Universidade do Minho
Escola de Medicina

Ana Alexandra de Portugal dos Santos Pereira

Impact of HIV-1 genetic diversity on the interaction with host cells

fevereiro 2022

Impact of HIV-1 genetic diversity on the interaction with host cells

Ana Pereira

UMinho | 2022

Universidade do Minho
Escola de Medicina

Ana Alexandra de Portugal dos Santos Pereira

Impact of HIV-1 genetic diversity on the interaction with host cells

Tese de Doutoramento
Doutoramento em Envelhecimento e Doenças Crónicas

Trabalho efetuado sob a orientação do
Professor Doutor Nuno S. Osório
e da
Professora Doutora Helena Soares

fevereiro 2022

Agradecimentos

"I've heard it said that people come into our lives for a reason, bringing something we must learn. And we are led to those who help us most to grow if we let them and we help them in return (...). So much of me is made of what I learned from you. You'll be with me like a handprint on my heart."

- *For Good, Wicked*

Durante os últimos quatro anos, foram várias as pessoas que deixaram a sua marca na minha vida e que foram os alicerces da realização desta tese. Aproveito para destacar, aqui, algumas delas, e deixar-lhes os meus mais sinceros agradecimentos.

Em primeiro lugar, quero agradecer aos meus orientadores Doutor Nuno Osório e Doutora Helena Soares, não só por me terem aceitado nos seus grupos de investigação, mas, também, por todo o conhecimento, apoio e amizade que me ofereceram, durante todo este percurso.

Agradeço, também, à minha segunda família do i3.09: Bernardino, Deisy, Carlos, Pedro, Joana e Vera, por toda a vossa amizade e apoio. Obrigada, também, a todos os funcionários do bar da Escola de Medicina. Não houve dia mau que uma ida ao bar não curasse.

À Maria João Barreto, à Sara, ao Joca, ao Carlos e ao Neto. Por muito que o tempo passe, a nossa amizade não muda e, por isso, estou-vos eternamente grata.

Às minhas avós, por terem cuidado sempre de mim e por todas as nossas confidências.

À minha tia Margarida, por ser uma segunda mãe, uma amiga e uma irmã, e por estar sempre disponível para me ajudar.

Ao irmão que ganhei quando entrei nesta academia e que levo comigo para a vida toda.

Aos meus pais, que são os grandes responsáveis pela pessoa que sou e por tudo o que alcancei até hoje.

À minha irmã, por todo o seu apoio incondicional e por ser o par de todos os meus duetos.

Ao meu João, por ser o meu maior apoio, o meu melhor amigo e o meu companheiro de aventuras para o resto da vida.

O trabalho desta tese foi realizado no Instituto de Investigação em Ciências da Vida e de Saúde (ICVS), Universidade do Minho. O estudo foi financiado pela bolsa PD/BD/127827/2016 e por fundos nacionais, através da Fundação para a Ciência e Tecnologia - projetos UIDB/50026/2020 e UIDP/50026/2020.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Impacto da diversidade genética do VIH-1 na interação com o hospedeiro

Resumo

Em 2019, estima-se que 38 milhões de pessoas viveriam infetadas pelo vírus da imunodeficiência humana (VIH). Atualmente, não existe uma cura para a infeção por VIH e a extensa variabilidade genética do vírus representa um dos maiores obstáculos para acabar com a proliferação da pandemia e desenvolver uma terapia eficiente. O Brasil é o país com o maior número de pessoas infetadas por VIH na América Latina, apresentando uma distribuição heterogénea de subtipos e recombinantes virais pelas suas diferentes regiões geográficas. Perceber a epidemia do VIH-1 no Brasil poderá dar-nos informações valiosas sobre a diversidade genética do vírus e da sua relação com as características genéticas e sociodemográficas da população. Assim, primeiramente, neste estudo, tentamos perceber a dinâmica das mutações de resistência em pacientes em falência terapêutica no Brasil, entre 2008 e 2017, e, posteriormente, tentamos investigar a epidemiologia molecular e a evolução dos subtipos de VIH neste país.

