência da reação, no intuito de melhorar a especificidade dos resultados obtidos. Na ausência de melhoria dos mesmos, foi decidido retornar ao protocolo inicial, descrito previamente.

2.4 Visualização dos produtos de PCR

Com o intuito de verificar os resultados da reação de PCR, procedeu-se à visualização dos produtos de PCR através da realização de um protocolo de eletroforese em gel de agarose.

O gel foi preparado a 1,5% (p/v) de agarose, em tampão TAE 0,5X (Anexo 7). Para um volume final de 100 mL de gel, adicionou-se 2 µL de *GreenSafe Premium* (NZYTech, Lisboa, Portugal), como agente marcador de DNA, antes da sua polimerização. Após 30 min de polimerização do gel, este foi introduzido na tina de eletroforese com tampão TAE 0,5X e foram aplicados 5 µL de cada produto de PCR em poços independentes, bem como 4 µL de marcador de massa molecular, tendo sido utilizado o *NZYDNA Ladder V* (NZYTech, Lisboa, Portugal). Não foi necessária a adição de tampão de deposição conjuntamente com as amostras, uma vez que a mistura reacional utilizada pode ser depositada diretamente nos poços do gel de agarose (entre outros, possui na sua constituição os corantes *Orange G* e xileno de cianol que funcionam como marcadores de migração). A eletroforese realizou-se a 100 V, durante 1 h e 15 min, utilizando o equipamento *Power Pac 300* (Bio-Rad, EUA) como fonte de corrente elétrica.

Após a migração das amostras, o gel foi observado sob luz ultravioleta (UV) ($\lambda=302$ nm) através de um transiluminador associado a um sistema digital de captação de imagem (*Gel Doc XR*, Bio Rad, EUA).

2.5 Purificação dos produtos de PCR

Após a visualização dos produtos de PCR, como descrito na secção anterior, foram selecionadas as amostras que apresentavam bandas inespecíficas, para além do fragmento de DNA pretendido de 723 pb, para fim da sua purificação a partir de gel de agarose. Procedeu-se à realização de um protocolo de extração utilizando o kit *NZYGelpure* (*NZYTech*, Lisboa, Portugal), de acordo com as especificações do fabricante, com pequenas modificações.

As amostras selecionadas para purificação foram submetidas a uma eletroforese a 80 V, durante 1 h 30 min, e posteriormente visualizadas como descrito previamente. Os fragmentos selecionados (723 pb) foram excisados a partir do gel com um bisturi limpo e afiado, sendo as fatias de gel transferidas para um microtubo *ependorf* de 2 mL e pesadas. Durante o protocolo, todos os passos de centrifugação foram realizados à TA numa microcentrífuga (*Eppendorf 5415D Digital Centrifuge*, Marshall Scientific, Hampton, EUA) a 12400 x g.

Primeiramente, procedeu-se à adição de 300 μ L de tampão de ligação por cada 100 mg de gel, após o que se seguiu uma incubação a 60 °C durante 10 min, com agitação ocasional, até à total dissolução do gel de agarose. Posteriormente, aplicou-se a mistura produzida numa minicoluna de purificação *NZYTech spin* colocada num tubo coletor de 2 mL e centrifugou-se, nas condições descritas acima, durante 1 min. Descartou-se o líquido coletado e prosseguiu-se para os passos de lavagem, adicionando 500 μ L de tampão de lavagem e centrifugando durante 1 min à TA, descartando-se o líquido coletado. Repetiu-se o passo de lavagem, mas utilizando agora um volume de tampão de lavagem de 600 μ L. Após esta fase foi realizado um passo de centrifugação adicional durante 1 min, com o intuito de eliminar os resíduos de etanol que se pudessem encontrar na membrana. Seguidamente, colocou-se a minicoluna de purificação *NZYTech spin* num microtubo *ependorf* estéril de 1,5 mL e adicionou-se 25 μ L de tampão de eluição ao centro da coluna (este tampão foi pré-aquecido a 60 °C, de modo a aumentar o rendimento da eluição). Incubou-se a 37 °C durante 1 min. Por fim, procedeu-se a mais um passo de

centrifugação durante 1 min para eluir o DNA que foi armazenado a -20 °C até posterior utilização.

2.6 Sequenciação nucleotídica dos produtos de PCR nas duas cadeias pelo método de Sanger

Os produtos de PCR que apresentaram uma banda específica de 723 pb foram enviados diretamente para a empresa Stab Vida (Caparica, Portugal), onde foram purificados. Após a obtenção de todos produtos de PCR purificados, estes foram submetidos a sequenciação, em ambas as cadeias, tendo esta sido levada a cabo pela empresa Stab Vida (Caparica, Portugal). Foi utilizado o método de terminação de cadeia, com a utilização de ddNTPs (do inglês, *dideoxynucleotide*), desenvolvido pelo grupo de Sanger (119). Desta forma, foram enviados 20 µL dos produtos de amplificação de forma a permitir a obtenção da quantidade mínima estimada de 400 ng de DNA por reação de sequenciação, isto porque cada fragmento seria sequenciado nos dois sentidos (direto e reverso). Para além disso, foram também enviados 50 pmoles de cada primer interno (INPS3 e INPR9) por reação de sequenciação.

2.7 Análise bioinformática das sequências nucleotídicas

2.7.1 Edição e confirmação de identidade

Primeiramente, foi realizada a edição, com base na análise dos cromatogramas correspondentes, das sequências nucleotídicas diretas e complementares, de cada molécula de DNA, obtidas por meio da sequenciação dos produtos de PCR. O objetivo dessa análise foi criar uma sequência consenso, a qual seria utilizada nas etapas de análise subsequentes. A fim de realizar essa tarefa, foi empregue o software *BioEdit Sequence Alignment Editor* (versão 7.2.5), que está acessível em <https://bioedit.software.informer.com/7.2/>. Posteriormente, procedeu-se à verificação da identidade das sequências consenso obtidas para cada amostra, por meio de comparação com sequências do *GenBank*, utilizando o programa *BLAST* (do inglês *Basic Local Alignment Search Tool*). Este programa encontra-se disponível em <http://blast.ncbi.nlm.nih.gov/Blast.cgi>, procede ao cálculo do grau de semelhança entre

sequências (em porcentagem) e faz também uma estimativa da probabilidade de as semelhanças encontradas serem devidas ao acaso (*E value*, de 0 a 1).

2.7.2 Genotipagem automatizada

Utilizando o *software REGA HIV-1 Subtyping Tool* (versão 3.46), disponível em <https://www.genomedetective.com/app/typingtool/hiv>, foi realizada uma genotipagem automatizada das sequências obtidas.

Este *software* emprega métodos filogenéticos e de *bootscanning* para identificar o subtipo de uma sequência específica e analisar possíveis eventos de recombinação. No contexto da classificação genotípica de sequências do HIV-1, o programa constrói árvores filogenéticas utilizando um conjunto de sequências de referência do grupo M. Essas árvores são complementadas por análises de *bootstrapping*, tendo em consideração apenas valores de *bootstrap* acima de 70% para a identificação do subtipo. A análise de eventos de recombinação genética só é realizada se as sequências em estudo tiverem tamanho superior a 800 pb, o que não foi o nosso caso.

2.7.3 Análise filogenética

Para uma inferência mais robusta e detalhada dos genótipos das sequências codificantes da IN, foi realizada a construção de árvores filogenéticas através de métodos manuais. Inicialmente, foi construído um alinhamento múltiplo, recorrendo ao programa *MAFFT* (versão 7, disponível em <https://mafft.cbrc.jp/alignment/server/>), referente à região codificante da IN, com as sequências em estudo, sequências obtidas por pesquisa de semelhança local entre sequências (análise *BLAST*) e sequências de referência do grupo M (subtipos, sub-subtipos e CRFs) e ainda uma sequência pertencente ao grupo O (como *outgroup*), retiradas da plataforma *Los Alamos HIV sequence database* (disponível em <https://www.hiv.lanl.gov/content/index>). Seguidamente, o alinhamento foi introduzido na plataforma *Gblocks* (versão 0.91.1, disponível em <https://ngphylogeny.fr/tools/tool/276/form>), de forma a eliminar posições mal alinhadas e regiões divergentes do alinhamento, para que este se torne mais adequado para a análise filogenética. Para a construção das árvores filogenéticas foi utilizado o *software MEGA* (versão 11), segundo o método de *neighbor-joining* (120). Para calcular as matrizes de distância genética foi utilizado o método de dois parâmetros de Kimura (121),

selecionando os parâmetros: frequência de bases igual; taxa de transições igual à de transversões; distribuição *Gamma* para a taxa entre posições; modelo de substituição GTR+ I + G e a opção *pairwise deletion* para as colunas contendo espaços abertos (*gaps*). Através do método de *bootstrapping*, foi avaliada a robustez das topologias sugeridas, utilizando um número de 1000 réplicas. Os agrupamentos formados foram considerados como consistentes, do ponto de vista filogenético, quando apresentavam valores de *bootstrap* iguais ou superiores a 70%.

2.8 Análise bioinformática das sequências aminoacídicas

2.8.1 Pesquisa de mutações associadas a resistência aos INSTI

A partir das sequências nucleotídicas foram obtidas, após tradução, sequências putativas de aminoácidos, procedendo-se à pesquisa de mutações associadas à diminuição da suscetibilidade aos ARVs inibidores da IN e detecção de outros polimorfismos genéticos, utilizando o programa *HIVdb – Genotypic Resistance Interpretation Algorithm* da *Stanford University HIV Drug Resistance Database* (versão 9.4), disponível em <https://hivdb.stanford.edu/hivdb/by-sequences/>. O recurso a este programa teve como objetivo identificar mutações principais (*major*) e mutações acessórias (*minor*), bem como outras mutações não relacionadas com a resistência aos ARVs. Esta identificação ocorre por meio de comparação das sequências em análise com uma sequência consenso de referência do subtipo B, totalmente suscetível aos ARVs, considerando cada uma das regiões codificantes analisadas. A partir desse processo de comparação, é atribuída a cada mutação uma pontuação individual, para cada fármaco analisado. Essa pontuação reflete, com base em diversos estudos publicados, a probabilidade de a mutação encontrada estar associada à resistência ao respectivo fármaco. A soma das pontuações atribuídas a cada mutação, relativamente a determinado fármaco, permite estimar o nível de resistência fenotípica para esse fármaco associada a cada sequência aminoacídica putativa em estudo. Com base no valor total obtido, a resistência é classificada numa das seguintes cinco categorias:

- **Suscetibilidade** (pontuação de 0 a 9) – Não há evidências de redução da suscetibilidade ao ARV em comparação com a estirpe selvagem.

- **Resistência potencial de nível baixo** (pontuação de 10 a 14) – O vírus não demonstra suscetibilidade diminuída ao fármaco, mas exibe mutações que indicam exposição prévia a ARVs ou que só conferem resistência quando combinadas com mutações adicionais.
- **Resistência de nível baixo** (pontuação de 15 a 29) – O vírus pode manifestar uma redução na suscetibilidade, *in vitro*, ao fármaco, o que pode resultar numa resposta virológica subótima no tratamento com o mesmo de indivíduos infectados.
- **Resistência de nível intermédio** (pontuação de 30 a 59) – Sugere uma alta probabilidade de suscetibilidade reduzida ao fármaco, porém irá manter uma significativa atividade antiviral remanescente.
- **Resistência de nível elevado** (pontuação maior ou igual a 60) – Indica um nível de resistência semelhante ao observado *in vitro*, em vírus altamente resistentes ao fármaco. Essa resistência pode resultar numa resposta virológica significativamente reduzida, ou até mesmo nula, em termos clínicos, nos indivíduos infectados.

2.9 Análise estatística

Para a análise dos dados demográficos e epidemiológicos que caracterizam a amostra populacional em estudo, incluindo variáveis como sexo, idade, estado civil, nível de escolaridade, área de residência, local de residência e outras infecções diagnosticadas, procedeu-se a uma análise de distribuição, dada a natureza categórica dessas variáveis. No que diz respeito à idade, além da análise de distribuição, foram calculadas a média e o desvio padrão.

A associação entre os diversos polimorfismos genéticos naturais observados e os genótipos virais mais frequentes na amostra populacional em estudo (A, C e F1), através da pesquisa de assinaturas genéticas, foi analisada recorrendo ao teste exato de Fisher bilateral ou ao teste de χ^2 , aplicados à análise de tabelas de contingência 2x2

(disponíveis em <https://www.socscistatistics.com/>), considerando um valor de significância de 1%.

3. Resultados e discussão

3.1 Caracterização da amostra populacional em estudo

Este trabalho incluiu amostras de sangue coletadas de um grupo de 226 indivíduos internados, utentes de consultas externas ou utentes de urgência do HGB, na cidade de Benguela, em Angola, que foram diagnosticados como seropositivos para o HIV-1.

Todos os indivíduos incluídos tinham nacionalidade angolana e, à data da colheita das amostras, tinham idades compreendidas entre 0 e 95 anos (idade média de $38 \pm 19,7$ anos). Eram 42% do sexo masculino e 58% do sexo feminino apresentando uma maior percentagem de solteiros (39%) ou casados (37%). A grande maioria possuía, pelo menos, o ensino básico (33%), existindo uma percentagem considerável sem escolaridade e com o ensino secundário concluído (ambos 30%), os restantes 7% representavam os indivíduos que concluíram o ensino superior (dados apresentados na Figura 11).

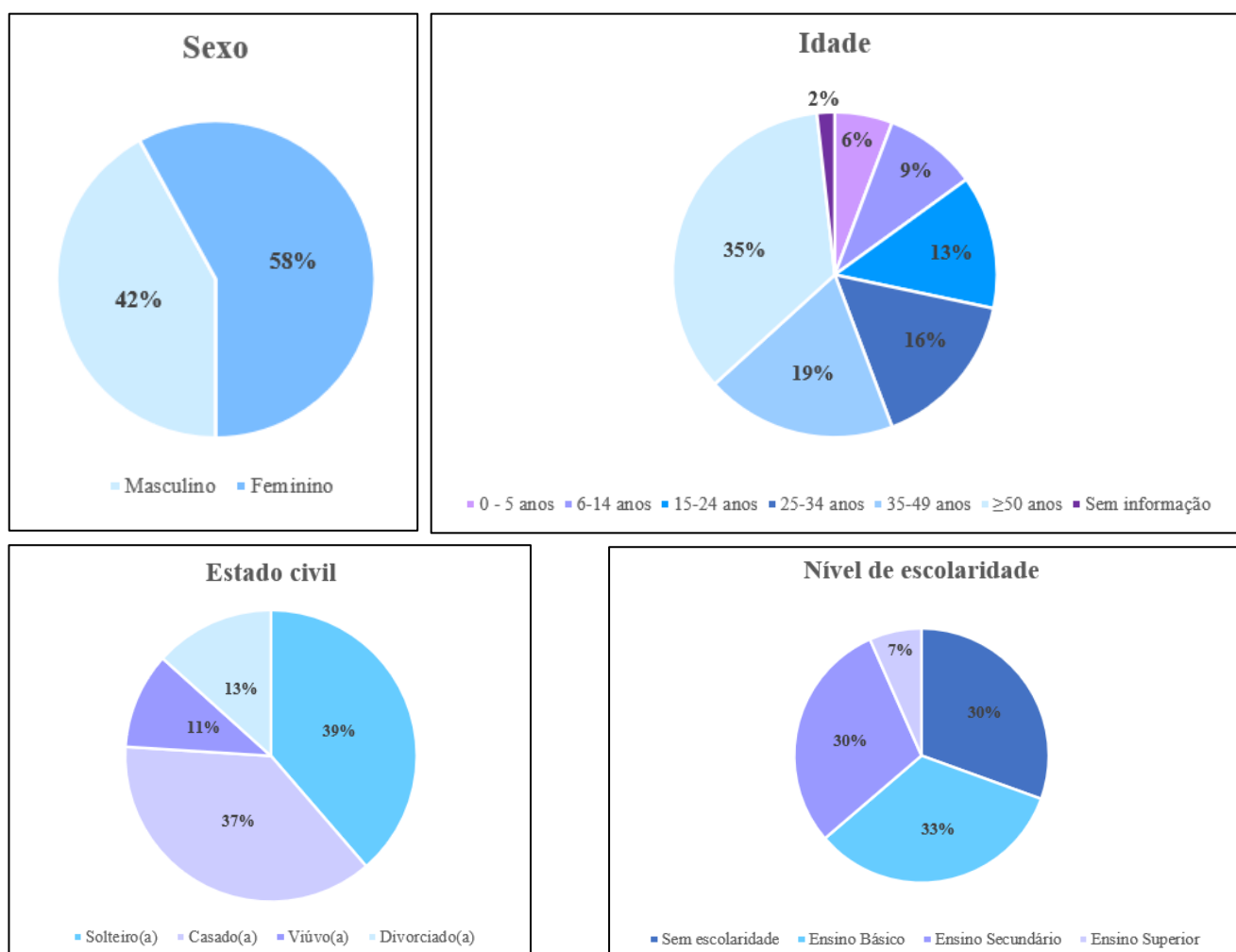


Figura 11. Representação gráfica (em percentagem) dos dados demográficos da amostra populacional em estudo, relativos ao sexo, idade, estado civil e nível de escolaridade.

Todos os participantes neste estudo tinham nacionalidade angolana e eram residentes na província de Benguela, sendo que 41% vivia na cidade de Benguela, enquanto os restantes residiam noutras partes da província, como se pode verificar na Figura 12. Em geral, a maioria dos indivíduos residia em áreas urbanas (61%). Uma vez que esta proposta de trabalho está incluída num estudo transversal, de maior amplitude, cujo objetivo consiste na caracterização de algumas infeções de transmissão sexual nesta população, foram também colhidos dados disponíveis sobre coinfeção com sífilis e hepatite B (conforme mencionado na secção 2.1). Foi possível observar que 17% (39/226) dos participantes apresentavam outras infeções além da infeção pelo HIV-1. Entre estes, 54% (21/39) tinham sífilis, 41% (16/39) estavam infetados com o vírus da hepatite B e 5% (2/39) apresentavam coinfeção HIV-1/sífilis/hepatite B (Figura 12).

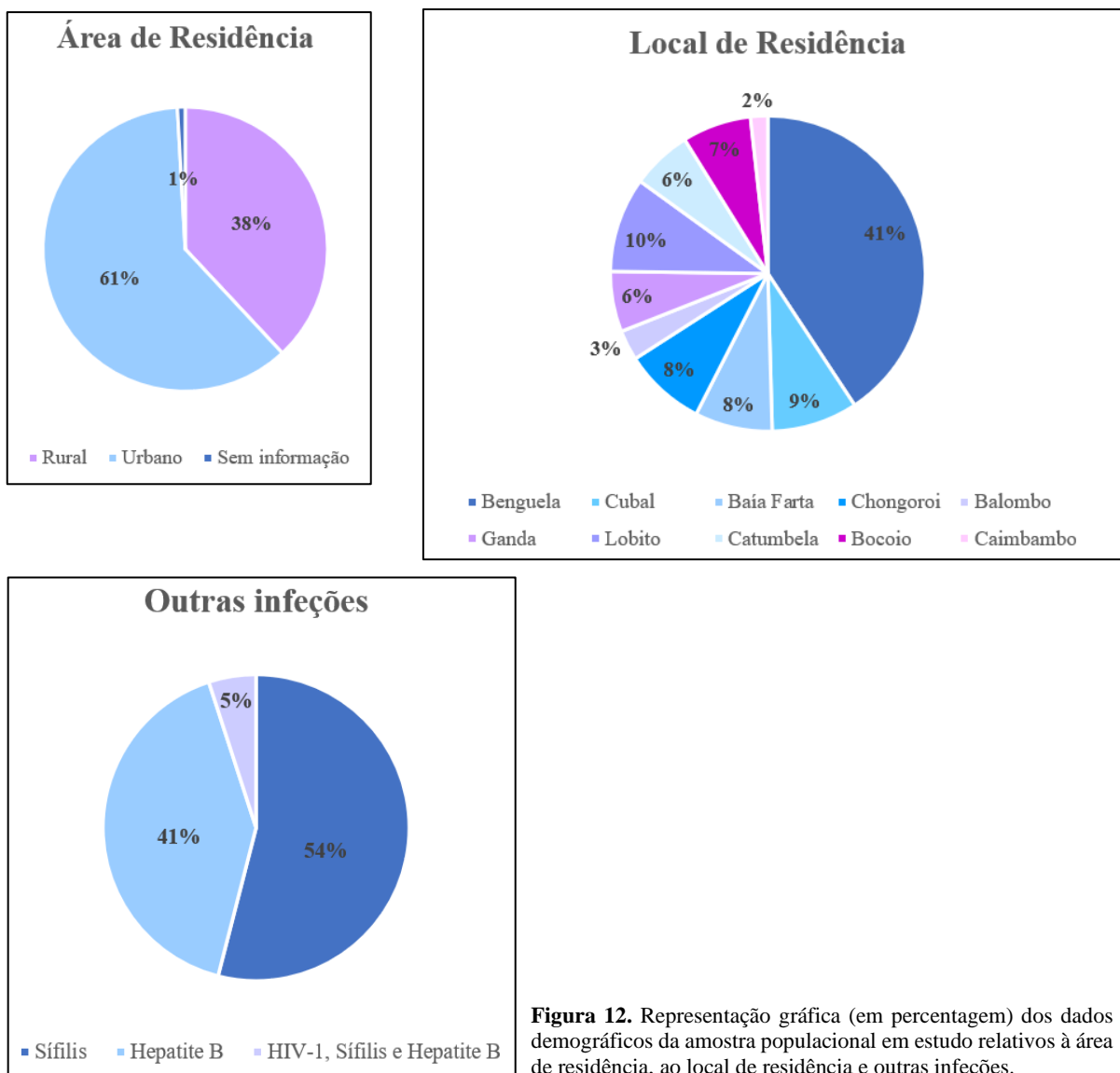


Figura 12. Representação gráfica (em percentagem) dos dados demográficos da amostra populacional em estudo relativos à área de residência, ao local de residência e outras infeções.