Os nossos resultados revelaram uma prevalência de mutações de resistência elevada na população estudada, apesar de um decréscimo ter sido observado ao longo dos anos. Contrariamente, um aumento significativo foi observado na prevalência da mutação de resistência a inibidores da transcriptase reversa K65R, seguindo as alterações efetuadas a nível dos fármacos recomendados para tratamento. Foi, também, encontrada evidencia de transmissão da mutação e os nossos resultados sugeriram que a mesma poderia aumentar o reconhecimento viral pelo HLA-B27, que tem uma prevalência relativamente reduzida na população braseira. Mais ainda, foi verificado um aumento na prevalência do subtipo C, especialmente no Sul, onde é dominante. Encontramos, também, evidencia de transmissão do subtipo C entre o Sul e outras regiões do Brasil, apesar de a sua proporção ser pequena nestas áreas. Adicionalmente, este subtipo foi mais associado a baixos níveis de imunossupressão e a transmissão de mulher para homem e mãe para filho, do que o subtipo B, sustentando a hipótese que alguns subtipos podem tirar partido de períodos assintomáticos mais longos e das características sociodemográficas da população para proliferar.

No geral, estes resultados reforçam a importância de se perceber a dinâmica da expansão do VIH e da monitorização das mutações de resistência, de modo a serem criadas diretrizes de controlo e tratamento mais específicas, para se reduzir o fardo da epidemia do VIH.

Palavras-chaves: VIH-1, Diversidade Genética, Subtipos, Mutações de Resistência, Brasil

Impact of HIV-1 genetic diversity on the interaction with host cells

Abstract

In 2019, 38 million people were estimated to be living with Human Immunodeficiency Virus (HIV) infection. Currently, there is still no cure for HIV infection and the extensive viral genetic diversity represents the one of the biggest obstacles to end the spread of the pandemics and to develop an effective therapy. Brazil represents the country with the largest number of people living with HIV in Latin America, presenting a heterogeneous distribution of HIV-1 subtypes and recombinant forms across different geographic regions. Understanding HIV-1 epidemics in a country as vast as Brazil, could give valuable insights about the viral genetic diversity and its relationship with genetic and sociodemographic characteristics of the host population. Thus, in this study, we firstly aimed at understanding the dynamics of the drug resistance mutations (DRMs) in HIV-1 infected individuals failing antiretroviral treatment (ART) in Brazil, between the years 2008 and 2017, and, posteriorly, at investigating the molecular epidemiology and evolution of HIV-1 subtypes, in the same country.

Our results revealed a high prevalence of DRMs in the studied population, although a mild decline was observed over the years. Contrastingly, a significant increase on the prevalence of the K65R reverse transcriptase mutation was noticed, following a shift on the used ART regimens. Evidence of K65R transmission was also verified and our results suggested that this mutation could enhance viral recognition by HLA-B27 that has relatively low prevalence in the Brazilian population. Moreover, our results indicated an increase on subtype C prevalence over the years, especially in the South of Brazil, where it dominates. We also observed evidence for subtype C transmission events between the South and other Brazilian regions, although this subtype was only present in small proportions in these areas. Additionally, subtype C was significantly associated with lower levels of immunodepression infection of women and women-to-child transmission, when compared with subtype B, sustaining the hypothesis that some subtypes might take advantage of longer asymptomatic periods and the sociodemographic characteristics of the population to proliferate.

Overall, our results reinforce the importance of understanding the dynamics of HIV-1 subtype expansion and monitoring DRMs prevalence to establish specific guidelines for prevention and treatment, aiming at decreasing the epidemic burden of the HIV-1 infection.

Keywords: HIV-1, Genetic Diversity, Subtypes, Drug Resistance Mutation, Brazil.

Table of Contents

List of Abbreviations and Acronyms

3TC - Lamivudine;

ABC - Abacavir;

AIDS - Acquired Immunodeficiency Syndrome;

ART - Antiretroviral Treatment;

ATV - Atazanavir;

AZT - Zidovudine;

CCR5 - CC-Chemokine Receptor 5;

CRF - Circulating Recombinant Form;

CXCR4 - CXC-Chemokine Receptor 4;

d4T - Stavudine;

DRM - Drug Resistance Mutation;

DRV - Darunavir;

DTG - Dolutegravir;

EFV - Efavirenz;

EVG - Elvitegravir;

FTC - Emtricitabine;

GALT - Gut-Associated Lymphoid Tissue;

HIV - Human Immunodeficiency Virus;

HIV-1 - HIV-type 1;

HIV-2 - HIV-type 2;

INSTI - Integrase Strand-Transfer Inhibitor;

LPV - Lopinavir;

LTR - Long Terminal Repeat;

MSM - Men That Have Sex With Men;

Mtb - *Mycobacterium tuberculosis*;

NNRTI - Non-Nucleoside Reverse Transcriptase Inhibitor;

NRTI - Nucleoside Reverse Transcriptase Inhibitor;

NVP - Nevirapine;

PI - Protease Inhibitor;

PLWH - People Living With HIV;

PR - Protease;

PrPE - Pre-exposure Prophylaxis;

RAL - Raltegravir;

RT - Reverse Transcriptase;

RTV - Ritonavir;

SIV - Simian Immunodeficiency Virus;

SDRM - Surveillance Drug Resistance Mutation;

TAM - Thymidine Analog-Associated Mutation;

TDF - Tenofovir Disoproxil Fumarate;

TB – Tuberculosis;

UNAIDS - Joint United Nations Programme on HIV/AIDS;

USA - United States of America;

VIH – Vírus da Imunodeficiência Humana;

VL - Viral Loads;

WHO - World Health Organization.

List of Figures

List of tables

Chapter 1 | General Introduction

Part of this work contributed to the publication of the following article:

Santos-Pereira, A.; Magalhães, C.; Araújo, P.M.M.; Osório, N.S. Evolutionary Genetics of *Mycobacterium tuberculosis* and HIV-1: "The Tortoise and the Hare". Microorganisms 2021, 9, 147. *https://doi.org/10.3390/microorganisms9010147*

1. Introduction

1.1. Human Immunodeficiency Virus (HIV)

On the early 1980s, the first cases of an inexplicable and severe immunodeficiency, that was posteriorly described as Acquired Immunodeficiency Syndrome (AIDS), were reported in the United States of America (USA), among men who had sex with men [1–4]. The virus responsible for AIDS was identified after two years of these reports and named, in 1986, by the International Committee on Taxonomy of Viruses, as the Human Immunodeficiency Virus (HIV) [1,2,5–8]. The AIDS epidemic rapidly spread and, in 2019, 38 million people were estimated to be living with HIV worldwide [9]. Although new cases of infection were globally reduced by 23% since 2010, there was still an alarming increase of 72% in eastern Europe and more than 20% in Central Latin America, Asia, the Middle East and North Africa, comprising a total of 1.7 million new infections in 2019 [9]. Even though HIV-related deaths have presented a 51% decrease in the last 20 years, HIV infection is still one of the top ten leading causes of death in low-income countries and was estimated to be responsible for 690 000 AIDS-related deaths in the same year [9,10].

1.1.1. HIV-1 origin and diversity

Being part of the *Retroviridae* family, HIV can be grouped into HIV-type 1 (HIV-1) and HIV-type 2 (HIV-2). Both HIV types were likely introduced by multiple cross-species transmission events of simian immunodeficiency viruses (SIVs) infecting African non-human primates. Hunting non-human primates for bushmeat is believed to be the potential reasons behind these zoonotic transmissions, as well as injuries related with capture, trade and keeping these animals as pets [11,12].

At least four zoonotic transmission events are associated with the initial genetic difference between the different HIV-1 groups (M, N, O and P), while no less than nine independent cross-species transmissions are behind the introduction of the HIV-2 groups (A-I) [11–13]. Studies have found that the HIV-1 pandemic group M is closely related with SIVcpz that infected chimpanzees (*Pan troglodytes troglodytes*) from Cameroon southeast, while the rare group N ancestors have been identified in SIVcpz from chimpanzee communities of south-central Cameroon [12–15]. Group O and P, which also have not been disseminated widely, were found to be originated from independent zoonotic transmission events

from SIVgor, that infected Western lowland gorillas (*Gorilla gorilla gorilla*) from Cameroon [12,13,16,17]. On the other hand, HIV-2, which mainly restricted to West Africa, possibly for being less transmissible, is more closely related with the SIVsmm that infects sooty mangabeys (*Cercocebus atys*) (Figure 1) [11,18,19].