3.2 Amplificação da região codificante da integrase por *nested* PCR

A amplificação da região codificante da IN foi realizada por *nested* PCR, a partir do DNA proviral presente em 226 amostras de sangue seco imobilizado em cartões FTA[®] (GE Healthcare, Reino Unido), cujo protocolo se encontra descrito na secção 2.3.2. Conforme descrito na secção 2.3.2.1, foram conduzidas diversas tentativas de otimização do protocolo utilizando a técnica de *bi-heminested* PCR, no entanto, devido à ausência de sucesso nos resultados obtidos, o protocolo foi abandonado em favor do retorno ao procedimento original. Os *primers* internos INPS3 e INPR9, usados na 2.^a ronda da reacção de PCR, possibilitaram a amplificação de um fragmento de DNA com 723 pb, cujas coordenadas estão descritas na Tabela 3, mencionada previamente, referentes à sequência de referência HXB2. Após a realização do protocolo de amplificação descrito na secção 2.3.2.1, foram obtidas 97 amostras positivas (42,9%) do total das 226 (Figura 13). Para as restantes amostras (n=129), apesar de várias tentativas de otimização do protocolo, não foi possível a obtenção de produtos de amplificação.

Amplificação da região codificante da integrase por *nested* PCR

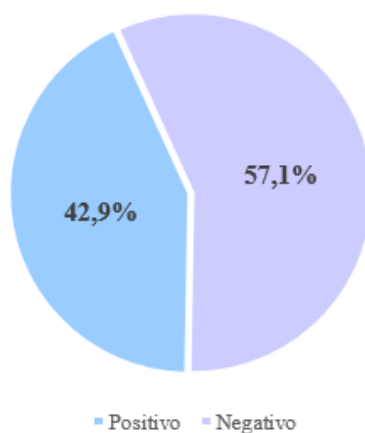


Figura 13. Representação gráfica (em percentagem) dos resultados obtidos após amplificação por *nested* PCR da região codificante da IN.

Observando a taxa de positividade de 42,9%, este valor relativamente baixo poderá estar relacionado com o suporte utilizado para recolha e preservação das amostras sanguíneas, nomeadamente, os cartões *Whatman Indicating FTA[®] Elute Micro Cards* (GE Healthcare, Reino Unido), referidos anteriormente. Efetivamente, foi referido num estudo realizado anteriormente a necessidade de abandonar o protocolo recomendado pelo fabricante para a extração do DNA proviral referente a este tipo de cartões (74), sendo então aplicado o

protocolo de preparação de discos desenvolvido no IHMT, que se encontra descrito na secção 2.2.

Como se pôde verificar na literatura, este tipo de cartões comporta-se de forma diferente dos originais quando exposto a condições de humidade relativa e temperatura elevadas, resultando, nalguns casos, no comprometimento da integridade do substrato e das amostras presentes nos mesmos, devendo assim ser armazenado em condições de baixa humidade e temperaturas inferiores a 4°C (122–124). Foi também descrito que, aquando da colheita das amostras, o armazenamento deve ser feito a temperatura inferior a -20°C, não se devendo deixar mais de 14 dias à temperatura ambiente (122), o que não foi realizado. Estes fatores influenciam a preservação das amostras e, por conseguinte, a capacidade de deteção por PCR das diferentes variantes de HIV-1, dada a elevada diversidade genética que este vírus apresenta (122–124). É, no entanto, importante realçar que as particularidades do HIV-1, em termos de diversidade genética, como também a localização geográfica na qual o estudo foi realizado (Angola), região com uma multiplicidade de subtipos em circulação, constituíram sempre um desafio para uma amplificação bem-sucedida. Um dos pontos fulcrais a considerar seria a capacidade dos *primers* hibridarem eficaz e especificamente à região alvo pretendida.

As amostras positivas foram seleccionadas através da visualização dos produtos de PCR após eletroforese em gel de agarose, sendo estas as que apresentavam uma banda com um tamanho aproximado de 723 pb. Observaram-se vários padrões de migração nas diferentes amostras. Como se pode verificar na Figura 14 (amostras 380-384, 386 e 387), obtiveram-se padrões de amplificação específicos, com apenas uma banda, com o tamanho pretendido. Porém, algumas reações de amplificação também apresentavam produtos inespecíficos, como se verifica, a título de exemplo, na Figura 15. O procedimento experimental subsequente variava de acordo com a especificidade dos produtos de amplificação, conforme mencionado previamente (secção 2.5).

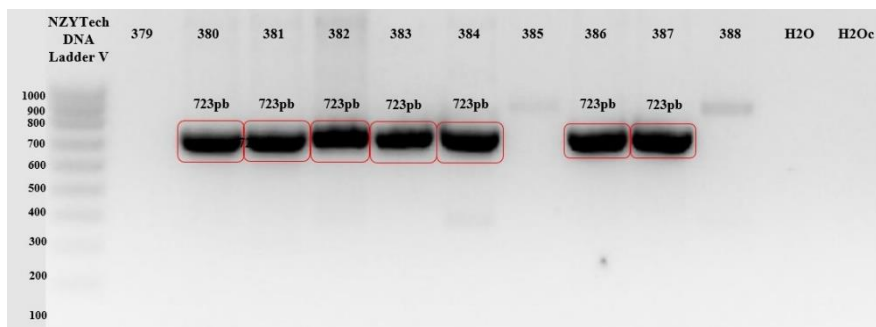


Figura 14. Visualização dos padrões de migração dos produtos de amplificação por *nested* PCR da região codificante da IN do HIV-1, após eletroforese em gel de agarose a 1,5% (p/v), em tampão TAE 0,5X, corado com *Green Safe Premium* (NZYTech, Portugal). Legenda: *NZYTech DNA Ladder V* – Marcador de massa molecular (100-1000 pb); 379-388 – Amostras em estudo; H₂O e H₂Oc– Controlos negativos.

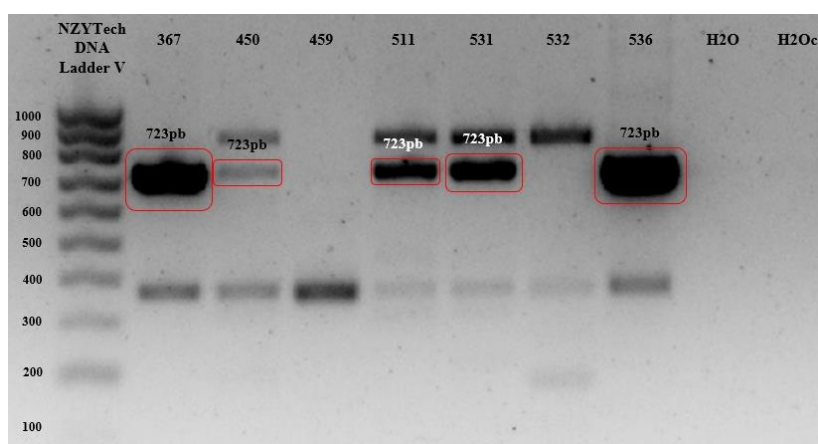


Figura 15. Visualização dos padrões de migração dos produtos de amplificação por *nested* PCR da região codificante da IN do HIV-1, após eletroforese em gel de agarose a 1,5% (p/v), em tampão TAE 0,5X, corado com *Green Safe Premium* (NZYTech, Portugal). Legenda: *NZYTech DNA Ladder V* – Marcador de massa molecular (100-1000 pb); 367-536 – Amostras em estudo; H₂O e H₂Oc– Controlos negativos.

Nas amostras que apresentavam uma banda específica, com o tamanho pretendido, e uma quantidade de DNA aceitável (88/97) (90,7%), este foi enviado para sequenciação nucleotídica diretamente a partir da mistura de PCR, com realização de uma purificação pela própria empresa de fornecimento de serviços de sequenciação (STAB VIDA, Caparica, Portugal). Nas amostras que apresentavam bandas inespecíficas (9/97) (9,3%), que poderiam interferir, desde logo, com a qualidade das sequências nucleotídicas a obter, os produtos de PCR de tamanho esperado foram submetidos a purificação a partir de gel de agarose, após separação eletroforética e, só depois, enviados para sequenciação.

3.3 Análise bioinformática das sequências nucleotídicas

3.3.1 Edição e confirmação de identidade

Após a sequenciação nucleotídica dos 97 fragmentos e obtenção de duas sequências, e respectivos cromatogramas, para cada amostra, uma no sentido direto (*sense*) e outra no sentido reverso (*antisense*), de forma a ser possível a sua comparação para obtenção final de uma sequência consenso, procedeu-se à sua análise e edição. A sequenciação nucleotídica em ambos os sentidos teve como objetivo principal aumentar a fiabilidade dos resultados, considerando o objetivo de caracterizar mutações ligadas à resistência aos fármacos antirretrovirais inibidores da IN. Para determinadas posições, observou-se a incapacidade de atribuição de apenas um nucleótido, devido à existência de *quasispecies*, sendo assim necessária a atribuição de uma base degenerada, com base no código IUPAC (do inglês *International Union of Pure and Applied Chemistry*) (Anexo 4). Na totalidade, foram obtidas 92 sequências nucleotídicas consenso, uma vez que em 5/97 fragmentos de DNA sequenciados (5,2%) não foi possível obter sequências devido à sua fraca qualidade (sequências/cromatogramas truncados, cromatogramas com elevado “ruído de fundo” e/ou um elevado número de degenerações, que tornavam difícil a sua interpretação).

Depois de obter uma sequência nucleotídica consenso para cada amostra, procedeu-se a uma confirmação preliminar da sua identidade através da comparação com sequências depositadas em bases de dados internacionais, utilizando o programa *BLAST* (disponível em <https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Posteriormente, foram registadas as sequências da base de dados que demonstraram maior similaridade com as sequências em análise, bem como o respetivo *E value*, para posterior utilização na análise filogenética. Como resultado, foram identificadas percentagens de similaridade com sequências de HIV-1 iguais ou superiores a 92,8% e *E value* igual a zero.

3.3.2 Genotipagem automatizada

Após obtenção das sequências em análise, prosseguiu-se para uma genotipagem preliminar automatizada, antes da construção da árvore filogenética manual. Para tal, foi utilizado o *software REGA HIV-1 Subtyping Tool* (versão 3.46), como descrito no ponto 2.7.2. Este procedimento teve como objetivo a obtenção de um conhecimento inicial

sobre a diversidade dos genótipos presentes na amostra populacional em estudo. Como pode ser observado na Figura 16, verificou-se a predominância do sub-subtipo F1 (22/92, 23,9%), seguido dos subtipos C (17/92, 18,5%), G (15/92, 16,3%) e A1 (8/92, 8,7%), observando-se ainda a presença do subtipo H (1/92, 1,1%). Uma vez que, para este *software*, a análise de eventos de recombinação genética só é efetuada para sequências de tamanho superior a 800 pb, e as sequências em análise apresentam um tamanho inferior (682 pb), não foi possível a sua realização. Assim, é importante realçar que para 29/92 (31,5%) sequências em estudo não foi possível a atribuição de qualquer genótipo/subtipo, o que pode indicar que estamos face a uma sequência recombinante (genoma mosaico), ou uma sequência muito divergente das sequências de referência utilizadas para a realização da análise.

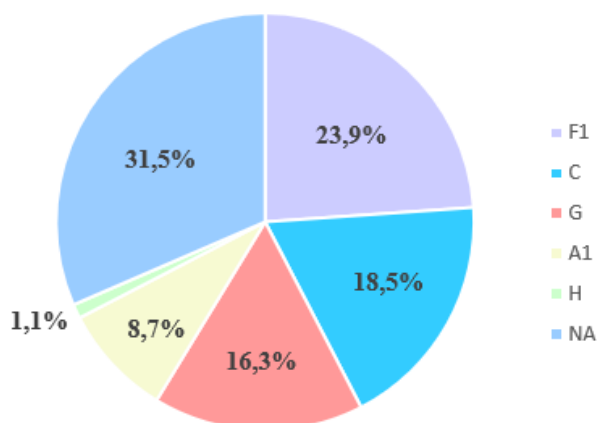


Figura 16. Representação gráfica (em percentagem) dos subtipos obtidos após genotipagem preliminar automatizada das sequências nucleotídicas em estudo, utilizando o *software REGA HIV-1 Subtyping Tool* (versão 3.46). NA é a designação para as sequências cujos genótipos não foram atribuídos.

Optou-se por utilizar o *software REGA HIV-1 Subtyping Tool* pois os resultados gerados são fundamentados em análise filogenética, ao contrário de outros programas disponíveis *online* que se baseiam principalmente na procura de semelhança direta entre sequências. Isso confere uma maior significância à análise, apesar de algumas restrições que possam existir (125–127).

3.3.3 Análise filogenética

Para uma caracterização mais significativa da diversidade genética das estirpes de HIV-1 presentes na amostra em estudo, procedeu-se à construção de uma árvore filogenética

por métodos não automatizados, como descrito no ponto 2.7.3 (Figura 17). Esta árvore filogenética, construída a partir de um alinhamento múltiplo de sequências nucleotídicas que abrange um total de 682 nts, compreende um conjunto de 150 sequências, incluindo as 92 sequências em análise, 30 sequências obtidas da base de dados *GenBank* por meio de pesquisa de regiões de semelhança local entre sequências, utilizando o programa *BLAST*, 27 sequências de referência do grupo M do HIV-1 e uma sequência de referência do grupo O do HIV-1, que foi utilizada como *outgroup* para esta análise. Resumidamente, para a construção desta árvore filogenética foi utilizado o método de *neighbor-joining* (120), com a matriz de distância genética calculada pelo método de dois parâmetros de Kimura (121). A robustez da topologia resultante foi avaliada por meio do método de *bootstrapping*, com a realização de 1000 réplicas. Os agrupamentos foram considerados filogeneticamente consistentes quando os valores de *bootstrap* obtidos foram iguais ou superiores a 70%.

Após análise da árvore filogenética obtida, foi possível atribuir um genótipo a 80 das sequências nucleotídicas em análise, estando estas posicionadas em nove agrupamentos filogenéticos distintos. Destes agrupamentos, três (CRF02_AG, H e CRF18_cpx, conforme ilustrado na Figura 17) são suportados por valores de *bootstrap* significativos, com percentagens de 76%, 90% e 95%, respetivamente. No entanto, mesmo nos casos de agrupamentos que não alcançaram significância do ponto de vista filogenético, as suas topologias apresentam-se, de maneira geral, bem definidas. Portanto, a identificação dos genótipos da grande maioria das sequências em estudo pôde ser realizada com um grau de certeza elevado. A um conjunto de 12 sequências, no entanto, não foi possível atribuir um genótipo, devido à sua posição ambígua na árvore filogenética ou à notável divergência genética apresentada em relação ao nó ancestral que as suporta. Estas sequências seriam naturais candidatas a uma posterior análise de *bootscanning*, com o objetivo de investigar a possível ocorrência de fenómenos de recombinação intergenotípica, contudo, esta não foi efetuada devido a restrições temporais.

Das 29 sequências não genotipáveis pelo método automatizado, 22 (AO_202, AO_210, AO_334, AO_356, AO_367, AO_380, AO_383, AO_384, AO_386, AO_387, AO_404, AO_406, AO_416, AO_433, AO_449, AO_456, AO_462, AO_472, AO_483, AO_488, AO_508 e AO_521) foram classificadas somente recorrendo à árvore filogenética. As restantes 7/29 sequências (AO_398, AO_474, AO_475, AO_480, AO_481, AO_527 e

AO_533) não genotipadas pelo método automatizado, juntamente com mais cinco (AO_434, AO_466, AO_494, AO_520 e AO_531) constituem o conjunto total de 12 sequências não genotipáveis na amostra populacional em estudo. As 58 sequências restantes foram classificadas por conjugação de informação entre a genotipagem automatizada e a posição relativa na árvore filogenética final.

Numa primeira análise da árvore filogenética construída, é evidente que uma grande parte dos grupos formados não apresentaram valores de *bootstrap* superiores a 70%, sugerindo que a fiabilidade da inferência filogenética realizada é algo limitada. Esta limitação pode ser atribuída, em parte, ao tamanho da região analisada, que compreende apenas 682 nts, sendo que para uma inferência mais robusta seria necessária a análise de uma sequência de maior tamanho, idealmente, o genoma completo. Para além disso, apesar da região codificante da IN ser mais conservada que as correspondentes das restantes enzimas, é importante realçar que as amostras em estudo são provenientes de uma região geográfica onde é reconhecida uma elevada diversidade genética de formas virais circulantes. Assim, será provável estar na presença de uma grande diversidade de genótipos divergentes, CRFs e/ou URFs, não sendo possível formar grupos monofiléticos suportados por valores de *bootstrap* estatisticamente significativos com genótipos puros. Tendo em consideração esses fatores, prossegue-se para uma análise mais detalhada dos resultados obtidos.

O agrupamento das sequências AO_202, AO_235, AO_245, AO_374, AO_383, AO_386, AO_387, AO_404, AO_416, AO_423, AO_432, AO_433, AO_449, AO_455, AO_462, AO_483, AO_489, AO_508 e AO_511, com sequências do subtipo A apresenta um suporte estatístico apoiado por valores baixos, devido a valores de *bootstrap* inferiores a 70%, o que indica falta de consistência. No entanto, a topologia deste agrupamento permitiu, para fins deste estudo, a classificação destas amostras como subtipo A (Figura 17). É importante realçar que a inferência filogenética neste caso possui uma fiabilidade limitada e seria necessário estudar uma região mais extensa da sequência, em vez de se restringir apenas à região codificante da IN. Este agrupamento demonstra, também, uma grande diversidade intragenotípica, inferência apoiada pelo padrão de distribuição dos diferentes ramos neste agrupamento da árvore filogenética. A sequência AO_508, por exemplo, difere claramente das demais, pois forma um grupo monofilético consistente com duas sequências de referência do sub-subtipo A2 (valor de *bootstrap* de 95%), sendo assim possível atribuir este sub-subtipo a essa sequência.

Em estudos realizados em 2009 foi verificado que o subtipo A era predominante em Angola o que não seria surpreendente dado os países que faziam fronteira com este país (RDC e RC) apresentarem também uma predominância deste subtipo (61,65). Em 2014, começa-se a observar uma diminuição da prevalência do subtipo A (66) acompanhada de um aumento da prevalência do subtipo C, que poderia ser explicado pelo aumento da sua prevalência nos países fronteiriços, nomeadamente a região sul da RDC e Zâmbia (66). Assim, ao longo do período 2000-2019 tem-se vindo a verificar uma diminuição na prevalência do subtipo A em Angola, juntamente com um aumento de formas recombinantes (71). Os resultados apresentados neste trabalho vão de encontro ao registado na literatura, uma vez que, como se pode observar na Tabela 6, para além da prevalência de ambos os subtipos ser equivalente, no seu conjunto, ambos os subtipos concorrem para cerca de 40% das sequências caracterizadas.

Tabela 6. Resultados da genotipagem para a região codificante da IN das 92 sequências em estudo, após construção da árvore filogenética (NA= genótipo não atribuído).

Genótipo	Número de sequências (%)
F1	26 (28,3)
A	19 (20,7)
C	18 (19,6)
CRF27_cpx	4 (4,3)
G	3 (3,3)
CRF02_AG	3 (3,3)
CRF19_cpx	3 (3,3)
D	2 (2,2)
H	1 (1,1)
CRF18_cpx	1 (1,1)
NA	12 (13,0)
Total	92 (100)

Ainda no agrupamento A, é importante observar a presença de um subgrupo monofilético consistente, que inclui três das sequências angolanas (AO_356, AO_384 e AO_406), juntamente com uma sequência CRF19_cpx, sustentado por um valor de *bootstrap* de 92%. Considerando este valor elevado e que a região específica do genoma que estamos a analisar tem na CRF19_cpx uma estrutura mosaico complexa (conforme representada no Anexo 5), na qual a região codificante da IN inclui dois pontos de recombinação,

originando um mosaico A/G/A, pode afirmar-se com alguma certeza que as amostras em causa, na região em estudo, têm uma estrutura mosaico CRF19_cpx. Dentro deste agrupamento, foram observados eventos de hipermutação pela APOBEC na sequência AO_356 (Figura 17).

Relativamente aos dois agrupamentos que incluem sequências da CRF124_cpx e de referência da CRF27_cpx, estes constituem grupos monofiléticos com valores de *bootstrap* estatisticamente significativos, sendo eles 91% e 74%, respetivamente. Foi observado que a estrutura da CRF124_cpx (Anexo 5) na região codificante da IN contém uma porção da CRF27_cpx (Anexo 5), o que suporta a posição relativa de ambas as CRFs na árvore filogenética. Como tal, na medida em que ambas as CRFs têm exatamente a mesma estrutura recombinante G/J, característica da CRF27_cpx, para a região genómica em estudo (região codificante da IN e não o genoma completo), as sequências AO_363, AO_448, AO_482 e AO_538 foram classificadas como CRF27_cpx. Um estudo desenvolvido em 2012 identificou vários recombinantes que apresentavam uma porção parcial do gene *pol* desconhecida, que não agrupavam com nenhuma CRF conhecida, mas entre eles constituíam um grupo monofilético bem suportado (67), sendo denominados de “recombinantes_UH”. Mais tarde, um estudo realizado em 2022 caracterizou melhor estas sequências e descreveu-as como uma nova CRF em Angola designada de CRF124_cpx (69). Esta CRF é composta por fragmentos dos subtipos A, G, H e da CRF27_cpx (69), o que vai de encontro com aos resultados apresentados na Figura 17, onde se consegue observar uma clara relação filogenética no posicionamento relativo das sequências em estudo agrupadas com a CRF124_cpx e com a CRF27_cpx. Finalmente, a estrutura mosaico G/J para a região codificante da IN, nesta última CRF, suporta uma maior proximidade filogenética com o subtipo G (e CRF02_AG), como se pode observar na árvore filogenética. Para além disto, este agrupamento revelou duas sequências (AO_363 e AO_482) que apresentaram eventos de hipermutação pela APOBEC (Figura 17).

No que se refere às sequências AO_7, AO_309 e AO_473, que foram genotipadas como CRF02_AG, elas constituem um grupo monofilético, com uma única sequência de referência dessa CRF, com origem nos Camarões, apresentando um valor de *bootstrap* de 76%. Assim, com base na estrutura mosaico da CRF02_AG (representada no Anexo 5), particularmente na região codificante da IN, que inclui um ponto de recombinação entre

o subtipo G e uma região A/G, e considerando ainda a sua posição relativa na árvore filogenética (próxima das sequências do grupo G), podemos afirmar com confiança que essas sequências serão, de facto CRF02_AG na região em análise. Relativamente ao agrupamento que inclui as sequências AO_237, AO_373 e AO_485 e sequências de referência do subtipo G, é relevante observar que ele não se configura como um agrupamento consistentemente robusto, visto que o valor de *bootstrap* que o suporta é inferior a 70%. Contudo, dentro desse mesmo agrupamento, as amostras AO_237, AO_373 e AO_485 formam um subgrupo consistente (valor de *bootstrap* de 84%) com sequências de referência do subtipo G com origem em Espanha e Portugal, demonstrando a sua grande proximidade evolutiva. Levando ainda em consideração a raridade da circulação do subtipo G na Europa, mas a sua elevada expressão na epidemia portuguesa (para Portugal, a *Los Alamos HIV Sequence Database* reporta 38,7% de sequências do subtipo G), podemos especular sobre a mais que provável origem em Angola deste subtipo no nosso país.

Quanto às sequências em estudo AO_367 e AO_488, estas emparelham com as sequências de referência do subtipo D (Figura 17). Ainda que o agrupamento formado não apresente um valor de *bootstrap* estatisticamente significativo, a topologia da árvore filogenética sugere uma clara similaridade das sequências em estudo com este subtipo, pelo que nele foram classificadas.

Na Figura 17, é também possível observar um agrupamento monofilético, estatisticamente significativo (suportado por um valor de *bootstrap* de 95%), que inclui a sequência AO_334 juntamente com duas sequências de referência classificadas como CRF18_cpx. Tendo ainda em conta a estrutura mosaico altamente complexa característica dessa CRF (Anexo 5) na região codificante da IN (U/G/A/G), apresentando dois pontos de recombinação e duas regiões de genótipo dúbio (U e A/G), podemos afirmar com elevada fiabilidade que a sequência AO_334, na região em análise, é uma CRF18_cpx. Esta forma recombinante foi pela primeira vez descrita, em 2005, em Cuba (128), país que manteve, e mantém, relações sociais e políticas com Angola, particularmente durante o período da guerra civil. A circulação desta CRF foi descrita em Angola pela primeira vez em 2012 (67).

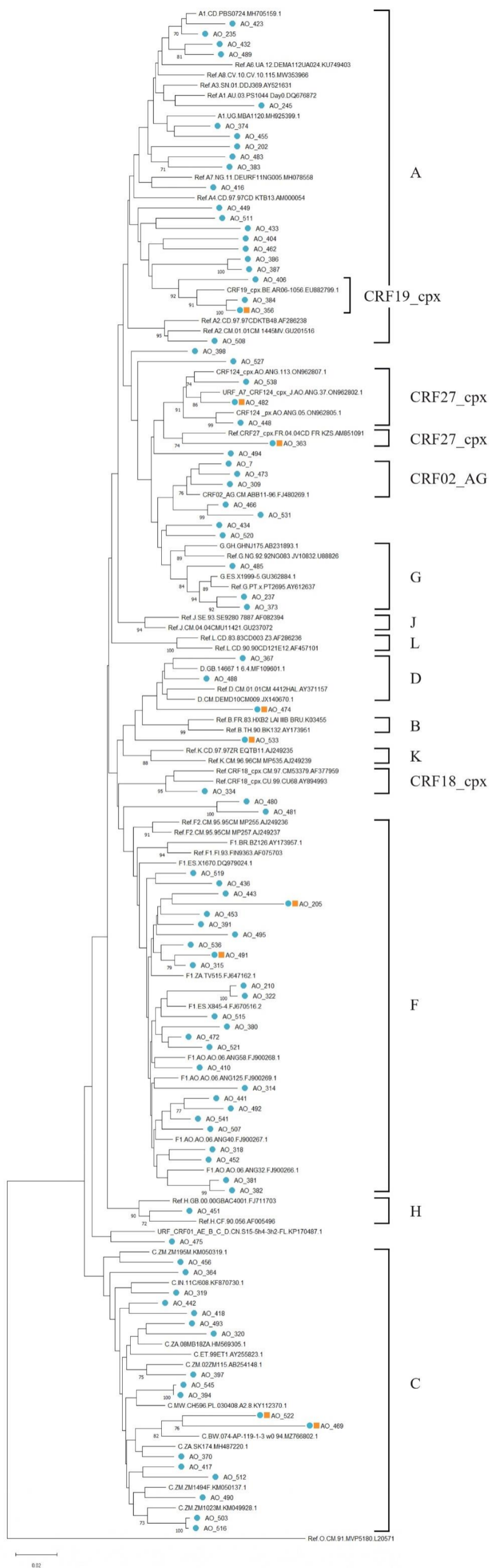


Figura 17. Relações filogenéticas entre 150 seqüências nucleotídicas da região codificante da integrase, baseadas num alinhamento múltiplo de 682 nucleótidos. Na árvore filogenética foram incluídas as 92 seqüências em estudo (indicadas com o símbolo ●), 27 seqüências de referência do grupo M (A1-A4, A6-A8, B, D, F1, F2, G, H, J, K, L, CRF18_cpx e CRF27_cpx) e uma do grupo O (como *outgroup*) e 30 seqüências obtidas por pesquisa de semelhança local entre seqüências (análise *BLAST*). Com exceção das seqüências em estudo, a designação de cada seqüência inicia-se com o respetivo genótipo, seguido do código do país de origem (AO – Angola, AU – Austrália, BE – Bélgica, BR – Brasil, BW – Botsuana, CD – República Democrática do Congo, CF – República Centro-Africana, CM – Camarões, CN – China, CU – Cuba, CV – Cabo Verde, ES – Espanha, ET – Etiópia, FI – Finlândia, FR – França, GB – Reino Unido, GH – Gana, IN – Índia, MW – Maláui, NG – Nigéria, PT – Portugal, SE – Suécia, SN – Senegal, TH – Tailândia, UA – Ucrânia, UG – Uganda, ZA – África do Sul, ZM – Zâmbia), nome da estirpe/isolado viral e número de acesso no *GenBank*. A inferência filogenética foi realizada com recurso ao *software MEGA* versão 11, 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 2% de distância genética). A robustez da inferência filogenética foi analisada por *bootstrapping* de 1000 amostragens aleatórias, considerando-se significativos valores $\geq 70\%$. As seqüências nas quais foi detetada a ocorrência de fenómenos de hipermutação estão indicadas com o símbolo ■.

No que diz respeito ao subtipo F, é notório que o grupo de sequências em estudo não inclui representantes do sub-subtipo F2. No que diz respeito a 26 sequências (AO_205, AO_210, AO_314, AO_315, AO_318, AO_322, AO_380, AO_381, AO_382, AO_391, AO_410, AO_436, AO_441, AO_443, AO_452, AO_453, AO_472, AO_491, AO_492, AO_495, AO_507, AO_515, AO_519, AO_521, AO_536 e AO_541), estas apresentam-se agrupadas juntamente com sequências de referência do sub-subtipo F1. Apesar deste agrupamento não estar suportado por um valor de *bootstrap* consistente (<70%), a topologia da árvore evidencia uma clara relação filogenética entre as sequências em estudo e o sub-subtipo F1, com uma elevada diversidade intragenotípica. As sequências AO_205, AO_210, AO_314 e AO_322, ainda que classificadas, nesta análise, como pertencentes ao sub-subtipo F1, seriam passíveis de uma posterior análise de recombinantes, dada a divergência genética elevada sugerida pelos ramos que as suportam na árvore filogenética (Figura 17), não tendo sido realizada pelas razões referidas anteriormente. Entre estas, a sequência AO_205, juntamente com a sequência AO_491, evidenciando esta notável divergência genética, foram selecionadas para pesquisa de possíveis eventos de hipermutação, tendo sido verificados. O sub-subtipo F1 é descrito como o mais frequente na região norte de Angola (62,67,71), sendo que há uma diminuição da sua prevalência de norte para sul deste país (67). Os subtipos C e F1 foram apresentados como os mais frequentes na maioria dos estudos revistos por Sebastião *et al.* (2021) em Angola, foi também relatado um aumento da prevalência do sub-subtipo F1 entre os anos de 2000-2019 (129), o que pode justificar a predominância deste subtipo na amostra populacional em estudo. Na árvore filogenética (Figura 17), foram evidenciadas relações filogenéticas com sequências de referência originárias do Brasil, Angola, Camarões, África do Sul e alguns países europeus. Estas observações vão de encontro aos resultados do estudo de Bello *et al.* (2012), que refere que o subtipo F1 é excepcionalmente prevalente em Angola, Brasil e Roménia. O contexto epidemiológico em que se dispersou está relacionado com o longo período de guerra civil em Angola e a infeção de um grande número de crianças na Roménia (68). As três subepidemias do subtipo F1 começaram a propagar-se mais ou menos na mesma altura, entre o final da década de 1970 e o início da década de 1980. Este período coincide com as importantes mudanças sociopolíticas que ocorreram em Angola após o início da guerra civil em 1975. Esta, não só esteve associada a uma importante vaga de emigração, mas também recebeu

intervenção de várias potências estrangeiras, nomeadamente, Cuba, países do Bloco de Leste (incluindo a Roménia), EUA, África do Sul e Zaire (atual RDC) (68). Estas circunstâncias conferiram uma dimensão internacional ao conflito angolano que pode ter espoletado a migração de vírus do sub-subtipo F1 da África Central para a Europa e a América do Sul, bem como a sua larga disseminação em Angola (68), especialmente na região norte.

O grupo monofilético relativo ao subtipo H é sustentado por um valor de *bootstrap* de 90% (Figura 17), o que permite uma classificação genotípica inequívoca da sequência AO_451 como pertencente a esse subtipo, relativamente à região em estudo. O subtipo H tem sido reportado em Angola por vários estudos (71), embora, tal como no presente trabalho, em prevalências relativamente baixas.

A sequência AO_475 emparelha com uma sequência de referência da China reportada como uma URF complexa em 2015 (130), apesar deste agrupamento não ser suportado por um valor de *bootstrap* estatisticamente significativo (<70%). Esta URF é uma forma recombinante de terceira geração, resultante das CRF55_01B, CRF01_AE e CRF07_BC, na qual, na estrutura mosaico final, segmentos genómicos dos subtipos B e C foram inseridos na matriz genómica da CRF01_AE (130). A CRF01_AE teve origem na África Central e é agora muito mais prevalente no leste e sudeste Asiático, tendo tido uma rápida expansão em diferentes grupos de risco e regiões geográficas na China, desde 1980 (131). Para além disto, na árvore filogenética de genoma completo apresentada por Gui *et al.* (2015), evidencia-se um agrupamento monofilético entre esta URF e a CRF01_AE proveniente dos Camarões, com um valor de *bootstrap* de 100% (130). Ambos os aspetos mencionados parecem corroborar uma possível relação filogenética entre a sequência AO_475 de Benguela e esta URF complexa descrita na China. Neste contexto, digno ainda de realce, o notável aumento, nos últimos anos, da emigração com origem na China para o continente africano, particularmente para Angola (132), que conta já com mais de 20 000 indivíduos residentes com origem nesse país asiático (129).

No que diz respeito às demais sequências sob análise, nomeadamente AO_319, AO_320, AO_364, AO_370, AO_394, AO_397, AO_417, AO_418, AO_442, AO_456, AO_469, AO_490, AO_493, AO_503, AO_512, AO_516, AO_522 e AO_545, observa-se que formam um agrupamento com diversas sequências de referência do subtipo C. No entanto, é relevante destacar que esse agrupamento não é apoiado por um valor de

bootstrap consistente, permanecendo abaixo do limite de 70%, à exceção das sequências AO_469 e AO_522 que se encontram num agrupamento monofilético com uma sequência de referência do subtipo C, suportado por um valor de *bootstrap* de 82%, bem como a sequência AO_397 pertencente a um agrupamento monofilético com uma sequência do subtipo C, apoiado por um valor de *bootstrap* de 75%. O grau de divergência dos ramos onde se encontram as sequências AO_469 e AO_522 na árvore filogenética consta-se de elevada significância, pelo que, segundo a análise, representa eventos de hipermutação induzida pela APOBEC. Tanto estas duas sequências, como a AO_364, seriam sequências que, dada a posição relativa dos ramos na árvore filogenética, faria sentido prosseguir para uma análise de recombinação, como passo futuro, não tendo sido realizada por razões já mencionadas. Não obstante, a topologia da árvore filogenética sugere uma relação filogenética evidente dessas sequências com o subtipo C, embora com divergência genética elevada (conforme ilustrado na Figura 17). O subtipo C é prevalente na maioria dos estudos citados por Sebastião *et al.* (2021) em Angola, nomeadamente, nas províncias da região sul e sudeste, tal como na região meridional de África (63) e, portanto, o aumento na prevalência deste subtipo desde 2009 poderá ser explicado pela sua circulação disseminada nos países que fazem fronteira com Angola, particularmente, Zâmbia, Namíbia e a região sul da RDC (66,67). Os resultados apresentados, no que diz respeito à elevada prevalência do subtipo C na nossa amostra, seriam expectáveis, dada a localização geográfica de Benguela. As sequências do subtipo C que formam agrupamentos com as sequências em estudo têm, na sua maioria, origem em estirpes de HIV-1 em circulação em países do sul e sudeste de África, como Zâmbia, África do Sul, Maláui e Botsuana, o que vai de encontro à literatura referida anteriormente.

Por último, é importante notar que para um conjunto de 12 sequências, nomeadamente AO_398, AO_434, AO_466, AO_474, AO_475, AO_480, AO_481, AO_494, AO_520, AO_527, AO_531 e AO_533, não foi possível atribuir nenhum genótipo, pelo facto de não se encontrarem incluídas em nenhum agrupamento monofilético significativo, de acordo com a inferência filogenética realizada. Consequentemente, essas sequências seriam candidatas a uma análise futura mais aprofundada, a fim de investigar possíveis fenómenos de divergência genética extrema e recombinação, não tendo sido realizada por razões já mencionadas. Sendo que duas delas (AO_474 e AO_533) estão associadas a fenómenos de hipermutação pela APOBEC. No geral, essas sequências são suportadas

por ramos precocemente divergentes, ou formam mesmo agrupamentos divergentes, em comparação com a topologia global da árvore filogenética construída, e, na sua maioria, apresentam distâncias genéticas relativas aos respectivos ancestrais (nós da árvore) substancialmente elevadas, conforme ilustrado na Figura 17.

Foi referido anteriormente que algumas árvores filogenéticas com sequências de HIV-1 com origem em Angola têm uma fraca organização subestrutural, sendo constituídas por agrupamentos menos consistentes do que os formados pelas sequências de referência globais, não permitindo, por isso, uma classificação inequívoca dos subtipos (61), como se pode observar na árvore filogenética construída neste estudo. Como consequência, há um grau de incerteza associado à classificação global dos subtipos neste estudo, que pode não refletir toda a extensão da diversidade genética nesta região. Assim, uma análise mais aprofundada será necessária, por exemplo, para clarificar se estas sequências são subtipos puros ou representam novas formas genéticas recombinantes do HIV-1. Apesar de tudo, estes resultados reforçam que Angola, tal como a RDC, é um dos epicentros de diversidade desta epidemia. Alguns subtipos com distribuição fora de África poderão ter resultado da disseminação de estirpes fundadoras, com origem na região de Angola, para outras regiões geográficas (e.g. subtipo G, em Portugal, depois na base também da CRF14_BG, com alguma disseminação europeia, particularmente, entre a população toxicodependente por via intravenosa). Os resultados publicados têm vindo a indicar que a complexidade genética da epidemia do HIV-1 em Angola continua a aumentar, sugerindo altas taxas de coinfeção e/ou superinfecção (66), o que é consistente com o aumento da incidência e prevalência destes vírus com genomas recombinantes.

3.4 Análise bioinformática das sequências aminoacídicas e pesquisa de mutações de resistência associadas aos inibidores da integrase

De forma a realizar a análise bioinformática das 92 sequências aminoacídicas traduzidas putativas e pesquisa das mutações associadas a resistência aos antirretrovirais inibidores da IN, utilizou-se o programa HIVdb, nomeadamente o *Genotypic Resistance Interpretation Algorithm da Stanford University HIV Drug Resistance Database* (versão 9.4) (disponível em: <https://hivdb.stanford.edu/hivdb/by-sequences/>), como foi referido no ponto 2.8.1. O objetivo principal seria fazer um levantamento das mutações associadas

a resistência aos inibidores da IN (mutações principais e acessórias), bem como de outros polimorfismos genéticos naturais. O programa utilizado gera um relatório interpretativo individual para cada uma das sequências nucleotídicas submetidas para análise, referentes à região codificante da IN, sendo os resultados obtidos por meio de comparação com uma sequência consenso do subtipo B.

Na sua grande maioria (90/92, 97,8%), as sequências aminoacídicas em análise correspondem a um fragmento de 227 dos 288 resíduos de aminoácidos que compõem a estrutura primária da IN (78,8%), ainda assim, englobando a quase totalidade das posições nas quais já foram descritas mutações associadas a resistência aos inibidores desta enzima (principais e acessórias) (posições 54-281). Utilizando o mesmo algoritmo, também foram reunidos os perfis fenotípicos previstos para cada um dos inibidores da IN considerados, em relação a cada sequência analisada. No relatório emitido, com base na mutação identificada, foram fornecidas pontuações individuais relativamente a cada INSTI. Em seguida, considerando a soma das pontuações de cada uma das mutações encontradas, em cada sequência submetida, foi atribuída uma classificação do nível fenotípico de resistência correspondente (susceptibilidade, potencial de nível baixo, nível baixo, intermédio ou elevado), descritas no subcapítulo 2.8.1. Essa classificação foi determinada para os cinco inibidores da IN (RAL, EVG, DTG, BIC e CAB). É importante realçar que o RNA genómico viral isolado a partir do plasma é considerado o material biológico preferencial para pesquisa de mutações de resistência aos ARVs (99,133,134). Contudo, no presente estudo, as sequências em análise foram obtidas a partir da amplificação do DNA proviral que se encontrava em células mononucleadas do sangue periférico infetadas, existentes na mancha de sangue preservada nos cartões FTA[®] (GE Healthcare, Reino Unido). Desta forma, o vírus representado pelo DNA proviral pode não ser inteiramente representativo da população de vírus replicativos, em circulação na corrente sanguínea, podendo representar apenas populações virais latentes e sequências “arquivadas”. A utilização, para este tipo de testes, de RNA genómico viral presente no plasma, permite a identificação preferencial das mutações que conferem resistência ao regime terapêutico presente e que poderão estar na base de uma eventual falência terapêutica (134). Outro tipo de resistência, como reflexo de falhas terapêuticas anteriores ou mutações associadas a TDR, pode encontrar-se no DNA proviral, em células mononucleadas de sangue periférico, sendo que essas variantes virais que se encontram

latentes, podem reemergir sob a pressão seletiva do fármaco (134). Foi demonstrado que a resistência aos fármacos no reservatório proviral prevê a recuperação virológica (135), ainda que a avaliação de resultados desta natureza requeira algum cuidado, já que as mutações encontradas nem sempre correspondem às encontradas no RNA genómico do vírus (embora alguns estudos indiquem valores de concordância entre os resultados obtidos com o RNA viral e o DNA proviral que variam entre 28% e 100% (134)). Independentemente do tipo de amostra biológica utilizada, existe uma outra limitação para estes testes, especialmente quando implementados através do método clássico de sequenciação de Sanger, que é o facto de poderem existir estirpes minoritárias que possuem mutações associadas à resistência aos fármacos antirretrovirais no conjunto da *quasispecies* que não são caracterizadas, isto devido ao limite de deteção (15–20%) dos testes genotípicos empregues para rastrear essas mutações de resistência (99,134,136). Apesar destas e outras limitações, os testes genotípicos são preferencialmente utilizados devido ao seu custo reduzido e facilidade e rapidez de implementação (99).

3.4.1 Identificação de mutações associadas a resistência aos INSTI e outros polimorfismos genéticos

De entre o conjunto de sequências aminoacídicas em análise foram encontradas cinco mutações associadas a uma diminuição da suscetibilidade aos inibidores da IN, num total de 10 sequências: uma mutação principal, E92G (1,2%) e quatro mutações acessórias, L74M (3,6%), Q95K (2,4%), T97A (3,6%) e E157Q (1,2%), como se pode observar na Figura 18, sendo que cada uma das sequências apresentava apenas uma das mutações representadas.

É importante realçar que nesta análise foram excluídas 9/92 (9,8%) sequências em estudo (AO_205, AO_356, AO_363, AO_469, AO_474, AO_482, AO_491, AO_522 e AO_533), por apresentarem indícios de hipermutação por hiperadenilação associada à atividade da APOBEC (137,138). Esta é uma enzima do hospedeiro humano a que está associado um fenómeno de hiperadenilação do ácido nucleico do vírus (C→A), com efeitos mutagénicos (138,139). A introdução destas mutações, nomeadamente conduzindo a codões de terminação, ao longo da sequência codificante da IN, conduz a uma enzima defetiva. Uma vez que esta desempenha um papel fundamental no ciclo

replicativo do vírus, é possível que nem se cheguem a formar quaisquer novas partículas virais, a partir destes provírus contendo regiões codificantes da IN com evidência de hiperadenilação, dependendo da localização na grelha de leitura aberta do gene *pol* do primeiro codão de terminação introduzido.

a)

L	E	Q	T	E
74	92	95	97	157
M	G	K	A	Q

b)

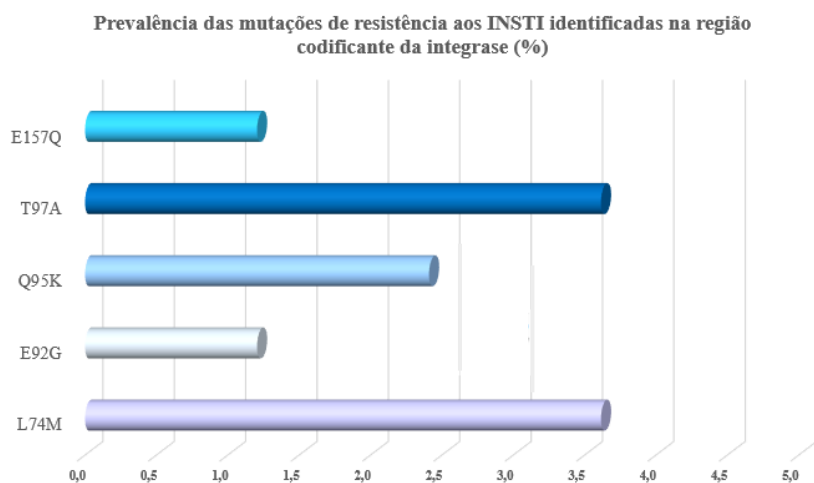


Figura 18. a) Representação esquemática de todas as mutações encontradas associadas a resistência aos INSTI, na sequência codificante da IN. As letras superiores representam o aminoácido (código de uma letra) da sequência consenso do subtipo B (susceptibilidade), o número refere-se à sua posição na sequência e as letras inferiores indicam o aminoácido associado à mutação que confere resistência aos INSTI. A negrito está destacada a mutação principal, enquanto as restantes são mutações acessórias. b) Representação gráfica da prevalência das mutações associadas a resistência aos INSTI identificadas na região codificante da IN das sequências em estudo.