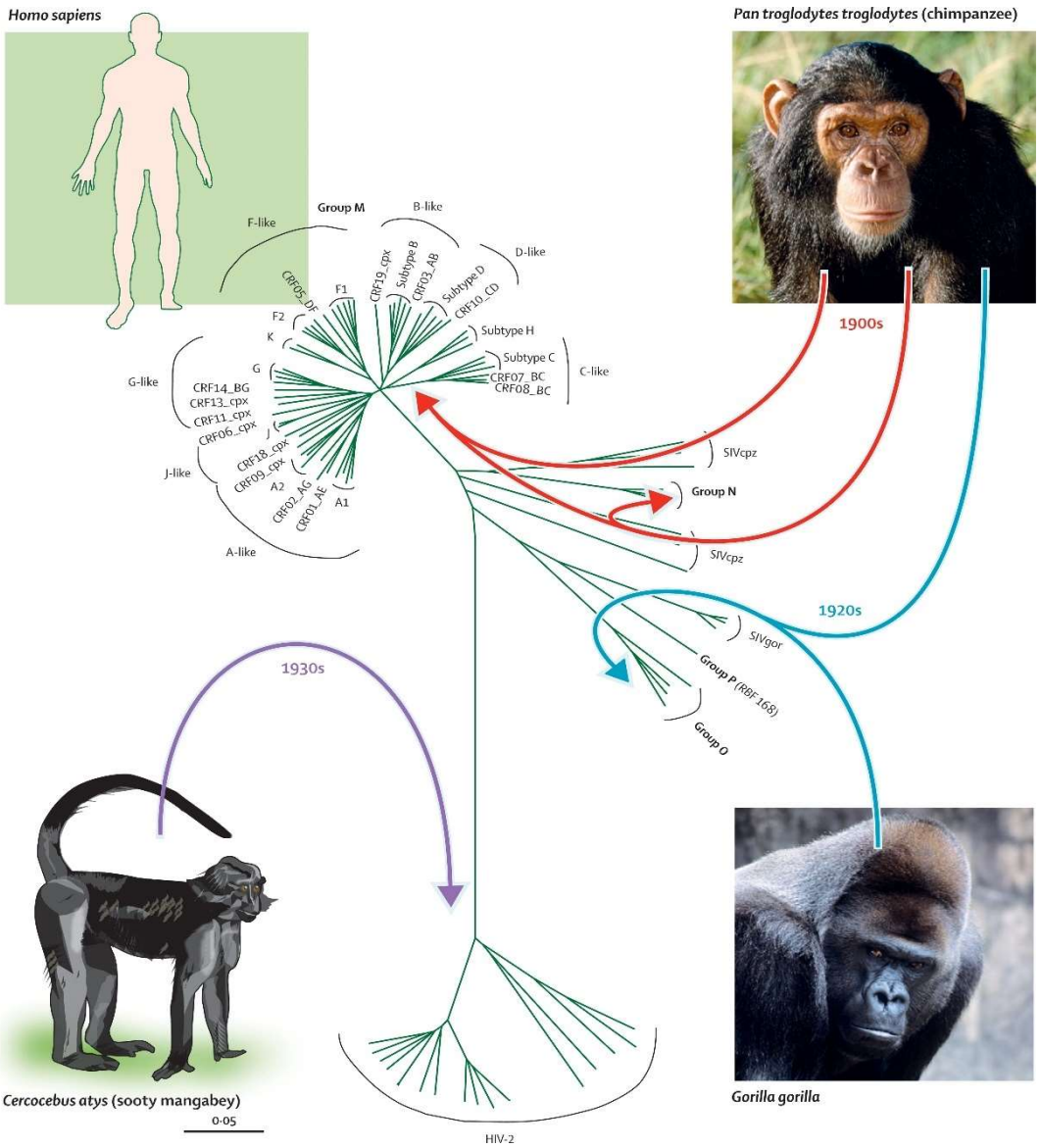


Figure 1. HIV genetic diversity and cross-species transmission events. Evolutionary relation between different SIV and HIV (reprinted from Tebit and Arts, 2011 [20]).1

Although infection of the pandemic group M was first reported in the early 1980s, genetic evidence supports events of group M infection around the final years of the 1950s decade in the

1 Reprinted from The Lancet Infectious Diseases, 11, Denis M. Tebit and Eric J. Arts, Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease, 45-56, Copyright (2011), with permission from Elsevier.

Democratic Republic of Congo [21,22]. In fact, phylogenetic studies suggest that HIV-1 group M originated in Kinshasa on early twentieth century, with a relative slow exponential growth till 1960, where it transitioned to a faster exponential growth phase, expanding to outside of the Democratic Republic of Congo [22,23]. The genetic diversification and the introduction of the virus in other geographic regions, resulted on the heterogeneous distribution of the ten HIV-1 group M subtypes (A-D, F-H, J-L), with ones comprising sub-subtypes (A1, A2, A3, A4, A6 and F1, F2), and more than a hundred circulating recombinant forms (CRFs) [24,25]. The main factors behind the group M extended genetic diversity are believed to be the high replication levels using an error-prone reverse transcriptase (RT) and the recombination events that can take place during replication [26]. Remarkably, the genetic distance between subtypes can vary between 25 and 35% and 15 to 20% within a subtype [27].