É de notar, também, que 60 posições aminoacídicas da sequência primária da IN não foram analisadas neste estudo (posições 1-53 e 282-288), sendo uma delas, a posição 51, associada a uma mutação de resistência aos INSTI (H51Y, uma mutação acessória pouco comum e não polimórfica).

As sequências AO_391, AO_434 e AO_472 apresentaram a mutação L74M. Nas sequências AO_448, AO_473 e AO_480, foi constatada a presença da mutação T97A. Ambas são mutações acessórias, polimórficas, associadas a resistência aos INSTI, sendo a primeira delas comum. Segundo os dados da *Stanford University HIV Drug Resistance Database*, apresentam uma prevalência entre 1-5% em pessoas que nunca receberam HAART, dependendo do subtipo viral. Por si só, não reduzem a susceptibilidade aos

INSTI, contudo, em conjugação com outras mutações de resistência principais, contribuem para a diminuição da suscetibilidade (140).

Em relação às sequências AO_456 e AO_520, identificou-se a mutação Q95K, que é uma mutação acessória, não polimórfica, e que, sozinha, contribui pouco, ou nada, para a diminuição da suscetibilidade aos INSTI (140). A amostra AO_404 exibiu a mutação E157Q, sendo esta uma mutação polimórfica, também acessória, selecionada em indivíduos sob terapia incluindo EVG e RAL, mas que parece ter também pouco efeito na redução da suscetibilidade aos INSTI (140).

Por último, a mutação E92G foi detetada exclusivamente na sequência AO_416. E92G é uma mutação principal, não polimórfica, que reduz a suscetibilidade ao EVG mais de 10 vezes, mas não parecendo contribuir para uma redução da suscetibilidade aos outros INSTI (140). Esta é, também, uma mutação relevante para a vigilância epidemiológica da resistência transmitida aos ARVs no HIV-1 (SDRM) (110,141), tendo sido selecionada *in vivo* em indivíduos sob tratamento com INSTI (141). As SDRMs, como mencionado anteriormente, são uma lista de mutações que cumprem os critérios necessários para comparar com exatidão as taxas de resistência transmitida aos ARVs entre regiões geográficas e períodos de tempo distintos: 1) são reconhecidas como DRMs por um grupo de peritos em resistência aos fármacos contra o HIV; 2) não são polimórficas, independentemente do subtipo viral; e 3) não são excessivamente raras (ou seja, foram excluídas as mutações que resultam muito raramente da pressão dos fármacos) (141).

É de notar que na população angolana, tanto quanto é do nosso conhecimento, ainda não foram conduzidas pesquisas de resistência aos INSTI envolvendo grupos de indivíduos nunca submetidos a terapia antirretroviral com base neste grupo de ARVs, considerando até a recente introdução do DTG nos esquemas terapêuticos do país. Portanto, os resultados aqui apresentados, indicando a presença de uma mutação principal e quatro mutações acessórias, associadas a uma redução na suscetibilidade aos INSTI, não foram, até à data, previamente reportados nesta mesma população. Contudo, foram realizados estudos noutras regiões (e.g. Marrocos, Rússia, Suíça), nos quais já foram descritas estas mutações, como por exemplo a L74M, T97A e E157Q, surgindo em proporções semelhantes às relatadas no presente estudo (142–146).

Com recurso ao algoritmo da *Sandford University HIV Drug Resistance Database*, a cada mutação foi atribuída uma pontuação, associada a cada inibidor da IN, por forma a estimar a probabilidade de resistência fenotípica para cada uma das sequências em estudo.

A grande maioria das sequências analisadas apresentava total suscetibilidade aos diferentes INSTI (76/83; 91,6%), enquanto para 7 sequências (7/83; 8,4%) foi previsto algum nível associado de resistência fenotípica (Figura 19).

Para a mutação E92G (na sequência AO_416) foram estimados níveis intermédio, baixo e potencial de nível baixo de resistência ao EVG, RAL e CAB, respetivamente. Como referido anteriormente, uma mutação de resistência de nível intermédio indica uma elevada probabilidade de que a atividade do fármaco, neste caso o EVG, seja reduzida, mas que este irá, provavelmente, reter uma significativa atividade antiviral remanescente (147). A resistência de nível baixo, associada a esta mutação, indica que o vírus pode ter uma suscetibilidade reduzida ao RAL *in vitro*, ou que o indivíduo infetado com este vírus pode ter uma resposta virológica subótima ao tratamento com este ARV (147). Por fim, uma mutação de resistência potencial de nível baixo indica uma possível exposição anterior do vírus a ARVs, sendo que, pelas características da população angolana em estudo, estaríamos perante uma mutação de resistência transmitida (147). Relativamente às mutações acessórias Q95K, T97A e E157Q, encontradas nas restantes seis sequências (AO_404, AO_448, AO_456, AO_473, AO_480 e AO_520), foram associados perfis fenotípicos de resistência potencial de nível baixo para o EVG e RAL. A mutação L74M, presente nas sequências AO_391, AO_434 e AO_472, foi associada a perfis fenotípicos de suscetibilidade. É notável observar a completa suscetibilidade, das sequências em estudo, manifestada em relação ao BIC e ao DTG, especialmente considerando a introdução recente, deste último, nos esquemas terapêuticos em Angola (2021) (91). Este resultado parece indicar uma perspetiva promissora em termos de eficácia terapêutica, uma vez que a suscetibilidade abrangente a este ARV pode sugerir a viabilidade do recurso ao mesmo como componente essencial nos regimes de tratamento. Esta observação pode ser relevante para a definição de diretrizes e planeamento de esquemas terapêuticos no controlo da disseminação do HIV-1 em Angola.

Neste estudo, a mutação L74M foi descrita em sequências classificadas nos subtipos F1 (AO_391 e AO_472) e G (AO_434). Esta mutação foi mencionada num estudo de 2011,

no qual se verificou a sua seleção, bem como a de outras mutações acessórias, como T97A e E157Q, devido à pressão seletiva associada à terapia com RAL, podendo contribuir para a falha na terapia com recurso a esse ARV (148). Esta mutação foi observada em vírus dos subtipos A1, C, D, F1, G e CRF02_AG, a níveis inferiores a 10% (149), o que vai de encontro ao observado no presente estudo, em que a mesma se encontra, com uma baixa prevalência, associadas aos subtipos F1 e G. Num estudo realizado em 2016, esta mutação foi observada maioritariamente em vírus pertencentes ao subtipo C (150). Relativamente às mutações E157Q e E92G, estas estão presentes nas sequências AO_404 e AO_416, respetivamente, classificadas como subtipo A. A mutação E157Q foi observada em 35% dos indivíduos infetados que nunca foram submetidos a HAART infetados com o subtipo C (148). Num estudo de 2016, foi também observada a presença desta mutação em indivíduos infetados com este subtipo (149), o que não se observa no presente trabalho. E157Q tem sido, ainda, reportada em vírus do subtipo B (150,151) e CRF02_AG (150), o que não parece ir de encontro aos resultados obtidos. É, no entanto, importante referir novamente que apenas foi estudada a região codificante da IN e que não foi feita uma análise de recombinação, pelo que a comparação entre resultados pode não ser inteiramente significativa. Segundo a literatura, a prevalência da mutação E92G varia substancialmente, tendo sido relatada associada a diferentes subtipos, com diferentes prevalências, como por exemplo, o subtipo F (18,5%) (152) e a CRF02_AG (0,1%) (141). Em 2018, verificou-se uma associação entre indivíduos infetados com vírus do subtipo A e a presença desta mutação com uma prevalência de 1,7% (153), na África subsariana. Este resultado vai de encontro ao observado no presente trabalho, que apresenta uma percentagem semelhante. Quanto à mutação Q95K, esta encontrava-se presente nas amostras AO_456 e AO_520, classificadas nos subtipos C e G, respetivamente. Num estudo realizado em 2016, foi observada esta mutação, maioritariamente, em vírus pertencentes ao subtipo C (150). Por fim, as sequências AO_448, AO_473 e AO_480, que apresentavam a mutação T97A, foram classificadas como sendo CRF27_cpx, CRF02_AG e NA (genótipo não atribuído), respetivamente. Dados anteriores, associam esta mutação, a níveis abaixo de 10%, a vírus pertencentes à CRF02_AG (149).

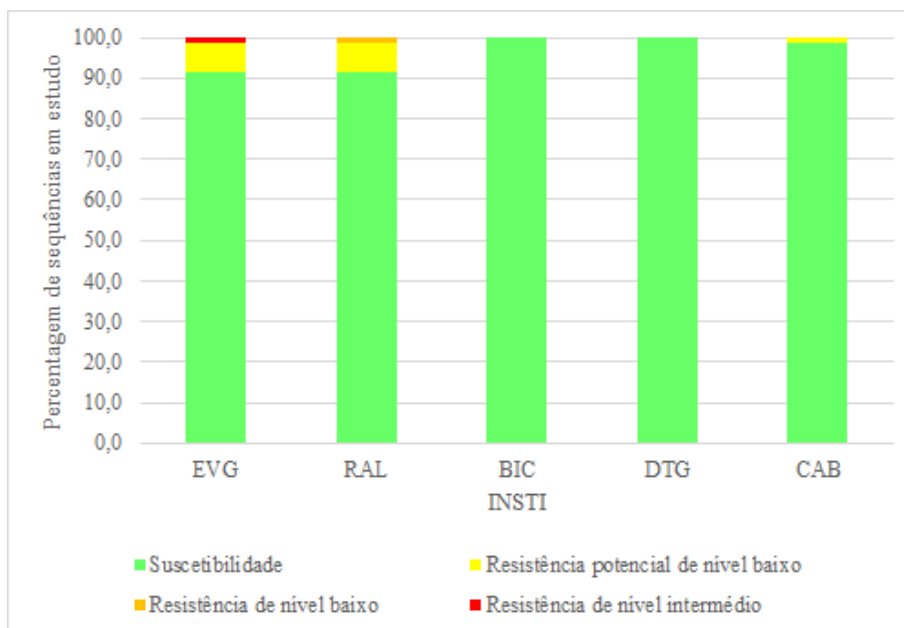


Figura 19. Representação gráfica da distribuição relativa dos níveis de resistência fenotípicos putativos aos cinco INSTI: Elvitegravir (EVG), Raltegravir (RAL), Bictegravir (BIC), Dolutegravir (DTG) e Cabotegravir (CAB); associados às sequências da região codificante da IN analisadas.

De nota que a sequência AO_373 apresentou uma inserção na posição 253, com ocorrência de *frameshift*, contudo, não demonstrando afetar o nível de suscetibilidade a nenhum dos INSTI.

Finalmente, procedeu-se a uma análise dos polimorfismos genéticos não associados a diminuição da suscetibilidade aos INSTI, presentes nas 92 sequências em estudo, segundo o algoritmo da *Stanford University HIV Drug Resistance Database*. Foi também realizado um alinhamento múltiplo das sequências aminoacídicas putativas em análise, tendo como referência uma sequência consenso do subtipo B, o que permitiu uma mais fácil comparação entre sequências. Após recolha de todos os polimorfismos genéticos a partir do alinhamento múltiplo, foi realizada uma análise estatística (como descrito na secção 2.10), com o propósito de analisar uma potencial associação entre os polimorfismos genéticos identificados em cada posição e os diferentes genótipos.

Alguns polimorfismos genéticos demonstraram uma associação com significado estatístico ($p < 0,01$) com os subtipos A, C e F1 (Anexo 6). Estes mesmos polimorfismos foram, então, considerados como assinaturas genéticas de subtipo. Para o subtipo A foram identificados três destas assinaturas, sendo elas I59V, S65P e D113E. Para o subtipo C, cinco, nomeadamente, A37E, F46Y, S65T, C162H e D225A, apresentado duas delas

(F46Y e D225A) uma associação altamente significativa ($p < 0,00001$). Por fim, para o subtipo F1 foram identificadas sete assinaturas genéticas, A44G, T58I, K82C, A151S, T164I, L180V e D203E, com três delas (T58I, L180V e D203E) com uma associação altamente significativa ($p < 0,00001$) (Anexo 6 e Tabela 7). É importante referir que a importância destes polimorfismos genéticos ainda se apresenta controversa, uma vez que estes apresentam um impacto nulo, a nível fenotípico, na resistência aos ARVs. Contudo, já foi demonstrada uma eventual importância clínica dos mesmos, no que diz respeito à seleção de mutações de resistência, que se verificou ser acompanhada por um aumento da prevalência de polimorfismos específicos que ocorrem naturalmente no vírus, contribuindo para uma seleção mais rápida ou mais lenta de DRM (154). Para além disto, estes polimorfismos genéticos associados de forma significativa a determinados subtipos, podem ser usados como marcadores epidemiológicos (155).

Tabela 7. Assinaturas genéticas identificadas nas sequências da região codificante da IN em estudo pertencentes aos subtipos A, C e F1 ($p < 0,01$). Os polimorfismos assinalados a negrito indicam uma associação altamente significativa ao respetivo genótipo ($p < 0,00001$). O código internacional IUPAC de uma letra foi utilizado para representar os aminoácidos em cada posição (Anexo 4).

Subtipo	Vírus Selvagem (sequência consenso do subtipo B)	Posição aminoacídica	Vírus mutado
A	I	59	V
	S	65	P
	D	113	E
C	A	37	E
	F	46	Y
	S	65	T
	C	162	H
	D	225	A
F1	A	44	G
	T	58	I
	K	82	C
	A	151	S
	T	164	I
	L	180	V
	D	203	E

4. Conclusões

A pandemia do HIV-1 continua a representar um desafio significativo para a saúde global, mas são de notar os avanços realizados a nível da prevenção, tratamento e monitorização desta infeção viral. No entanto, apesar de todos os progressos nos últimos anos, a infeção pelo HIV-1 permanece uma ameaça persistente, especialmente na região africana.

O HIV-1 apresenta, cada vez mais, um aumento na diversidade genética e complexidade das formas virais circulantes, observando-se, ao longo dos tempos, um aumento na prevalência das formas virais recombinantes, acompanhado de uma diminuição na prevalência dos subtipos puros (23,30,32). Esta complexidade genética é bastante significativa em Angola, país localizado na região geográfica epicentro desta complexidade. O presente estudo permitiu verificar, exatamente, essa grande diversidade genética de formas virais circulantes, para a região codificante da IN, no gene *pol*, nomeadamente, na província de Benguela, em Angola. É importante ressaltar a limitação de dados anteriores, presentes na literatura, disponíveis referentes a esta região. Os resultados obtidos confirmam uma maior prevalência do subtipo F1 (28,3%), seguido dos subtipos A (20,7%) e C (19,6%). Foram ainda encontrados a CRF27_cpx (4,3%), o subtipo G (3,3%), a CRF02_AG (3,3%), a CRF19_cpx (3,3%), o subtipo D (2,2%), o subtipo H (1,1%) e a CRF18_cpx (1,1%). Nesta região (sul de Angola), seria razoável antecipar uma maior prevalência do subtipo C, no entanto, o predomínio conjunto daqueles três genótipos virais vai de encontro ao descrito na literatura para esta região. Para além disso, é imperativo reiterar que os subtipos atribuídos referem-se apenas à região codificante da IN, não abrangendo o genoma total. Entre as formas virais encontradas, foi constatada uma relativa diversidade de formas recombinantes (CRFs ou URFs) para a região em estudo. Uma vez que, por razões temporais, não houve oportunidade de implementar uma análise de recombinação, por técnica de *bootscanning*, é plausível que a contagem atual de formas recombinantes possa subestimar a verdadeira incidência de eventos de recombinação genómica nesta população viral.

Uma das principais conquistas no combate à pandemia pelo HIV-1 tem sido o desenvolvimento e implementação mais generalizada da HAART, o que permitiu a transição de um estado de doença quase sempre fatal para uma condição crónica controlável. A terapia antirretroviral não melhora apenas a qualidade de vida das pessoas

infetadas, como também reduz significativamente a transmissão vertical e sexual, contribuindo assim também, em última instância, para a redução do impacto desta infecção ao nível dos sistemas de saúde, à escala global. No entanto, a acessibilidade a estes fármacos, a adesão terapêutica individual e a implementação de uma monitorização sistemática, com recurso a testes de carga viral e genotípicos/fenotípicos de resistência, constituem um grande desafio em muitas partes do mundo, nomeadamente, no continente africano. No âmbito deste estudo, observou-se um grau de resistência primária aos INSTI não negligenciável, tendo-se reportado a circulação na província de Benguela de estirpes virais com as mutações associadas à diminuição da suscetibilidade aos INSTI L74M, Q95K, T97A, E157Q e E92G, sendo esta última uma mutação principal. No cômputo geral 12,0% (10/83) das sequências estudadas continham mutações associadas à diminuição da suscetibilidade aos INSTI, 8,4% (7/83) associadas a níveis fenotípicos de resistência. Dada a introdução tardia destes ARVs nos regimes de HAART em Angola, nomeadamente o DTG (91) como esquema de primeira linha (93), a presença considerável de níveis de resistência aos INSTI em indivíduos que ainda não foram submetidos a terapia, apesar de não ser extremamente alarmante, sugere, de forma indicativa, a existência de resistência aos ARVs transmitida. Além disso, o grau de resistência manifestado neste estudo, em relação ao DTG, pode fornecer uma estimativa da probabilidade de sucesso da HAART que inclua este ARV.

Apesar de todos os progressos dos últimos anos, continuam a persistir desafios significativos. A acessibilidade universal a tratamentos eficazes e preventivos ainda está longe de ser ideal, especialmente em regiões com recursos limitados, como os países de médio/baixo rendimento. Adicionalmente, a falta de uma monitorização terapêutica sistematizada, como observada na maioria da população infetada, nomeadamente nestes países, representa um desafio significativo, podendo contribuir para o desenvolvimento de resistência aos ARVs, tanto primária, como transmitida. A resistência aos antirretrovirais continua a ser uma preocupação, destacando-se a necessidade de desenvolvimento contínuo de novos fármacos. A estigmatização social relacionada com a infecção pelo HIV-1 também permanece como um obstáculo, que pode resultar em barreiras significativas para a prevenção, testagem e tratamento. Assim, uma abordagem abrangente para o HIV-1 não inclui apenas avanços científicos, mas exige também esforços para combater o estigma e promover a educação pública.

É imperativo salientar que este estudo apresenta algumas limitações, incluindo o número mediano de amostras analisadas (n=92) e a relativamente baixa taxa de sucesso de amplificação da região codificante da IN (42,9%). As razões subjacentes a essas limitações foram previamente explanadas e estão associadas, principalmente, a desafios enfrentados durante a fase inicial de otimização do protocolo laboratorial para a obtenção dos fragmentos de DNA a serem analisados por sequenciação nucleotídica de Sanger. Adicionalmente, é importante observar que os resultados obtidos não incluem informações sobre eventuais estirpes virais minoritárias em circulação em cada um dos indivíduos da amostra populacional estudada, devido às características do método de genotipagem (sequenciação de Sanger) e pesquisa de mutações de resistência aos INSTI utilizados, que não possibilita a sua detecção. Por fim, foram impostas restrições temporais para a conclusão da parte laboratorial desta dissertação, que resultaram na exclusão de algumas amostras e procedimentos (e.g. análise de recombinação), a fim de concluir o estudo dentro dos prazos estipulados.

O estado atual da pandemia pelo HIV-1 reflete um quadro complexo de avanços promissores e desafios persistentes. Enquanto a evolução constante da eficácia das terapias antirretrovirais e estratégias preventivas, como a PrEP, oferecem esperança, a equidade global no acesso a essas intervenções e a luta contra o estigma continuam a ser cruciais. A colaboração internacional, investimentos contínuos em pesquisa e ações coordenadas a nível global, nacional e comunitário são essenciais para superar os obstáculos que ainda permanecem e alcançar o objetivo de pôr fim à pandemia até 2030 (96).

Para concluir, os resultados obtidos neste trabalho reforçam a importância da monitorização da emergência de mutações de resistência primária aos diferentes ARVs e sua transmissão em Angola, considerando o seu impacto direto na efetividade terapêutica e conseqüentemente controlo da infeção.

5. Referências bibliográficas

1. Vahlne A. A historical reflection on the discovery of human retroviruses. *Retrovirology*. 2009 May 1;6.
2. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. The early spread and epidemic ignition of HIV-1 in human populations. *Science* (1979). 2014 Oct 3;346(6205):56–61.
3. Barré-Sinoussi F, Chermann J, Rey F, Nugeyre M, Chamaret S, Gruest J, et al. Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science* (1979) [Internet]. 1983;220(4599):868–71. Available from: <http://science.sciencemag.org/>
4. Hemelaar J. The origin and diversity of the HIV-1 pandemic. *Trends Mol Med*. 2012 Mar;18(3):182–92.
5. Sharp M. Paul, Hahn H. Beatrice. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med*. 2011;1(1).
6. Sharp M. Paul, Hahn H. Beatrice. The evolution of HIV-1 and the origin of AIDS. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2010 Aug 27;365(1552):2487–94.
7. Nagata S, Imai J, Makino G, Tomita M, Kanai A. Evolutionary analysis of HIV-1 pol proteins reveals representative residues for viral subtype differentiation. *Front Microbiol*. 2017 Nov 2;8.
8. Vidal N, Peeters M, Mulanga-kabeya C, Nzilambi N, Robertson D, Ilunga W, et al. Unprecedented Degree of Human Immunodeficiency Virus Type 1 (HIV-1) Group M Genetic Diversity in the Democratic Republic of Congo Suggests that the HIV-1 Pandemic Originated in Central Africa. *J Virol* [Internet]. 2000;74(22):10498–507. Available from: <http://igs-server.cnrs-mrs.fr>
9. National Center for Biotechnology Information. Taxonomy [Internet]. [cited 2023 Feb 23]. Available from: <https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=11676>

10. National Library of Medicine. Taxonomy Browser [Internet]. [cited 2023 Feb 23]. Available from: <https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/tree/?taxon=11676>
11. Maertens GN, Engelman AN, Cherepanov P. Structure and function of retroviral integrase. *Nat Rev Microbiol.* 2022 Jan 1;20(1):20–34.
12. ViralZone. Human immunodeficiency virus 1 [Internet]. [cited 2023 Feb 23]. Available from: https://viralzone.expasy.org/7?outline=all_by_species
13. Freed EO. HIV-1 assembly, release and maturation. *Nat Rev Microbiol.* 2015 Jul 16;13(8):484–96.
14. Moore MD, Hu WS. HIV-1 RNA Dimerization: It Takes Two to Tango. *AIDS.* 2009 Dec 14;11(2):91–102.
15. National Library of Medicine. Human immunodeficiency virus 1 [Internet]. [cited 2023 Feb 23]. Available from: <https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/11676/>
16. Sükösd Z, Andersen ES, Seemann SE, Jensen MK, Hansen M, Gorodkin J, et al. Full-length RNA structure prediction of the HIV-1 genome reveals a conserved core domain. *Nucleic Acids Res.* 2015;43(21):10168–79.
17. Science Direct. Gag Protein [Internet]. [cited 2023 Feb 23]. Available from: <https://www.sciencedirect.com/topics/medicine-and-dentistry/gag-protein>
18. Engelman A, Cherepanov P. Retroviral Integrase Structure and DNA Recombination Mechanism. *Microbiol Spectr.* 2014;2(6):1–22.
19. Science Direct. Pol Protein [Internet]. [cited 2023 Feb 23]. Available from: <https://www.sciencedirect.com/topics/medicine-and-dentistry/pol-protein>
20. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature.* 1988;331:280–3.
21. Qiao Q, Yan Y, Guo J, Du S, Zhang J, Jia R, et al. A review on architecture of the gag-pol ribosomal frameshifting RNA in human immunodeficiency virus: a

- variability survey of virus genotypes. *J Biomol Struct Dyn*. 2017 Jun 11;35(8):1629–53.
22. Science Direct. *Env (Gene)* [Internet]. [cited 2023 Feb 23]. Available from: <https://www.sciencedirect.com/topics/immunology-and-microbiology/env-gene>
 23. Giovanetti M, Ciccozzi M, Parolin C, Borsetti A. Molecular Epidemiology of HIV-1 in African Countries:A Comprehensive Overview. *Pathogens*. 2020 Dec;9(12):1072.
 24. Alessandri-Gradt E, Moisan A, Plantier JC. HIV-1 Non-Group M Strains and ART. *Viruses*. 2023 Mar 1;15(3).
 25. Hemelaar J, Elangovan R, Yun J, Dickson-Tetteh L, Fleminger I, Kirtley S, et al. Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *Lancet Infect Dis*. 2019 Feb 1;19(2):143–55.
 26. Abidi SH, Aibekova L, Davlidova S, Amangeldiyeva A, Foley B, Ali S. Origin and evolution of HIV-1 subtype A6. *PLoS One*. 2021 Dec 13;16(12).
 27. Désiré N, Cerutti L, Le Hingrat Q, Perrier M, Emler S, Calvez V, et al. Characterization update of HIV-1 M subtypes diversity and proposal for subtypes A and D sub-subtypes reclassification. *Retrovirology*. 2018 Dec 22;15(1):80.
 28. Yamaguchi J, Vallari A, McArthur C, Sthreshley L, Cloherty GA, Berg MG, et al. Complete Genome Sequence of CG-0018a-01 Establishes HIV-1 Subtype L. *J Acquir Immune Defic Syndr (1988)* [Internet]. 2020 Mar 1;83(3):319–22. Available from: www.hiv.lanl.gov
 29. Mendes Da Silva RK, Monteiro de Pina Araujo II, Venegas Maciera K, Gonçalves Morgado M, Lindenmeyer Guimarães M. Genetic characterization of a new hiv-1 sub-subtype a in cabo verde, denominated A8. *Viruses*. 2021 Jun 1;13(6):1093.
 30. Abongwa LE, Nyamache AK, Torimiro JN, Okemo P, Charles F. Human immunodeficiency virus type 1 ((HIV-1) subtypes in the northwest region, Cameroon. *Virol J* [Internet]. 2019 Aug 15 [cited 2023 Feb 14];16(1):103. Available from: <https://pubmed.ncbi.nlm.nih.gov/31416460/>

31. Yongjian L, Lei J, Bin S, Hanping L, Zhen L, Jingwan H, et al. The Genetic Diversity of HIV-1 Quasispecies Within Primary Infected Individuals. *AIDS Research and Human Retroviruses* . 2020 May;36(5):440–9.
32. Bbosa N, Kaleebu P, Ssemwanga D. HIV subtype diversity worldwide. *Curr Opin HIV AIDS*. 2019 May 1;14(3):153–60.
33. HIV sequence database. HIV Circulating Recombinant Forms (CRFs) [Internet]. 2022 [cited 2023 Feb 23]. Available from: <https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html>
34. Jumare J, Ndembu N, El-Kamary SS, Magder L, Hungerford L, Burdo T, et al. Cognitive Function Among Antiretroviral Treatment-Naive Individuals Infected With Human Immunodeficiency Virus Type 1 Subtype G Versus CRF02_AG in Nigeria. *Clinical infectious diseases*. 2018 Apr 17;66(9):1448–53.
35. Young-Keol C, Jung-Eun K, Sun-Hee L, Brian T F, Byeong-Sun C. Impact of HIV-1 subtype and Korean Red Ginseng on AIDS progression: comparison of subtype B and subtype D. *J Ginseng Res*. 2018 Apr;43(2):312–8.
36. Farinre O, Gounder K, Reddy T, Tongo M, Hare J, Chaplin B, et al. Subtype-specific differences in Gag-protease replication capacity of HIV-1 isolates from East and West Africa. *Retrovirology*. 2021 Dec 1;18(1).
37. Tongo M, de Oliveira T, Martin DP. Patterns of genomic site inheritance in HIV-1M inter-subtype recombinants delineate the most likely genomic sites of subtype-specific adaptation. *Virus Evol*. 2018 Jan 1;4(1).
38. Kanki PJ, Hamel DJ, Sankalé JL, Hsieh CC, Thior I, Barin F, et al. Human Immunodeficiency Virus Type 1 Subtypes Differ in Disease Progression. *J Infect Dis* [Internet]. 1999 Jan;179(1):68–73. Available from: <https://academic.oup.com/jid/article/179/1/68/883777>
39. Kiguoya MW, Mann JK, Chopera D, Gounder K, Lee GQ, Hunt PW, et al. Subtype-Specific Differences in Gag-Protease-Driven Replication Capacity Are Consistent with Intersubtype Differences in HIV-1 Disease Progression. *J Virol*. 2017 Jul;91(13).

40. Iordanskiy S, Waltke M, Feng Y, Wood C. Subtype-associated differences in HIV-1 reverse transcription affect the viral replication. *Retrovirology*. 2010 Oct 12;7.
41. Ye J, Chen J, Wang J, Wang Y, Xing H, Yu F, et al. CRF07_BC is associated with slow HIV disease progression in Chinese patients. *Sci Rep*. 2022 Dec 1;12(1).
42. Ng OT, Laeyendecker O, Redd AD, Munshaw S, Grabowski MK, Paquet AC, et al. HIV type 1 polymerase gene polymorphisms are associated with phenotypic differences in replication capacity and disease progression. *Journal of Infectious Diseases*. 2014 Jan 1;209(1):66–73.
43. Jumare J, Ndembi N, El-Kamary SS, Magder L, Hungerford L, Burdo T, et al. Cognitive Function Among Antiretroviral Treatment-Naive Individuals Infected With Human Immunodeficiency Virus Type 1 Subtype G Versus CRF02_AG in Nigeria. *Clinical infectious diseases*. 2018 Apr 17;66(9):1448–53.
44. Vasylyeva TI, Du Plessis L, Pineda-Peña AC, Kühnert D, Lemey P, Vandamme AM, et al. Tracing the Impact of Public Health Interventions on HIV-1 Transmission in Portugal Using Molecular Epidemiology. *Journal of Infectious Diseases*. 2019 Jun 19;220(2):233–43.
45. Cheng Z, Yan H, Li Q, Ablan SD, Kleinpeter A, Freed EO, et al. Enhanced Transmissibility and Decreased Virulence of HIV-1 CRF07_BC May Explain Its Rapid Expansion in China. *Microbiol Spectr*. 2022 Aug 31;10(4).
46. Martin A, Moritz C, Car-Son H, Strathmann R, Hadfield M. What is a Quasispecies. *TREE*. 1992;7(4):118–21.
47. Arts EJ, Hazuda DJ. HIV-1 antiretroviral drug therapy. *Cold Spring Harb Perspect Med*. 2012;2(4).
48. Bale MJ, Kearney MF. Review: HIV-1 phylogeny during suppressive antiretroviral therapy. *Curr Opin HIV AIDS*. 2019 May 1;14(3):188–93.
49. Alteri C, Fabeni L, Scutari R, Berno G, Di Carlo D, Gori C, et al. Genetic divergence of HIV-1 B subtype in Italy over the years 2003–2016 and impact on CTL escape prevalence. *Sci Rep*. 2018 Dec 1;8(1).

50. Rolland M. HIV-1 transmission and phylogenetics. *Curr Opin HIV AIDS*. 2019 May 1;14(3):151–2.
51. Esteves A, Parreira R, Venenno T, Franco M, Piedade J, Germano De Sousa J, et al. Molecular Epidemiology of HIV Type 1 Infection in Portugal: High Prevalence of Non-B Subtypes. *AIDS Res Hum Retroviruses*. 2002;18(5):313–25.
52. Pimentel V, Pingarilho M, Alves D, Diogo I, Fernandes S, Miranda M, et al. Molecular epidemiology of hiv-1 infected migrants followed up in Portugal: Trends between 2001-2017. *Viruses*. 2020;12(3).
53. Souto B, Triunfante V, Santos-Pereira A, Martins J, Araújo PMM, Osório NS. Evolutionary dynamics of HIV-1 subtype C in Brazil. *Sci Rep*. 2021 Dec 1;11(1).
54. Crispim MAE, Reis MN da G, Stefani MM de A. HIV-1 genetic diversity and drug resistance mutations in the northern Brazilian region. *Brazilian Journal of Infectious Diseases*. 2021 May 1;25(3).
55. Wagner T, Zuckerman NS, Halperin T, Chemtob D, Levy I, Elbirt D, et al. Epidemiology and Transmitted HIV-1 Drug Resistance among Treatment-Naïve Individuals in Israel, 2010–2018. *Viruses*. 2022 Jan 1;14(1).
56. World Health Organization. HIV/AIDS Overview. 2018 [cited 2024 Apr 1]. *Who Africa*. Available from: <https://www.afro.who.int/health-topics/hivaids>
57. Amer AN, Gaballah A, Emad R, Ghazal A, Attia N. Molecular Epidemiology of HIV-1 Virus in Egypt: A Major Change in the Circulating Subtypes. *Curr HIV Res*. 2021 Aug 6;19(5):448–56.
58. Rubio-Garrido M, González-Alba JM, Reina G, Ndarabu A, Barquín D, Carlos S, et al. Current and historic HIV-1 molecular epidemiology in paediatric and adult population from Kinshasa in the Democratic Republic of Congo. *Sci Rep*. 2020 Dec 1;10(1).
59. Lumngwena EN, Shuping L, Bernitz N, Woodman Z. HIV-1 subtype C Envelope function becomes less sensitive to N-glycosylation deletion during disease progression. *BMC Res Notes*. 2019 Jun 17;12(1).

60. Faria NR, Vidal N, Lourenco J, Raghwani J, Sigaloff KCE, Tatem AJ, et al. Distinct rates and patterns of spread of the major HIV-1 subtypes in Central and East Africa. *PLoS Pathog.* 2019;15(12).
61. 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. *Infection, Genetics and Evolution.* 2009 Jul;9(4):672–82.
62. Abecasis A, Paraskevis D, Epalanga M, Fonseca M, Burity F, Bartolomeu J, et al. HIV-1 genetic variants circulation in the North of Angola. *Infection, Genetics and Evolution.* 2005 Apr;5(3):231–7.
63. Bártolo I, Epalanga M, Bartolomeu J, Fonseca M, Mendes A, Gama A, et al. Sequence Note High Genetic Diversity of Human Immunodeficiency Virus Type 1 in Angola. *AIDS Res Hum Retroviruses.* 2005;21(4):306–10.
64. 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. *Journal of Antimicrobial Chemotherapy.* 2008 Mar;61(3):694–8.
65. Bártolo I, Rocha C, Bartolomeu J, Gama A, Fonseca M, Mendes A, et al. Antiretroviral Drug Resistance Surveillance among Treatment-Naive Human Immunodeficiency Virus Type 1 Infected Individuals in Angola: Evidence for Low Level of Transmitted Drug Resistance. *Antimicrob Agents Chemother.* 2009 Jul;53.
66. 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 Dec 5;9(12).
67. Afonso Morais J, Bello G, Guimarães L. M, Sojka M, Morgado G. M. 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).

68. Bello G, Afonso JM, Morgado MG. Phylodynamics of HIV-1 subtype F1 in Angola, Brazil and Romania. *Infection, Genetics and Evolution*. 2012 Jul;12(5):1079–86.
69. Mendes Da Silva RK, Morais JA, Foley BT, Bello G, Morgado MG, Guimarães M. Identification of a new circulating recombinant form of human immunodeficiency virus type 1, CRF124_cpx involving subtypes A, G, H, and CRF27_cpx in Angola. *Front Microbiol*. 2022 Oct 17;13.
70. Sebastião CS, Neto Z, de Jesus CS, Mirandela M, Jandondo D, Couto-Fernandez JC, et al. Genetic diversity and drug resistance of HIV-1 among infected pregnant women newly diagnosed in Luanda, Angola. *PLoS One*. 2019 Nov 1;14(11).
71. Sebastião CS, Morais J, Brito M. Clinical and public health implications of HIV-1 genetic diversity and drug resistance mutations in Angola: A systematic review. *AIDS Rev*. 2021;22(4):1–9.
72. 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 Aug 1;32(8):822–8.
73. Joint United Nations Programme on HIV/AIDS (UNAIDS). HIV in Angola [Internet]. 2022 [cited 2024 Mar 16]. Available from: <https://www.unaids.org/en/regionscountries/countries/angola>
74. Quitéria R. Diversidade genética e resistência aos antirretrovirais no vírus da imunodeficiência humana tipo 1 em circulação em Benguela, Angola. 2019.
75. Los Alamos Data Base. HIV sequence locator [Internet]. 2019 [cited 2023 Feb 23]. Available from: <https://www.hiv.lanl.gov/content/sequence/LOCATE/locate.html>
76. Thierry E, Deprez E, Delelis O. Different pathways leading to integrase inhibitors resistance. *Front Microbiol*. 2017 Jan 11;7.
77. Mbhele N, Chimukangara B, Gordon M. HIV-1 integrase strand transfer inhibitors: a review of current drugs, recent advances and drug resistance. *Int J Antimicrob Agents*. 2021 May 1;57(5).

78. Almeida Machado L, Ramos Guimarães AC. Evidence for Disruption of Mg²⁺ Pair as a Resistance Mechanism Against HIV-1 Integrase Strand Transfer Inhibitors. *Front Mol Biosci*. 2020 Aug 20;7.
79. Videira e Castro S. Diversidade genética e resistência aos anti-retrovirais inibidores enzimáticos de vírus da imunodeficiência humana tipo 1 (HIV-1) de um população de toxicod dependentes por via endovenosa da grande Lisboa. 2010.
80. Andrade MD, Skalka AM. Retroviral Integrase: Then and Now. Vol. 2, Annual Review of Virology. Annual Reviews Inc.; 2015. p. 241–64.
81. Krishnan L, Engelman A. Retroviral integrase proteins and HIV-1 DNA integration. *Journal of Biological Chemistry*. 2012 Nov 30;287(49):40858–66.
82. Demeulemeester J, Chaltin P, Marchand A, De Maeyer M, Debyser Z, Christ F. LEDGINS, non-catalytic site inhibitors of HIV-1 integrase: A patent review (2006-2014). *Expert Opin Ther Pat*. 2014;24(6):609–32.
83. Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature*. 2010 Mar 11;464(7286):232–6.
84. Ohata Y, Tomonaga M, Watanabe Y, Tomura K, Kimura K, Akaki T, et al. Antiviral Activity and Resistance Profile of the Novel HIV-1 Non-Catalytic Site Integrase Inhibitor JTP-0157602. *J Virol* [Internet]. 2022;96(6). Available from: <https://journals.asm.org/journal/jvi>
85. Balakrishnan M, Yant SR, Tsai L, O’Sullivan C, Bam RA, Tsai A, et al. Non-Catalytic Site HIV-1 Integrase Inhibitors Disrupt Core Maturation and Induce a Reverse Transcription Block in Target Cells. *PLoS One*. 2013 Sep 9;8(9).
86. Scarsi K. Kimberly, Havens P. Joshua, Podany T. Anthony, Avedissian N. Sean, Fletcher V. Courtney. HIV-1 Integrase Inhibitors: A Comparative Review of Efficacy and Safety. *Drugs*. 2020;80(16):1649–76.
87. Zhao AV, Crutchley RD, Guduru RC, Ton K, Lam T, Min AC. A clinical review of HIV integrase strand transfer inhibitors (INSTIs) for the prevention and treatment of HIV-1 infection. *Retrovirology*. 2022 Dec 1;19(1).

88. Zulfiqar HF, Javed A, Sumbal A, Afroze B, Ali Q, Akbar K, et al. HIV Diagnosis and Treatment through Advanced Technologies. *Frontiers in Public Health*. 2017 Mar 7;5.
89. Kulpa DA, Chomont N. HIV persistence in the setting of antiretroviral therapy: when, where and how does HIV hide? *J Virus Erad*. 2015;1(2):59–66.
90. Iyidogan P, Anderson KS. Current perspectives on HIV-1 antiretroviral drug resistance. *Viruses*. 2014 Oct 24;6(10):4095–139.
91. República de Angola. Nota Técnica N° 01/2021. Ministério da Saúde. 2021.
92. ICAP Global Health. ICAP supports transition to dolutegravir based antirretroviral regimens throughout angola [Internet]. [cited 2023 Feb 23]. Available from: <https://icap.columbia.edu/news-events/with-new-regimens-icap-supports-transition-to-dolutegravir-based-antirretroviral-regimens-throughout-angola/>
93. Organização Pan-Americana da Saúde, Organização Mundial de Saúde. OMS recomenda dolutegravir como principal opção de tratamento para HIV entre todas as populações [Internet]. 2019 [cited 2023 Feb 23]. Available from: <https://www.paho.org/pt/noticias/22-7-2019-oms-recomenda-dolutegravir-como-principal-opcao-tratamento-para-hiv-entre-todas>
94. Józwik IK, Passos DO, Lyumkis D. Structural Biology of HIV Integrase Strand Transfer Inhibitors. *Trends Pharmacol Sci*. 2020 Sep 1;41(9):611–26.
95. Cook NJ, Li W, Berta D, Badaoui M, Ballandras-Colas A, Nans A, et al. Structural basis of second-generation HIV integrase inhibitor action and viral resistance. *Science* (1979). 2020 Feb 14;367(6479):806–10.
96. Organização das Nações Unidas. ONU News. 2021 [cited 2023 Feb 23]. Ainda é possível acabar com a Aids até 2030, diz ONU. Available from: <https://news.un.org/pt/story/2021/12/1772082>
97. Organização das Nações Unidas. ONU News. 2021 [cited 2023 Feb 23]. Mundo deve reduzir novas infeções anuais por HIV para menos de 370 mil . Available from: <https://news.un.org/pt/story/2021/04/1749302>

98. Global HIV Prevention Coalition. Prevenção do VIH 2025. 2021.
99. Clutter DS, Jordan MR, Bertagnolio S, Shafer RW. HIV-1 Drug Resistance and Resistance Testing. *Infection, genetics and evolution*. 2016 Dec;46:292–307.
100. Coelho LPO, Matsuda EM, Nogueira RS, de Moraes MJ, Jamal LF, Madruga JVR, et al. Prevalence of HIV-1 transmitted drug resistance and viral suppression among recently diagnosed adults in São Paulo, Brazil. *Arch Virol*. 2019 Mar 1;164(3):699–706.
101. Baxter J, Dunn D, White E, Sharma S, Geretti A, Kozal M, et al. Global HIV-1 transmitted drug resistance in the INSIGHT Strategic Timing of AntiRetroviral Treatment (START) trial. *HIV Med*. 2015 Apr 1;16(S1):77–87.
102. Rhee SY, Kassaye SG, Barrow G, Sundaramurthi JC, Jordan MR, Shafer RW. HIV-1 transmitted drug resistance surveillance: shifting trends in study design and prevalence estimates. *Journal of the International AIDS Society (JIAS)* [Internet]. 2020 Sep;23(9). Available from: <http://onlinelibrary.wiley.com/doi/10.1002/jia2.25611/full>
103. Wensing AM, Calvez V, Ceccherini-Silberstein F, Charpentier C, Günthard HF, Paredes R, et al. 2022 Update of the Drug Resistance Mutations in HIV-1. IAS-USA [Internet]. 2022 Oct;30(4). Available from: www.iasusa.org
104. O. A. Shadrina, T. S. Zatsepin, Yu. Yu. Agapkina, M. G. Isaguliants, M. B. Gottikh. Influence of drug resistance mutations on the activity of HIV-1 subtypes A and B integrases_A comparative study. *Acta Naturae*. 2015;7(1).
105. Delelis O, Malet I, Na L, Tchertanov L, Calvez V, Marcelin AG, et al. The G140S mutation in HIV integrases from raltegravir-resistant patients rescues catalytic defect due to the resistance Q148H mutation. *Nucleic Acids Res*. 2009;37(4):1193–201.
106. Loosli T, Hossmann S, Ingle SM, Okhai H, Kusejko K, Mouton J, et al. HIV-1 drug resistance in people on dolutegravir-based antiretroviral therapy: a collaborative cohort analysis. *Lancet HIV*. 2023 Nov 1;10(11):e733–41.

107. Rhee SY, Grant PM, Tzou PL, Barrow G, Harrigan PR, Ioannidis JPA, et al. A systematic review of the genetic mechanisms of dolutegravir resistance. *Journal of Antimicrobial Chemotherapy*. 2019 Nov 1;74(11):3135–49.
108. Oliveira M, Ibanescu RI, Anstett K, Mésplède T, Routy JP, Robbins MA, et al. Selective resistance profiles emerging in patient-derived clinical isolates with Cabotegravir, Bictegravir, Dolutegravir, and Elvitegravir. *Retrovirology*. 2018 Aug 17;15(1).
109. HIV Drug Resistance Database. INSTI SRDMS LIST 2019 [Internet]. [cited 2023 Jun 9]. Available from: <https://hivdb.stanford.edu/page/insti-sdrm-list/>
110. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. *PLoS One*. 2009 Mar 6;4(3).
111. WHO. World Health Organization [Internet]. [cited 2023 Feb 23]. Available from: <https://www.who.int/pt>
112. Organização Mundial de Saúde. Prevenção e combate contra a resistência medicamentosa do VIH na região Africana [Internet]. 2023 [cited 2023 Feb 23]. Available from: <https://hivdb.stanford.edu/>
113. Healgen Scientific. HIV Rapid Test: Healgen [Internet]. 2017 [cited 2023 Jun 2]. Available from: <http://www.healgen.com/infectious-disease/hiv/>
114. Türklab. Anti-HIV 1/2 TEST | TURKLAB [Internet]. 2017 [cited 2023 Jun 2]. Available from: <https://www.turklab.com.tr/hiv-tests/anti-hiv-1-2-test>
115. GE Healthcare. Whatman Indicating FTA Elute micro cards [Internet]. 2017 [cited 2023 Jun 2]. Available from: [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)
116. Human Diagnostics. Hexagon HIV [Internet]. 2017 [cited 2023 Jun 2]. Available from: https://www.human.de/fileadmin/content/flyer/en/981565_Hexagon_HIV_EN.pdf

117. Human Diagnostics. Hexagon HIV - Dispositivo Médico para Diagnóstico In vitro. 2017.
118. ANRS AC43 HIV Resistance Study Group PCR and Sequencing Procedures. HIV-1 Version 2022.1 PCR and sequencing Primers. 2022.
119. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors (DNA polymerase/nucleotide sequences/bacteriophage 4X174). *Proc Natl Acad Sci U S A* [Internet]. 1977;74(12):5463–7. Available from: <https://www.pnas.org>
120. Saitou N, Nei M. The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol Biol Evol* [Internet]. 1987 Jul;4(4):406–25. Available from: <https://academic.oup.com/mbe/article/4/4/406/1029664>
121. Kimura M. Journal of Molecular Evolution A Simple Method for Estimating Evolutionary Rates of Base Substitutions Through Comparative Studies of Nucleotide Sequences. *J Mol Evol*. 1980;16:111–20.
122. Lim MD. Dried blood spots for global health diagnostics and surveillance: Opportunities and challenges. *American Journal of Tropical Medicine and Hygiene*. 2018;99(2):256–65.
123. Denniff P, Spooner N. Effect of storage conditions on the weight and appearance of dried blood spot samples on various cellulose-based substrates. *Bioanalysis*. 2010 Nov;2(11):1817–22.
124. Hearps AC, Ryan CE, Morris LM, Plate MM, Greengrass V, Crowe SM. Stability of Dried Blood Spots for HIV-1 Drug Resistance Analysis. *Current HIV Research* [Internet]. 2010;8:134–40. Available from: <http://hivdb>.
125. 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* [Internet]. 2017 Sep;55(9). Available from: <https://www.hiv.lanl>
126. Pineda-Peña AC, 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. *Infection, Genetics and Evolution*. 2013;19:337–48.
127. Genome Detective, REGA HIV-1 subtyping tool [Internet]. [cited 2023 Sep 11]. Available from: <https://www.genomedetective.com/app/typingtool/hiv>
 128. Thomson MM, Casado G, Posada D, Sierra M, Nájera R. Identification of a novel HIV-1 complex circulating recombinant form (CRF18_cpx) of Central African origin in Cuba. *AIDS* [Internet]. 2005;19(11). Available from: <http://www.mbio.ncsu.edu/BioEdit/bioedit>.
 129. Cruz S, Brito R, Afonso J, Soares H. HIV-1 infection among pregnant women in Angola: molecular epidemiology and test of novel AZT Triazole derivatives. Universidade Nova de Lisboa; 2021.
 130. Gui T, Zhao J, Sun C, Chen L, Liu Y, Zheng C, et al. Genetic characterization of a unique recombinant originating from CRF55-01B, CRF01-AE, and CRF07-BC in Shenzhen, China. *AIDS Res Hum Retroviruses*. 2015 May 1;31(5):559–63.
 131. Li X, Liu H, Liu L, Feng Y, Kalish ML, Ho SYW, et al. Tracing the epidemic history of HIV-1 CRF01-AE clusters using near-complete genome sequences. *Sci Rep*. 2017 Dec 1;7(1).
 132. Helena C, Barroso M. Os acordos de financiamento entre a China e Angola: uma reconstrução pós-conflito sem reformas políticas. [Lisboa]: Escola de Sociologia e Políticas Públicas, ISCTE-IUL; 2015.
 133. Derache A, Shin HS, 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 Jan 30;10(1).
 134. Curanovic D, Martens SK, Rodriguez MA, Hammill HA, Petropoulos CJ, Walworth CM. HIV-1 DNA Testing in Viremic Patients Identifies More Drug Resistance Than HIV-1 RNA Testing. *Open Forum Infect Dis*. 2023 Apr 1;10(4).
 135. Armenia D, Zaccarelli M, Borghi V, Gennari W, Di Carlo D, Giannetti A, et al. Resistance detected in PBMCs predicts virological rebound in HIV-1 suppressed patients switching treatment. *Journal of Clinical Virology*. 2018 Jul 1;104:61–4.

136. Cortez KJ, Maldarelli F. Clinical management of HIV drug resistance. *Viruses*. 2011 Apr;3(4):347–78.
137. Li YL, Langley CA, Azumaya CM, Echeverria I, Chesarino NM, Emerman M, et al. The structural basis for HIV-1 Vif antagonism of human APOBEC3G. *Nature*. 2023 Mar 23;
138. Sharma S, Patnaik SK, Taggart RT, Baysal BE. The double-domain cytidine deaminase APOBEC3G is a cellular site-specific RNA editing enzyme. *Sci Rep*. 2016 Dec 15;6.
139. Goila-Gaur R, Strebel K. HIV-1 Vif, APOBEC, and intrinsic immunity. *Retrovirology*. 2008 Jun 24;5.
140. Stanford University. HIV Drug Resistance Database [Internet]. 2023 [cited 2024 Apr 3]. Available from: <https://hivdb.stanford.edu/>
141. Tzou PL, Rhee SY, Descamps D, Clutter DS, Hare B, Mor O, et al. Integrase strand transfer inhibitor (INSTI)-resistance mutations for the surveillance of transmitted HIV-1 drug resistance. *Journal of Antimicrobial Chemotherapy*. 2020 Jan 1;75(1):170–82.
142. Alaoui N, El Alaoui MA, Touil N, El Annaz H, Melloul M, Tagajdid R, et al. Prevalence of resistance to integrase strand-transfer inhibitors (INSTIs) among untreated HIV-1 infected patients in Morocco. *BMC Res Notes*. 2018 Jun 8;11(1).
143. Kirichenko A, Lapovok I, Baryshev P, van de Vijver DAMC, van Kampen JJA, Boucher CAB, et al. Genetic features of HIV-1 integrase sub-subtype A6 predominant in Russia and predicted susceptibility to instis. *Viruses*. 2020 Aug 1;12(8).
144. Abram ME, Hluhanich RM, Goodman DD, Andreatta KN, Margot NA, Ye L, et al. Impact of primary elvitegravir resistance-associated mutations in HIV-1 integrase on drug susceptibility and viral replication fitness. *Antimicrob Agents Chemother*. 2013 Jun;57(6):2654–63.

145. Scherrer AU, Yang WL, Kouyos RD, Böni J, Yerly S, Klimkait T, et al. Successful Prevention of Transmission of Integrase Resistance in the Swiss HIV Cohort Study. *Journal of Infectious Diseases*. 2016 Aug 1;214(3):399–402.
146. Gatell JM, Katlama C, Grinsztejn B, Eron JJ, Lazzarin A, Vittecoq D, et al. Long-Term Efficacy and Safety of the HIV Integrase Inhibitor Raltegravir in Patients With Limited Treatment Options in a Phase II Study. *J Acquir Immune Defic Syndr (1988)*. 2010 Apr 1;53(4):456–63.
147. Release notes, HIVdb [Internet]. [cited 2023 Sep 11]. Available from: <https://hivdb.stanford.edu/page/release-notes/>
148. Brenner BG, Lowe M, Moisi D, Hardy I, Gagnon S, Charest H, et al. Subtype diversity associated with the development of HIV-1 resistance to integrase inhibitors. *J Med Virol*. 2011 May;83(5):751–9.
149. Delicado TL, Torrecilla E, Holguín África. Deep analysis of HIV-1 natural variability across HIV-1 variants at residues associated with integrase inhibitor (INI) resistance in INI-naive individuals. *Journal of Antimicrobial Chemotherapy*. 2016 Feb 1;71(2):362–6.
150. Han YS, Mesplède T, Wainberg MA. Differences among HIV-1 subtypes in drug resistance against integrase inhibitors. *Infection, Genetics and Evolution*. 2016 Dec 1;46:286–91.
151. Lapovok I, Laga V, Kazennova E, Bobkova M. HIV Type 1 Integrase Natural Polymorphisms in Viral Variants Circulating in FSU Countries. *Curr HIV Res*. 2017 Aug 17;15(5).
152. Wenk BM, Mbunkah HA, Nsanwe NN, Mbu ET, Besong LM, Sama BA, et al. Prevalence of integrase strand transfer inhibitor resistance mutations in antiretroviral-naive HIV-1-infected individuals in Cameroon. *Journal of Antimicrobial Chemotherapy*. 2021;76(1):124–9.
153. Inzaule SC, Hamers RL, Noguera-Julian M, Casadellà M, Parera M, de Wit TFR, et al. Primary resistance to integrase strand transfer inhibitors in patients infected

- with diverse HIV-1 subtypes in sub-Saharan Africa. *Journal of Antimicrobial Chemotherapy*. 2018 May 1;73(5):1167–72.
154. Theys K, Abecasis AB, Vandamme AM. HIV-1 drug resistance: Where do polymorphisms fit in? *Future Microbiol*. 2013 Mar;8(3):303–6.
 155. Lapovok I, Laga V, Kazennova E, Bobkova M. HIV Type 1 Integrase Natural Polymorphisms in Viral Variants Circulating in FSU Countries. *Curr HIV Res*. 2017 Aug 17;15(5).
 156. Casado G, Thomson MM, Sierra M, Nájera R. Identification of a Novel HIV-1 Circulating ADG Intersubtype Recombinant Form (CRF19_cpx) in Cuba. *J Acquir Immune Defic Syndr (1988)* [Internet]. 2005 Dec 15;40(5):532–7. Available from: <http://www.mbio.ncsu>.
 157. Vidal N, Frange P, Chaix ML, Mulanga C, Lepira F, Bazepeo SE, et al. Characterization of an old complex circulating recombinant form, CRF27_cpx, originating from the Democratic Republic of Congo (DRC) and circulating in France. *AIDS Res Hum Retroviruses*. 2008 Feb 1;24(2):315–21.
 158. Carr JK, Salminen MO, Albert J, Sanders-Buell E, Gotte D, Birx DL, et al. Full Genome Sequences of Human Immunodeficiency Virus Type 1 Subtypes G and A/G Intersubtype Recombinants. *Virology*. 1998 Jul 20;247(1):22–31.

6. Anexos

Anexo 1. Póster apresentado no 6.º Congresso Nacional de Medicina Tropical

Resistência primária aos inibidores da integrase no vírus da imunodeficiência humana do tipo 1 (HIV-1) na província de Benguela (Angola)

Isabel S. Godinho^{1*}, Raquel Quitéria¹, Lesya Yefimenko^{1,2}, Filomena M. Pereira², João Piedade¹

¹Unidade de Microbiologia Médica e ²Unidade de Clínica Tropical, GHTM, Instituto de Higiene e Medicina Tropical, Universidade NOVA de Lisboa
*marisargodinho@gmail.com



Introdução

No final de 2021, o HIV-1 infetava aproximadamente 38,4 milhões de pessoas (1).

A terapia antirretroviral combinada tem permitido uma redução significativa da mortalidade, morbidade e transmissibilidade viral. Contudo, devido às elevadas taxas de mutação, o HIV-1 adquire facilmente resistência aos antirretrovirais (ARVs). Entre estes, contam-se os inibidores de transferência de cadeia da integrase (INSTI), para os quais estão descritas as mutações associadas a resistência representadas na figura 1. Desta forma, reveste-se de elevada importância o estudo da resistência primária a esta classe de fármacos, em locais em que a sua introdução é muito recente, como em Angola.

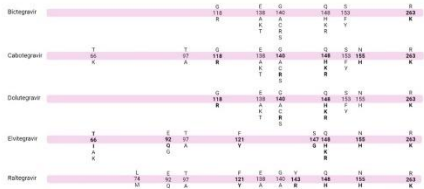
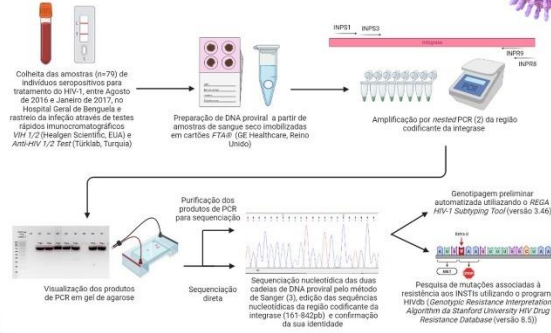


Figura 1. Mutações na região codificante da integrase associadas a resistência aos inibidores de transferência de cadeia da integrase.

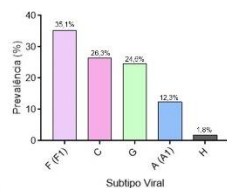
Este trabalho tem como objetivos a pesquisa de polimorfismos genéticos naturais na região codificante da integrase e sua associação com a resistência primária aos INSTI, bem como estimar a prevalência dos genótipos virais presentes no HIV-1 em circulação na província de Benguela, em Angola.

Materiais e Métodos



Resultados

1. Genotipagem



A genotipagem automatizada permitiu a atribuição de um genótipo a 57/79 sequências em análise. Os resultados encontram-se apresentados na figura 2.

Figura 2. Distribuição de genótipos para a região codificante da integrase do HIV-1 circulante na província de Benguela (Angola), obtidos com recurso ao REGA HIV-1 Subtyping Tool (versão 3.46) (N = 57).

Resultados

2. Mutações associadas a resistência aos inibidores de transferência de cadeia da integrase (INSTI)

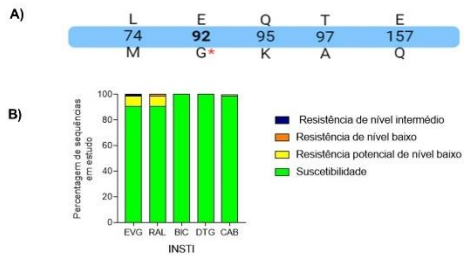


Figura 3. A) Mutações associadas a resistência aos INSTI. As letras superiores representam o aminoácido da sequência de consenso do subtipo B (código de uma letra), o número refere-se à sua posição na sequência e as letras inferiores indicam o aminoácido associado a resistência aos INSTI. A negrito está destacada a mutação principal, enquanto que as restantes são acessórias. O asterisco vermelho especifica uma mutação associada a resistência aos INSTI de vigilância epidemiológica da resistência transmitida no HIV-1. B) Representação gráfica da distribuição relativa dos níveis de resistência putativos aos cinco INSTI. Elvitegravir (EVG), Raltegravir (RAL), Bictegravir (BIC), Dolastegravir (DTG) e Cabotegravir (CAB), associados às sequências da região codificante da integrase analisadas. Nesta análise, foram excluídas 6/79 sequências em estudo, por apresentarem indícios de mutação por hiperamplificação associada à atividade da APOBEC3G. (4)

Conclusões

- A genotipagem automatizada preliminar de 57 sequências nucleotídicas da região codificante da integrase do HIV-1 circulante na província de Benguela (Angola), em indivíduos não submetidos a terapia antirretroviral, demonstrou uma maior prevalência do subtipo F1 (35,1%), seguida dos subtipos C (26,3%), G (24,6%), A1 (12,3%) e H (1,8%).
- A vasta maioria das sequências nucleotídicas estudadas (90,2%) estão associadas a uma total suscetibilidade aos INSTIs.
- Foi observada uma mutação principal (E92G), relevante para a vigilância epidemiológica, associada a níveis intermédio e baixo de resistência a Elvitegravir (EVG) e Raltegravir (RAL), respetivamente.
- Foram também encontradas as mutações acessórias L74M, Q95K, T97A e E157Q, associadas a níveis de resistência potencial de nível baixo, entre outros polimorfismos genéticos (dados não apresentados).
- Os resultados obtidos reforçam a importância da monitorização da emergência de mutações de resistência primária aos diferentes ARVs e sua transmissão em Angola, considerando o seu impacto direto na efetividade terapêutica e controlo da infeção.

Referências

- World health organization. HIV data and statistic, 2022. <https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv-strategic-information/hiv-data-and-statistics>
- ANRS AC43 HIV Resistance Study Group (2022). PCR and Sequencing Procedures: HIV-1 (v2022.1). <https://hivfrenchresistance.org/>
- Sanger F et al. DNA sequencing with chain-terminating inhibitors. 1977. *Biochemistry*. Vol. 16, No. 12, pp. 5463-5467. <https://doi.org/10.1073/pnas.74.12.5463>
- Yen-Li Li et al. The structural basis for HIV-1 Vif antagonism of human APOBEC3G. 2003. *Nature*. Vol 415, pp. 728-733. <https://doi.org/10.1038/n41586-023-05779-1>

Agradecimentos

- A todos os utentes do Hospital Geral de Benguela (Angola) que deram o seu consentimento informado, livre e esclarecido para participação neste estudo.

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_6	Não sabe	Feminino	Viúvo	Sem Escolaridade	Rural		X
AO_7	58	Masculino	Solteiro	Ensino Secundário	Urbana		
AO_25	38	Feminino	Casado	Ensino Básico	Urbana		X
AO_26	48	Masculino	Casado	Ensino Básico	Urbana		X
AO_51	52	Feminino	Solteiro	Ensino Secundário	Urbana	X	
AO_61	6	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_80	22	Masculino	Solteiro	Ensino Básico	Urbana		
AO_84	52	Feminino	Viúvo	Sem Escolaridade	Rural		
AO_91	25	Feminino	Casado	Ensino Secundário	Urbana		
AO_92	30	Masculino	Casado	Ensino Superior	Urbana		
AO_116	53	Feminino	Casado	Ensino Básico	Urbana		
AO_119	26	Feminino	Solteiro	Ensino Secundário	Urbana		
AO_141	53	Feminino	Divorciado	Ensino Secundário	Urbana	X	
AO_153	31	Masculino	Casado	Ensino Superior	Urbana		
AO_177	6	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_201	35	Feminino	Casado	Ensino Superior	Urbana	X	

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_202	46	Masculino	Divorciado	Ensino Básico	Urbana		X
AO_205	45	Masculino	Casado	Ensino Básico	Rural		
AO_206	65	Feminino	Viúvo	Ensino Secundário	Urbana		
AO_209	3 meses	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_210	67	Feminino	Viúvo	Ensino Básico	Urbana		
AO_211	52	Feminino	Casado	Ensino Básico	Urbana		
AO_216	18	Masculino	Solteiro	Ensino Básico	Rural		
AO_217	6	Masculino	Solteiro	Sem Escolaridade	Urbana		
AO_224	8	Masculino	Solteiro	Sem Escolaridade	Urbana		
AO_235	49	Feminino	Divorciado	Ensino Básico	Rural		
AO_237	63	Feminino	Viúvo	Sem Escolaridade	Rural	X	
AO_245	64	Feminino	Viúvo	Sem Escolaridade	Rural		X
AO_252	57	Feminino	Casado	Ensino Secundário	Urbana	X	
AO_269	29	Masculino	Solteiro	Ensino Secundário	Urbana	X	
AO_277	38	Masculino	Divorciado	Ensino Básico	Rural		
AO_289	69	Feminino	Viúvo	Ensino Superior	Urbana	X	
AO_307	60	Feminino	Casado	Ensino Secundário	Urbana		

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_309	28	Masculino	Casado	Ensino Secundário	Urbana		
AO_313	22	Masculino	Solteiro	Ensino Secundário	Rural		
AO_314	56	Masculino	Viúvo	Ensino Básico	Urbana		
AO_315	66	Feminino	Viúvo	Ensino Secundário	Urbana		
AO_316	52	Feminino	Casado	Ensino Básico	Rural		
AO_317	65	Feminino	Casado	Ensino Secundário	Urbana		
AO_318	35	Masculino	Solteiro	Ensino Secundário	Urbana		
AO_319	25	Feminino	Solteiro	Ensino Secundário	Urbana		
AO_320	33	Masculino	Casado	Ensino Básico	Rural		
AO_321	50	Feminino	Casado	Ensino Secundário	Urbana		
AO_322	21	Masculino	Solteiro	Ensino Básico	Rural		
AO_324	53	Feminino	Viúvo	Ensino Secundário	Urbana		
AO_325	43	Masculino	Casado	Ensino Secundário	Urbana		
AO_327	61	Feminino	Casado	Ensino Secundário	Urbana		
AO_330	24	Masculino	Solteiro	Ensino Secundário	Urbana		
AO_333	35	Feminino	Divorciado	Ensino Superior	Urbana		
AO_334	54	Masculino	Casado	Ensino Secundário	Rural		
AO_335	25	Feminino	Solteiro	Ensino Secundário	Rural		

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_337	51	Feminino	Casado	Ensino Secundário	Urbana		
AO_338	43	Feminino	Divorciado	Ensino Básico	Rural		
AO_340	67	Feminino	Viúvo	Ensino Básico	Rural		
AO_341	18	Feminino	Solteiro	Sem Escolaridade	Rural		X
AO_342	33	Masculino	Casado	Ensino Secundário	Urbana		
AO_343	57	Feminino	Casado	Ensino Básico	Rural		
AO_344	32	Masculino	Casado	Ensino Secundário	Urbana		
AO_347	27	Feminino	Casado	Ensino Secundário	Urbana		
AO_348	50	Feminino	Casado	Ensino Secundário	Urbana		
AO_349	7	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_350	54	Feminino	Casado	Ensino Secundário	Urbana		
AO_351	42	Masculino	Casado	Ensino Secundário	Rural		
AO_352	54	Feminino	Casado	Ensino Superior	Urbana		
AO_353	11	Masculino	Solteiro	Ensino Básico	Rural		
AO_354	36	Masculino	Divorciado	Ensino Superior	Urbana		
AO_355	28	Feminino	Solteiro	Ensino Secundário	Urbana		
AO_356	33	Masculino	Casado	Ensino Secundário	Urbana		
AO_357	33	Feminino	-	Sem Escolaridade	Urbana		X

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_358	24	Feminino	Solteiro	Ensino Secundário	Urbana		
AO_359	38	Masculino	Casado	Ensino Básico	Rural		
AO_360	14	Masculino	Solteiro	Sem Escolaridade	Urbana		X
AO_363	60	Feminino	Solteiro	Sem Escolaridade	Rural		X
AO_364	60	Feminino	Viúvo	Ensino Secundário	Rural		
AO_365	20	Masculino	Solteiro	Ensino Básico	Urbana		
AO_366	66	Feminino	Viúvo	Sem Escolaridade	Rural		
AO_367	34	Feminino	Casado	Ensino Secundário	Urbana		
AO_368	63	Feminino	Casado	Ensino Secundário	Urbana		
AO_369	53	Feminino	Casado	Ensino Básico	Rural		
AO_370	12	Feminino	Solteiro	Sem Escolaridade	Rural		X
AO_371	37	Feminino	Solteiro	Ensino Secundário	Urbana	X	X
AO_372	31	Feminino	Solteiro	Ensino Básico	Urbana		X
AO_373	32	Feminino	Solteiro	Ensino Superior	Urbana		X
AO_374	47	Feminino	Casado	Ensino Secundário	Urbana		
AO_377	48	Masculino	Divorciado	Ensino Secundário	Urbana		
AO_378	34	Masculino	Casado	Ensino Básico	Rural	X	
AO_379	44	Masculino	Divorciado	Ensino Básico	Rural	X	

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_380	10	Feminino	Solteiro	Sem Escolaridade	Rural		
AO_381	51	Masculino	Casado	Ensino Secundário	Urbana		
AO_382	34	Feminino	Casado	Ensino Secundário	Rural		
AO_383	63	Feminino	Casado	Ensino Secundário	Rural		
AO_384	21	Masculino	Solteiro	Ensino Básico	Urbana		
AO_385	32	Masculino	Casado	Ensino Superior	Urbana		
AO_386	55	Feminino	Divorciado	Ensino Secundário	Rural		
AO_387	37	Masculino	Casado	Ensino Superior	Urbana		
AO_388	48	Feminino	Casado	Ensino Secundário	Urbana		
AO_389	50	Feminino	Casado	Ensino Secundário	Urbana		
AO_390	29	Masculino	Casado	Ensino Básico	Urbana		
AO_391	17	Masculino	Solteiro	Ensino Básico	Urbana		
AO_392	65	Masculino	Viúvo	Ensino Básico	Rural		
AO_394	24	Masculino	Solteiro	Ensino Básico	Urbana		X
AO_395	51	Masculino	Casado	Ensino Secundário	Urbana		
AO_397	36	Masculino	Casado	Ensino Superior	Urbana		
AO_398	62	Feminino	Casado	Ensino Básico	Rural		
AO_399	37	Feminino	Casado	Ensino Secundário	Rural		

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_404	59	Feminino	Viúvo	Ensino Secundário	Rural	X	
AO_405	52	Feminino	Viúvo	Sem Escolaridade	Rural		
AO_406	50	Masculino	Casado	Ensino Básico	Urbana	X	
AO_409	30	Masculino	Casado	Ensino Superior	Urbana		
AO_410	35	Masculino	Divorciado	Ensino Secundário	Urbana		
AO_412	53	Feminino	Casado	Ensino Secundário	Urbana		
AO_414	37	Feminino	Solteiro	Sem Escolaridade	Rural		
AO_415	5	Feminino	Solteiro	Sem Escolaridade	Rural		X
AO_416	25	Feminino	Casado	Sem Escolaridade	Urbana		
AO_417	19	Masculino	Solteiro	Ensino Básico	Urbana		
AO_418	47	Feminino	Divorciado	Ensino Secundário	Urbana	X	
AO_419	Não sabe	Masculino	Divorciado	Sem Escolaridade	Rural		
AO_423	16	Feminino	Solteiro	Sem Escolaridade	Rural		
AO_426	80	Masculino	Casado	Ensino Básico	Rural		
AO_431	95	Masculino	Viúvo	Ensino Básico	Urbana		
AO_432	18	Feminino	Divorciado	Ensino Básico	Urbana		
AO_433	3	Masculino	Solteiro	Sem Escolaridade	Urbana		X
AO_434	67	Feminino	Viúvo	Sem Escolaridade	Rural		

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_436	Não sabe	Feminino	Viúvo	Sem Escolaridade	Rural		
AO_437	23	Feminino	Solteiro	Ensino Básico	Urbana		
AO_438	15	Masculino	Solteiro	Sem Escolaridade	Urbana		
AO_439	Não sabe	Feminino	Casado	Sem Escolaridade	Rural		
AO_440	52	Feminino	Casado	Ensino Básico	Rural	X	
AO_441	3 meses	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_442	15	Feminino	Solteiro	Sem Escolaridade	Rural		
AO_443	3	Masculino	Solteiro	Sem Escolaridade	Urbana		
AO_444	32	Masculino	Divorciado	Ensino Básico	Urbana		
AO_445	61	Masculino	Casado	Ensino Secundário	Urbana		
AO_446	78	Feminino	Viúvo	Sem Escolaridade	Urbana		
AO_447	41	Feminino	Casado	Ensino Superior	Urbana		
AO_448	50	Masculino	Solteiro	Ensino Básico	Urbana		
AO_449	52	Masculino	Divorciado	Ensino Básico	Urbana		
AO_450	2	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_451	53	Feminino	Casado	Ensino Básico	Urbana		
AO_452	16	Feminino	Solteiro	Sem Escolaridade	Urbana		
AO_453	5	Masculino	Solteiro	Sem Escolaridade	Urbana		

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_454	24	Masculino	Solteiro	Ensino Básico	Urbana		
AO_455	72	Feminino	Casado	Ensino Básico	Urbana		
AO_456	67	Masculino	Casado	Sem Escolaridade	Rural		
AO_457	18	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_458	1	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_459	23	Feminino	Solteiro	Ensino Superior	Urbana		
AO_460	59	Feminino	Casado	Ensino Secundário	Urbana		
AO_461	41	Masculino	Divorciado	Ensino Básico	Rural		
AO_462	14	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_463	66	Feminino	Divorciado	Ensino Básico	Urbana		
AO_464	49	Feminino	Solteiro	Ensino Básico	Urbana		
AO_465	66	Feminino	Casado	Ensino Secundário	Urbana		
AO_466	26	Masculino	Solteiro	Ensino Básico	Urbana		
AO_467	64	Feminino	Casado	Sem Escolaridade	Urbana		
AO_468	6	Masculino	Solteiro	Sem Escolaridade	Urbana		
AO_469	27	Feminino	Solteiro	Ensino Secundário	Urbana		
AO_470	12	Masculino	Solteiro	Sem Escolaridade	Urbana		
AO_471	53	Feminino	Solteiro	Ensino Secundário	Urbana		

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_472	1	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_473	57	Feminino	Casado	Sem Escolaridade	Rural		
AO_474	22	Feminino	Solteiro	Sem Escolaridade	Rural		
AO_475	32	Feminino	Divorciado	Ensino Secundário	Urbana		
AO_476	20	Masculino	Solteiro	Ensino Básico	Rural		
AO_477	57	Feminino	Casado	Ensino Básico	Urbana		
AO_478	6	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_479	27	Feminino	Solteiro	Ensino Secundário	Rural		
AO_480	31	Masculino	Casado	Ensino Básico	Urbana		
AO_481	50	Feminino	Casado	Ensino Básico	Urbana		
AO_482	9	Feminino	Solteiro	Sem Escolaridade	Urbana		X
AO_483	9	Feminino	Solteiro	Sem Escolaridade	Rural	X	
AO_484	40	Masculino	Solteiro	Ensino Básico	Urbana		
AO_485	36	Masculino	Casado	Ensino Básico	Rural		
AO_486	26	Feminino	Solteiro	Ensino Básico	Urbana		
AO_488	60	Feminino	Viúvo	Ensino Secundário	Rural		
AO_489	13	Feminino	Solteiro	Sem Escolaridade	Rural		X
AO_490	54	Masculino	Casado	Ensino Secundário	Urbana	X	

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_491	33	Feminino	Casado	Ensino Básico	Urbana		
AO_492	40	Feminino	Divorciado	Ensino Básico	Rural	X	
AO_493	46	Feminino	Casado	Ensino Secundário	Urbana		
AO_494	11	Masculino	Solteiro	Sem Escolaridade	Urbana		X
AO_495	42	Masculino	Divorciado	Ensino Básico	Rural		
AO_496	6	Feminino	Solteiro	Sem Escolaridade	Rural		
AO_497	35	Feminino	Casado	Ensino Básico	Rural	X	
AO_499	54	Feminino	Casado	Ensino Secundário	Rural		
AO_500	48	Feminino	Casado	Ensino Secundário	Urbana		
AO_501	42	Feminino	Casado	Ensino Secundário	Urbana		
AO_502	50	Masculino	Casado	Ensino Básico	Urbana		
AO_503	52	Feminino	Divorciado	Ensino Secundário	Urbana		
AO_505	15	Feminino	Solteiro	Sem Escolaridade	Rural		X
AO_506	71	Masculino	Viúvo	Sem Escolaridade	Urbana		
AO_507	0	Masculino	Solteiro	Sem Escolaridade	Rural		X
AO_508	17	Feminino	Solteiro	Sem Escolaridade	Urbana		
AO_510	23	Masculino	Solteiro	Ensino Básico	Rural		
AO_511	39	Feminino	Casado	Ensino Básico	Rural		

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_512	24	Feminino	Solteiro	Sem Escolaridade	Urbana		
AO_513	73	Feminino	Viúvo	Sem Escolaridade	Rural		
AO_515	57	Feminino	Casado	Sem Escolaridade	Rural		
AO_516	5 dias	Feminino	Solteiro	Sem Escolaridade	Rural		
AO_518	57	Masculino	Divorciado	Ensino Básico	Urbana		
AO_519	58	Masculino	Solteiro	Ensino Secundário	Urbana		
AO_520	5	Masculino	Solteiro	Sem Escolaridade	Urbana		
AO_521	42	Feminino	Divorciado	Ensino Básico	Urbana		
AO_522	35	Feminino	Divorciado	Ensino Superior	Urbana		
AO_523	4	Feminino	Solteiro	Sem Escolaridade	Rural		
AO_524	14	Feminino	Solteiro	Sem Escolaridade	Urbana		
AO_525	17	Feminino	Solteiro	Sem Escolaridade	Rural		
AO_526	19	Masculino	Solteiro	Ensino Básico	Urbana		
AO_527	7	Masculino	Solteiro	Sem Escolaridade	Rural	X	X
AO_528	52	Feminino	Solteiro	Sem Escolaridade	Urbana		
AO_529	33	Masculino	Solteiro	Sem Escolaridade	Rural		
AO_530	49	Feminino	Divorciado	Ensino Básico	Urbana		
AO_531	36	Feminino	Casado	Sem Escolaridade	Rural		

Anexo 2. Dados epidemiológicos da amostra populacional incluída neste estudo

Dados epidemiológicos dos 226 indivíduos seropositivos para o HIV-1 incluídos no presente estudo. No quadro é indicado o código do indivíduo, idade (em anos, relativamente ao ano de 2016), sexo, estado civil, nível de escolaridade (“sem escolaridade” = analfabetismo ou ciclo básico de escolaridade incompleto), área de residência (rural ou urbana) e outras infeções sexualmente transmissíveis (IST). A utilização do símbolo “-“ corresponde à não obtenção de resposta por parte do indivíduo.

Código do indivíduo	Idade	Sexo	Estado civil	Escolaridade	Residência	Outras IST	
						Hepatite B	Sífilis
AO_532	63	Feminino	Viúvo	Sem Escolaridade	Rural		
AO_533	29	Feminino	Divorciado	Ensino Básico	Urbana		
AO_534	59	Feminino	Casado	Ensino Secundário	Urbana		
AO_535	47	Feminino	Divorciado	Ensino Secundário	Urbana		
AO_536	18	Masculino	Solteiro	Ensino Básico	Urbana		
AO_537	7	Masculino	Solteiro	Sem Escolaridade	Urbana		
AO_538	51	Feminino	Casado	Ensino Básico	-		
AO_539	35	Feminino	Casado	Ensino Básico	Urbana		
AO_540	52	Feminino	Divorciado	Ensino Básico	Urbana		
AO_541	25	Feminino	Casado	Ensino Básico	Urbana		
AO_542	50	Feminino	Casado	Ensino Básico	Urbana		
AO_543	32	Feminino	Casado	Ensino Básico	Urbana		
AO_545	58	Masculino	Divorciado	Ensino Básico	Urbana		

Anexo 3. Protocolo para preparação de cartões *FTA*[®]

Cartões *FTA*[®] – Amostras de sangue

Preparação dos discos

1. Limpar o tapete de corte com etanol a 70% (v/v).
2. Limpar a extremidade do “puncher” (pressionar o êmbolo) com etanol a 70% (v/v) e de seguida fazer seis furos em papel Whatman.
3. Pressionar o “puncher” sobre a amostra de sangue no cartão *FTA* e introduzir o disco de cartão num *epENDORF* de 1,5 ml.
4. Repetir o passo 2.
5. Adicionar 200 µl de *FTA purification reagent* a cada disco.
Incubar durante 5 min. à temperatura ambiente, agitando o tubo de vez em quando.
Notas: Pode agitar-se no vortex.
É normal formar-se espuma, mas deve evitar-se.
O meio deve ficar amarelado, devido à saída de hemoglobina do disco.
Pretende-se obter um disco com cor clara.
6. Remover o meio (e a espuma) com auxílio de micropipeta com pontas com filtro.
7. Repetir os passos 5 e 6 mais duas vezes (fazer mais repetições, com tempos de incubação maiores, caso o disco não se apresente claro).
8. Adicionar 200 µl de tampão TE (10 mM Tris-HCl, 0,1 mM EDTA, pH 8,0) a cada disco e agitar o tubo.
Incubar durante 5 min. à temperatura ambiente, agitando o tubo de vez em quando.
Nota: Pode agitar-se no vortex.
É normal formar-se espuma, mas deve evitar-se.
9. Remover o meio (e a espuma) com auxílio de micropipeta com pontas com filtro.
10. Repetir os passos 8 e 9.
11. Colocar os tubos abertos no *SpeedVac (Eppendorf concentrator 5301)*, no programa para soluções aquosas, a 30°C durante 10 min.
Agitar o tubo para ver se o disco está bem seco (neste caso, deve soltar-se facilmente das paredes do tubo). Caso ainda não esteja seco, secar a 30°C mais 5 min.
12. O disco pode ser utilizado até 3h após o processamento. Caso contrário, pode ser conservado a 4°C (ou a -20°C) por um período máximo de uma semana.

Nota: Quanto mais “vermelhos” os discos ficarem, menor será o rendimento da PCR.

Anexo 4. Nomenclatura dos nucleótidos e aminoácidos

Código de uma letra da IUPAC para a nomenclatura dos nucleótidos.

Código IUPAC	Base orgânica
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

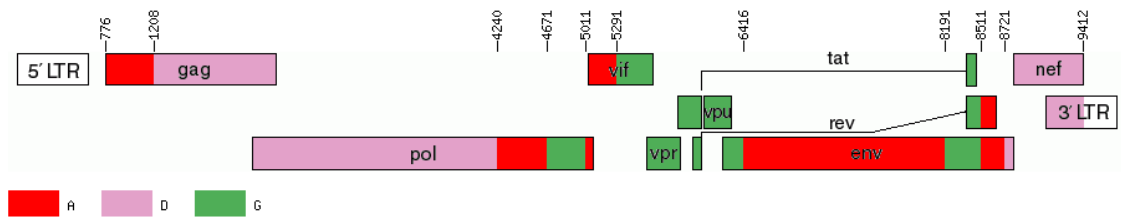
Código de três letras e de uma letra da IUPAC para a nomenclatura dos aminoácidos.

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 5. Representação do genoma mosaico das formas recombinantes circulantes CRF19_cpx, CRF124_cpx, CRF27_cpx, CRF02_AG e CRF18_cpx (*Los Alamos HIV Sequence Database*)

(disponível em, <https://www.hiv.lanl.gov/components/sequence/HIV/crfdb/crfs.comp>).

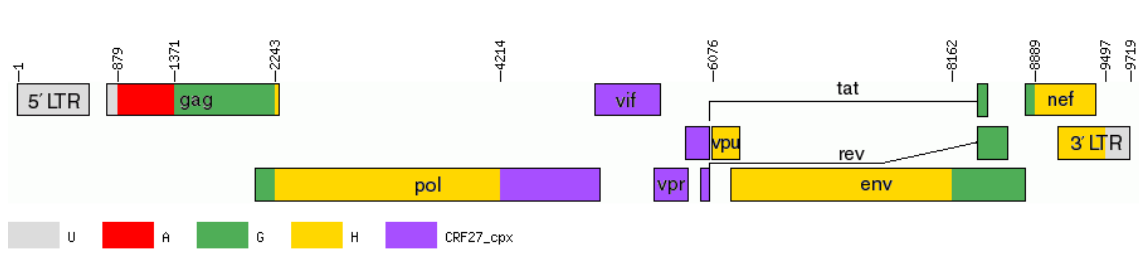
- CRF19_cpx



Pontos de recombinação em HXB2 (Casado *et al.*, 2005) (156):

- 776 1207 A
- 1208 4239 D
- 4240 4670 A
- 4671 5010 G
- 5011 5290 A*
- 5291 6415 G
- 6416 8190 A
- 8191 8510 G
- 8511 8720 A
- 8721 9412 D

○ CRF124_cpx



Pontos de recombinação em HXB2 (Mendes da Silva *et al.* 2022) (69):

1 878 U

879 1370 A

1371 2242 G

2243 4213 H

4214 6075 CRF27_cpx *

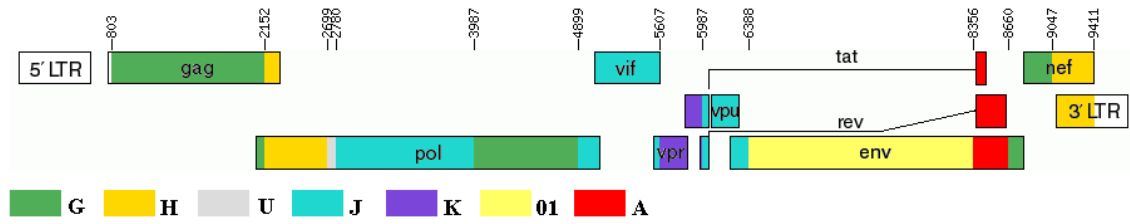
6076 8161 H

8162 8888 G

8889 9496 H

9497 9719 U

○ CRF27_cpx



Pontos de recombinação em HXB2 (Vidal *et al.*, 2008) (157):

803 2151 G

2152 2698 H

2699 2779 U

2780 3986 J

3987 4898 G

4899 5606 J *

5607 5986 K

5987 6387 J

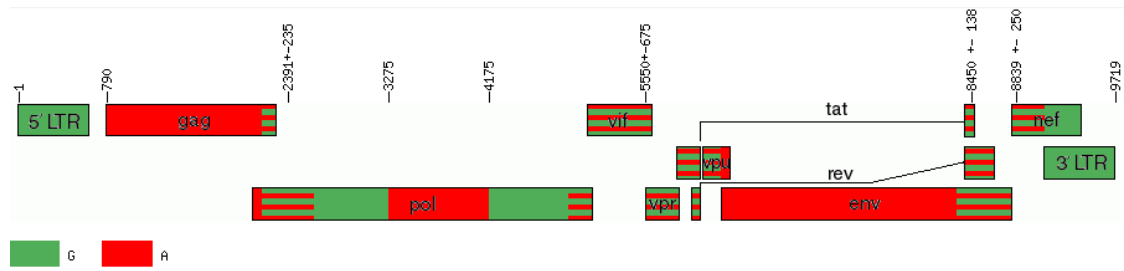
6388 8355 01_AE

8356 8659 A

8660 9046 G

9047 9411 H

○ CRF02_AG



Pontos de recombinação em HXB2 (Carr *et al.*, 1998) (158):

1 789 G

790 2155 A

2156 2625 A/G

2626 3274 G

3275 4174 A

4175 4874 G

4875 6224 A/G *

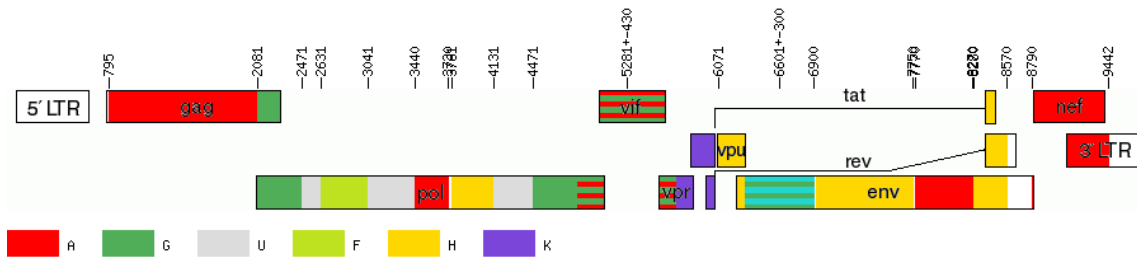
6225 8311 A

8312 8588 A/G

8589 9089 A/G

9090 9719 G

○ CRF18_cpx



Pontos de recombinação em HXB2 (Thomson *et al.*, 2005) (128):

795 2080 A

2081 2470 G

2471 2630 U

2631 3040 F

3041 3439 U

3440 3730 A

3761 4130 H

4131 4470 U

4471 4850 G

4851 5710 A/G *

5711 6070 K

6071 6300 H

6301 6900 G/J

6920 7750 H

7770 8270 A

8280 8570 H

8790 9442 A

*Segmentos nucleotídicos que incluem a região codificante da integrase em estudo no gene *pol* (posições nucleotídicas 4390 a 5071, em HXB2).

Anexo 6. Listagem dos polimorfismos genéticos identificados nas 92 sequências aminoacídicas putativas em estudo e associação com significado estatístico a subtipos ($p < 0,01$) (assinaturas genéticas de subtipo), recorrendo ao teste exato de Fisher bilateral ou χ^2 .

O quadro apresenta a comparação entre sequências do subtipo A e não-A. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas A ou não-A. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo A	Subtipos não-A	p
D1E	1	0	0,207
D1N	0	2	1,000
S3A	0	1	1,000
S3V	0	1	1,000
S3F	0	1	1,000
G5K	0	2	1,000
G5E	0	2	1,000
I6M	1	4	1,000
I6V	0	4	0,577
L9I	2	2	0,188
L9V	1	0	0,207
E15K	0	1	1,000
G16R	0	6	0,339
G16C	0	1	1,000
G16E	0	3	1,000
K17C	0	2	1,000
K17C	0	1	1,000
I18V	10	33	0,563*
I19V	0	1	1,000
L20I	5	9	0,131*
L20V	0	1	1,000
L20M	0	3	1,000
A22G	0	1	1,000
V23A	0	1	1,000
A26V	0	1	1,000
G28R	0	3	1,000
G28S	0	1	1,000
G28D	0	1	1,000
Y29F	0	1	1,000
I30M	1	3	1,000
E31K	0	1	1,000
E33K	0	2	1,000

O quadro apresenta a comparação entre sequências do subtipo A e não-A. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas A ou não-A. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo A	Subtipos não-A	p
A37S	1	0	0,207
A37C	0	1	1,000
A37E	0	3	1,000
A37T	0	2	1,000
E38K	0	1	1,000
E38G	1	0	0,207
T39I	1	0	0,207
G40R	0	2	1,000
C41K	0	2	1,000
C41Q	0	1	1,000
E42D	2	5	0,631
T43A	0	3	1,000
A44G	0	4	0,577
F46Y	0	13	0,063
L47I	3	57	< 0,00001
L47V	2	4	0,600
K49R	0	2	1,000
G52A	3	3	0,100
G52R	0	2	1,000
G52E	0	1	1,000
K57R	1	3	1,000
K57T	1	1	0,372
T58V	13	41	0,334*
T58I	2	20	0,225
T58M	1	3	1,000
T58A	0	1	1,000
I59V	7	6	0,001*
G64R	0	1	1,000
S65P	8	9	0,003*
S65T	1	10	0,449
S65R	0	6	0,339
S65G	0	1	1,000
T68I	3	2	0,058
T70A	11	54	0,170*
T70C	1	1	0,372
T70N	3	5	0,355

O quadro apresenta a comparação entre sequências do subtipo A e não-A. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas A ou não-A. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo A	Subtipos não-A	p
T70G	2	1	0,107
T70S	2	3	0,274
T71A	18	66	0,551*
T71V	0	1	1,000
V72M	0	1	1,000
V72A	0	1	1,000
K73R	0	4	0,577
G80D	7	10	0,021*
G80N	8	12	0,016*
G80S	1	0	0,207
I81V	5	22	0,745*
I81L	0	1	1,000
K82R	1	1	0,372
K82C	10	40	0,866*
K82H	3	3	0,100
K82N	0	3	1,000
K82T	0	13	0,063
K82Q	0	2	1,000
C83H	0	3	1,000
C83Q	0	1	1,000
C83R	0	1	1,000
E84D	0	3	1,000
F85Y	1	0	0,207
F85H	0	2	1,000
G86R	0	5	0,580
G86E	0	1	1,000
I87V	0	1	1,000
C94Q	0	1	1,000
G95R	0	4	0,577
V96I	0	2	1,000
V97I	0	1	1,000
E98K	0	2	1,000
M100I	0	3	1,000
E103K	0	2	1,000
E103C	1	0	0,207
K106C	1	0	0,207

O quadro apresenta a comparação entre sequências do subtipo A e não-A. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas A ou não-A. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo A	Subtipos não-A	p
K106R	0	1	1,000
I107L	0	1	1,000
I108V	0	1	1,000
G109R	0	3	1,000
G109T	0	1	1,000
G109N	0	1	1,000
G109E	0	1	1,000
C110Q	1	2	0,505
V111I	0	3	1,000
R112K	0	3	1,000
D113E	8	9	0,003*
D113K	0	1	1,000
C114Q	1	1	0,372
H117Y	0	1	1,000
K119R	0	1	1,000
V122I	0	1	1,000
F127L	1	4	1,000
F127R	0	1	1,000
I128V	0	1	1,000
K134R	0	8	0,198
G135E	0	1	1,000
G136K	0	1	1,000
I137M	1	0	0,207
I137V	0	1	1,000
G138R	0	2	1,000
G139K	0	1	1,000
G139R	0	1	1,000
G139E	0	2	1,000
G139D	0	1	1,000
Y140C	0	2	1,000
A142V	0	1	1,000
G143E	0	4	0,577
G143R	0	1	1,000
V147I	18	66	0,551*
I149M	2	1	0,107
A151S	0	14	0,066

O quadro apresenta a comparação entre sequências do subtipo A e não-A. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas A ou não-A. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo A	Subtipos não-A	p
T152S	8	31	0,977*
D153E	2	6	0,667
D153N	0	1	1,000
I154L	4	4	0,054
I154M	0	1	1,000
T156V	1	1	0,372
T156I	0	2	1,000
K157R	3	13	1,000
K157Q	1	0	0,207
K157T	0	1	1,000
K157C	0	2	1,000
E158A	1	3	1,000
E158I	0	2	1,000
K161N	1	3	1,000
C162R	2	0	0,041
C162H	1	0	0,207
T164S	1	4	1,000
T164I	3	32	0,033
K165N	4	2	0,016
K165C	0	2	1,000
I161V	1	2	0,505
I161L	0	2	1,000
C162T	2	0	0,041
C162S	1	0	0,207
C162H	0	3	1,000
C162E	0	1	1,000
N163K	2	3	0,274
N163A	0	1	1,000
N163H	0	2	1,000
R170C	1	5	1,000
R170Q	0	1	1,000
Y173F	1	10	0,449
D175G	0	1	1,000
S176N	0	2	1,000
D178E	0	2	1,000
L180I	16	50	0,175*

O quadro apresenta a comparação entre sequências do subtipo A e não-A. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas A ou não-A. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo A	Subtipos não-A	p
L180V	2	23	0,085
K182R	1	0	0,207
K186R	1	0	0,207
G193R	0	1	1,000
I197L	0	1	1,000
C198E	0	1	1,000
X199E	0	1	1,000
X200S	0	1	1,000
N201K	3	2	0,058
N201G	0	1	1,000
N201W	0	1	1,000
N201S	0	1	1,000
S202G	2	4	0,600
S202T	1	1	0,372
S202C	1	0	0,207
S202N	0	18	0,019
D203E	4	43	0,004
V206I	0	2	1,000
R209K	0	1	1,000
R210K	0	1	1,000
K211R	0	1	1,000
A212V	0	12	0,066
A212T	0	1	1,000
I214F	0	1	1,000
I215L	2	2	0,188
R216K	1	10	0,449
R216C	0	1	1,000
D217N	1	1	0,372
D217H	2	5	0,631
D217E	1	3	1,000
G219R	0	3	1,000
D225A	0	16	0,036
D225T	0	1	1,000
D225E	0	1	1,000
D225N	0	1	1,000
D226G	0	4	0,577

O quadro apresenta a comparação entre sequências do subtipo A e não-A. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas A ou não-A. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo A	Subtipos não-A	p
D226Y	0	1	1,000
C227G	1	0	0,207
C227W	0	1	1,000
C227R	0	1	1,000
C227S	0	1	1,000
Total	19	73	92

O quadro apresenta a comparação entre sequências do subtipo C e não-C. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas C ou não-C. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo C	Subtipos não-C	p
D1E	0	1	1,000
D1N	0	2	1,000
S3A	1	0	0,196
S3V	1	0	0,196
S3F	1	0	0,196
G5K	1	1	0,355
G5E	0	2	1,000
I6M	1	4	1,000
I6V	1	3	1,000
L9I	2	2	0,171
L9V	0	1	1,000
E15K	0	1	1,000
G16R	2	4	0,334
G16C	0	1	1,000
G16E	0	3	1,000
K17C	0	2	1,000
K17C	0	1	1,000
I18V	12	31	0,059*
I19V	0	1	1,000
L20I	2	12	0,730
L20V	1	0	0,196
L20M	0	3	1,000
A22G	0	1	1,000
V23A	0	1	1,000
A26V	0	1	1,000
G28R	1	2	0,484
G28S	1	0	0,196
G28D	0	1	1,000
Y29F	0	1	1,000
I30M	2	2	0,171
E31K	0	1	1,000
E33K	1	1	0,355
A37S	0	1	1,000
A37C	1	0	0,196
A37E	3	0	0,007
A37T	0	2	1,000

O quadro apresenta a comparação entre sequências do subtipo C e não-C. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas C ou não-C. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo C	Subtipos não-C	p
E38K	0	1	1,000
E38G	0	1	1,000
T39I	0	1	1,000
G40R	1	1	0,355
C41K	1	1	0,355
C41Q	0	1	1,000
E42D	0	7	0,338
T43A	0	3	1,000
A44G	0	4	0,582
F46Y	13	0	< 0,00001
L47I	16	44	0,019 *
L47V	1	5	1,000
K49R	1	1	0,355
G52A	0	6	0,593
G52R	1	1	0,355
G52E	1	0	0,196
K57R	1	3	1,000
K57T	1	1	0,355
T58V	15	39	0,018*
T58I	1	21	0,062
T58M	1	3	1,000
T58A	0	1	1,000
I59V	0	13	0,065
G64R	0	1	1,000
S65P	3	14	1,000
S65T	7	4	0,001
S65R	3	3	0,087
S65G	0	1	1,000
T68I	0	5	0,579
T70A*	8	57	0,006*
T70C	0	2	1,000
T70N	2	6	0,652
T70G	0	3	1,000
T70S	2	3	0,251
T71A	18	66	0,348
T71V	0	1	1,000

O quadro apresenta a comparação entre sequências do subtipo C e não-C. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas C ou não-C. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo C	Subtipos não-C	p
V72M	0	1	1,000
V72A	0	1	1,000
K73R	3	1	0,023
G80D	1	16	0,178
G80N	0	20	0,009
G80S	0	1	1,000
I81V	0	27	0,001
I81L	1	0	0,196
K82R	1	1	0,355
K82C	11	39	0,521*
K82H	1	5	1,000
K82N	2	1	0,097
K82T	0	13	0,065
K82Q	1	1	0,355
C83H	0	3	1,000
C83Q	0	1	1,000
C83R	0	1	1,000
E84D	0	3	1,000
F85Y	0	1	1,000
F85H	0	2	1,000
G86R	2	3	0,251
G86E	0	1	1,000
I87V	0	1	1,000
C94Q	0	1	1,000
G95R	2	2	0,171
V96I	1	1	0,355
V97I	1	0	0,196
E98K	0	2	1,000
M100I	1	2	0,484
E103K	1	1	0,355
E103C	0	1	1,000
K106C	0	1	1,000
K106R	0	1	1,000
I107L	0	1	1,000
I108V	0	1	1,000
G109R	2	1	0,097

O quadro apresenta a comparação entre sequências do subtipo C e não-C. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas C ou não-C. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo C	Subtipos não-C	p
G109T	0	1	1,000
G109N	0	1	1,000
G109E	0	1	1,000
C110Q	0	3	1,000
V111I	0	3	1,000
R112K	1	2	0,484
D113E	3	14	1,000
D113K	0	1	1,000
C114Q	0	2	1,000
H117Y	0	1	1,000
K119R	0	1	1,000
V122I	0	1	1,000
F127L	1	4	1,000
F127R	1	0	0,196
I128V	0	1	1,000
K134R	3	5	0,186
G135E	1	0	0,196
G136K	1	0	0,196
I137M	0	1	1,000
I137V	0	1	1,000
G138R	2	0	0,037
G139K	1	0	0,196
G139R	0	1	1,000
G139E	0	2	1,000
G139D	0	1	1,000
Y140C	1	1	0,355
A142V	0	1	1,000
G143E	2	2	0,1706
G143R	1	0	0,1957
V147I	17	67	0,598*
I149M	0	3	1,000
A151S	2	12	0,729
T152S	6	33	0,386*
D153E	2	6	0,652
D153N	0	1	1,000
I154L	1	7	1,000

O quadro apresenta a comparação entre sequências do subtipo C e não-C. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas C ou não-C. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo C	Subtipos não-C	p
I154M	0	1	1,000
T156V	0	2	1,000
T156I	0	2	1,000
K157R	2	14	0,729
K157Q	0	1	1,000
K157T	1	0	0,196
K157C	0	2	1,000
E158A	0	4	0,582
E158I	0	2	1,000
K161N	1	3	1,000
C162R	0	2	1,000
C162H	0	1	1,000
T164S	0	5	0,579
T164I	6	29	0,646*
K165N	0	6	0,593
K165C	0	2	1,000
I161V	0	3	1,000
I161L	1	1	0,355
C162T	0	2	1,000
C162S	0	1	1,000
C162H	3	0	0,007
C162E	0	1	1,000
N163K	2	3	0,251
N163A	1	0	0,196
N163H	2	0	0,037
R170C	2	4	0,334
R170Q	0	1	1,000
Y173F	1	10	0,685
D175G	0	1	1,000
S176N	0	2	1,000
D178E	0	2	1,000
L180I	17	49	0,017*
L180V	1	24	0,020
K182R	0	1	1,000
K186R	0	1	1,000
G193R	0	1	1,000

O quadro apresenta a comparação entre sequências do subtipo C e não-C. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas C ou não-C. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo C	Subtipos não-C	p
I197L	1	0	0,196
C198E	0	1	1,000
X199E	0	1	1,000
X200S	0	1	1,000
N201K	1	4	1,000
N201G	0	1	1,000
N201W	0	1	1,000
N201S	0	1	1,000
S202G	2	4	0,334
S202T	0	2	1,000
S202C	0	1	1,000
S202N	2	16	0,509
D203E	3	44	0,001
V206I	0	2	1,000
R209K	0	1	1,000
R210K	1	0	0,196
K211R	1	0	0,196
A212V	4	8	0,241
A212T	0	1	1,000
I214F	1	0	0,196
I215L	1	3	1,000
R216K	4	7	0,216
R216C	0	1	1,000
D217N	0	2	1,000
D217H	0	7	0,338
D217E	1	3	1,000
G219R	2	1	0,097
D225A	12	4	< 0,00001
D225T	1	0	0,196
D225E	0	1	1,000
D225N	0	1	1,000
D226G	0	4	0,582
D226Y	0	1	1,000
C227G	0	1	1,000
C227W	1	0	0,196
C227R	0	1	1,000

O quadro apresenta a comparação entre sequências do subtipo C e não-C. Nas células a fundo cinzento estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas C ou não-C. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo C	Subtipos não-C	p
C227S	0	1	1,000
Total	18	74	92

O quadro apresenta a comparação entre sequências do subtipo F1 e não-F1. Nas células a fundo lilás estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas F1 ou não-F1. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo F1	Subtipos não-F1	p
D1E	0	1	1,000
D1N	1	1	0,488
S3A	0	1	1,000
S3V	0	1	1,000
S3F	0	1	1,000
G5K	1	1	0,488
G5E	0	2	1,000
I6M	3	2	0,135
I6V	1	3	1,000
L9I	0	4	0,574
L9V	0	1	1,000
E15K	1	0	0,283
G16R	2	4	1,000
G16C	0	1	1,000
G16E	0	3	0,556
K17C	2	0	0,078
K17C	0	1	1,000
I18V	9	34	0,144*
I19V	1	0	0,283
L20I	3	11	0,750
L20V	0	1	1,000
L20M	2	1	0,192
A22G	1	0	0,283
V23A	0	1	1,000
A26V	0	1	1,000
G28R	2	1	0,192
G28S	0	1	1,000
G28D	0	1	1,000
Y29F	0	1	1,000
I30M	0	4	0,574
E31K	0	1	1,000
E33K	1	1	0,488
A37S	0	1	1,000
A37C	0	1	1,000
A37E	0	3	0,556
A37T	1	1	0,488

O quadro apresenta a comparação entre sequências do subtipo F1 e não-F1. Nas células a fundo lilás estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas F1 ou não-F1. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo F1	Subtipos não-F1	p
E38K	1	0	0,283
E38G	0	1	1,000
T39I	0	1	1,000
G40R	1	1	0,488
C41K	0	2	1,000
C41Q	1	0	0,283
E42D	2	5	1,000
T43A	0	3	0,556
A44G	4	0	0,005
F46Y	0	13	0,017
L47I	20	40	0,139*
L47V	1	5	0,672
K49R	1	1	0,488
G52A	1	5	0,672
G52R	0	2	1,000
G52E	0	1	1,000
K57R	1	3	1,000
K57T	0	2	1,000
T58V	4	50	< 0,00001
T58I	16	6	< 0,00001*
T58M	2	2	0,316
T58A	0	1	1,000
I59V	1	12	0,101
G64R	0	1	1,000
S65P	0	17	0,002
S65T	2	9	0,722
S65R	2	4	1,000
S65G	0	1	1,000
T68I	0	5	0,317
T70A	23	42	0,019*
T70C	0	2	1,000
T70N	2	6	1,000
T70G	1	2	1,000
T70S	0	5	0,317
T71A	24	60	0,830*
T71V	0	1	1,000

O quadro apresenta a comparação entre sequências do subtipo F1 e não-F1. Nas células a fundo lilás estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas F1 ou não-F1. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo F1	Subtipos não-F1	p
V72M	0	1	1,000
V72A	0	1	1,000
K73R	1	3	1,000
G80D	2	15	0,137
G80N	0	20	0,001
G80S	0	1	1,000
I81V	9	18	0,486*
I81L	0	1	1,000
K82R	0	2	1,000
K82C	23	27	0,000037*
K82H	2	4	1,000
K82N	0	3	0,556
K82T	0	13	0,017
K82Q	1	1	0,488
C83H	0	3	0,556
C83Q	1	0	0,283
C83R	1	0	0,283
E84D	1	2	1,000
F85Y	0	1	1,000
F85H	1	1	0,488
G86R	1	4	1,000
G86E	0	1	1,000
I87V	0	1	1,000
C94Q	0	1	1,000
G95R	1	3	1,000
V96I	0	2	1,000
V97I	0	1	1,000
E98K	0	2	1,000
M100I	1	2	1,000
E103K	0	2	1,000
E103C	0	1	1,000
K106C	0	1	1,000
K106R	0	1	1,000
I107L	1	0	0,283
I108V	0	1	1,000
G109R	1	2	1,000

O quadro apresenta a comparação entre sequências do subtipo F1 e não-F1. Nas células a fundo lilás estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas F1 ou não-F1. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo F1	Subtipos não-F1	p
G109T	1	0	0,283
G109N	0	1	1,000
G109E	0	1	1,000
C110Q	1	2	1,000
V111I	1	2	1,000
R112K	1	2	1,000
D113E	3	14	0,378
D113K	0	1	1,000
C114Q	1	1	0,488
H117Y	1	0	0,283
K119R	0	1	1,000
V122I	1	0	0,283
F127L	2	3	0,619
F127R	0	1	1,000
I128V	1	0	0,283
K134R	3	5	0,683
G135E	0	1	1,000
G136K	0	1	1,000
I137M	0	1	1,000
I137V	0	1	1,000
G138R	0	2	1,000
G139K	0	1	1,000
G139R	1	0	0,283
G139E	2	0	0,078
G139D	1	0	0,283
Y140C	0	2	1,000
A142V	1	0	0,283
G143E	0	4	0,574
G143R	0	1	1,000
V147I	24	60	0,830*
I149M	1	2	1,000
A151S	9	5	0,001*
T152S	5	34	0,005*
D153E	2	6	1,000
D153N	0	1	1,000
I154L	2	6	1,000

O quadro apresenta a comparação entre sequências do subtipo F1 e não-F1. Nas células a fundo lilás estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas F1 ou não-F1. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo F1	Subtipos não-F1	p
I154M	0	1	1,000
T156V	1	1	0,488
T156I	1	1	0,488
K157R	4	12	1,000
K157Q	0	1	1,000
K157T	0	1	1,000
K157C	0	2	1,000
E158A	2	2	0,316
E158I	1	1	0,488
K161N	0	4	0,574
C162R	0	2	1,000
C162H	0	1	1,000
T164S	0	5	0,317
T164I	17	18	0,001*
K165N	0	6	0,179
K165C	1	1	0,488
I161V	0	3	0,556
I161L	0	2	1,000
C162T	0	2	1,000
C162S	0	1	1,000
C162H	0	3	0,556
C162E	1	0	0,283
N163K	0	5	0,317
N163A	0	1	1,000
N163H	0	2	1,000
R170C	1	5	0,672
R170Q	0	1	1,000
Y173F	2	9	0,722
D175G	1	0	0,283
S176N	2	0	0,078
D178E	2	0	0,078
L180I	5	61	< 0,00001*
L180V	21	4	< 0,00001
K182R	0	1	1,000
K186R	0	1	1,000
G193R	1	0	0,283

O quadro apresenta a comparação entre sequências do subtipo F1 e não-F1. Nas células a fundo lilás estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas F1 ou não-F1. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo F1	Subtipos não-F1	p
I197L	0	1	1,000
C198E	0	1	1,000
X199E	0	1	1,000
X200S	0	1	1,000
N201K	0	5	0,317
N201G	1	0	0,283
N201W	1	0	0,283
N201S	0	1	1,000
S202G	1	5	0,672
S202T	1	1	0,488
S202C	0	1	1,000
S202N	2	16	0,086
D203E	24	23	< 0,00001*
V206I	2	0	0,078
R209K	0	1	1,000
R210K	0	1	1,000
K211R	0	1	1,000
A212V	4	8	0,735
A212T	1	0	0,283
I214F	0	1	1,000
I215L	0	4	0,574
R216K	3	8	1,000
R216C	1	0	0,283
D217N	1	1	0,488
D217H	3	4	0,399
D217E	1	3	1,000
G219R	1	2	1,000
D225A	2	14	0,220
D225T	0	1	1,000
D225E	0	1	1,000
D225N	0	1	1,000
D226G	3	1	0,067
D226Y	0	1	1,000
C227G	0	1	1,000
C227W	0	1	1,000
C227R	1	0	0,283

O quadro apresenta a comparação entre sequências do subtipo F1 e não-F1. Nas células a fundo lilás estão representados os valores correspondentes aos polimorfismos que foram classificados como assinaturas F1 ou não-F1. Os valores de p com * resultam da aplicação do teste de χ^2 . A negrito estão representados valores de p que comprovam uma associação altamente significativa ($p < 0,00001$).

Mutações	Subtipo F1	Subtipos não-F1	p
C227S	0	1	1,000
Total	26	66	92

Anexo 7. 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">▪ 242 g Tris-base▪ 57,1 ml ácido acético glacial▪ 100 ml 0,5 M EDTA, pH 8,0▪ Adicionar água destilada para 1000 ml
Tampão TE	<ul style="list-style-type: none">▪ 10 mM Tris-HCl▪ 0,1 mM EDTA, pH 8,0



NOVA

UNIVERSIDADE NOVA
DE LISBOA