Although HIV-1 subtype B is the most disseminated variant of the virus, being the main responsible for the western and central Europe, America and Oceania epidemics, subtype C is the most prevalent subtype, accounting for almost half of all HIV-1 infections worldwide (Figure 2) [28–30]. Different HIV-1 subtypes present variances on cell tropism, viral fitness and plasma viral loads, promoting differences on disease progression dynamics and transmission rates, and also on the response to antiretroviral treatment (ART), which can be a major obstacle on the development of an effective therapy [31–36].

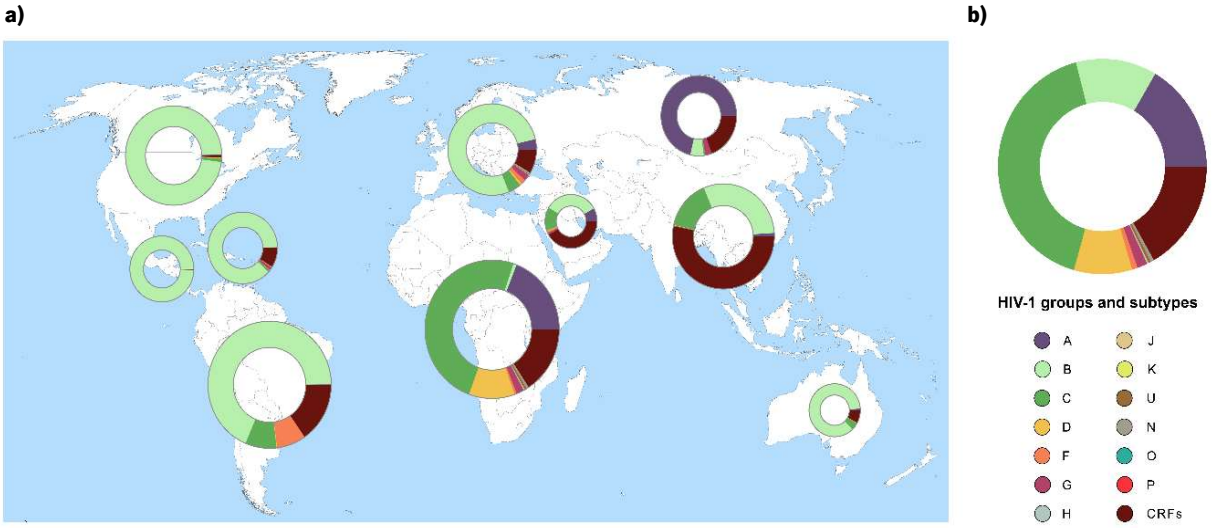


Figure 2. HIV-1 groups and subtypes distribution worldwide. (a) Frequencies of HIV-1 subtypes according to the sequences (n = 815,431) obtained from the Los Alamos HIV database (http://www.hiv.lanl.gov/), across 10 geographic regions [24]. Pie chart size is related with HIV

prevalence in each region (data obtained from the Joint United Nations Programme on HIV/AIDS (UNAIDS) data report from 2020 [9]) and the geographic regions are also according to Los Alamos HIV database. The global frequency of HIV-1 subtypes (b) was estimated by extrapolating the total number of sequences to the real number of reported cases number of people living with HIV-1 by continent (obtained from the UNAIDS report [9]). (Adapted from Santos-Pereira et al. [28]).

1.1.1.1. Phylogenetic analysis and phylogeography

As previously stated, the genetic variability among HIV-1 subtypes translates in differences on infection and disease progression, highlighting the importance of molecular epidemiology analysis to better understand the evolutionary dynamics of HIV-1 epidemics. With phylogenetic analysis it is possible to evaluate the genealogical relationship of viral strains, allowing the reconstruction of transmission chains and clusters, that provide valuable insights about the evolutionary history of the virus. There are several phylogenetic methods, although statistical phylogenetics, that include maximum-likelihood method - considering the phylogenetic tree with best fitting likelihood - and Bayesian phylogenetics – based on Markov chain Monte Carlo (MCMC) for calculation of the posterior probability - , are the most used when studying the HIV epidemiology, being highly accurate [37–39].

Moreover, software such as BEAST rely on molecular clock models that allow the estimation of time-scaled trees and infer the most recent common ancestor of the sample population. Information regarding the geographic localization of the sequences can be included in the BEAST analysis, being possible to analyze the phylogeography of the time-scaled tree [39–41].

1.1.2. HIV-1 virus and life cycle

As retrovirus, HIV-1 and HIV-2 present a similar basic structure, although some differences can be found on their genome organization [42]. HIV genome is composed by two identical single-stranded RNA molecules, protected inside a capsid (Figure 3). As a provirus, HIV RNA molecules are posteriorly converted to double-stranded DNA by RT and integrated in the human genome.

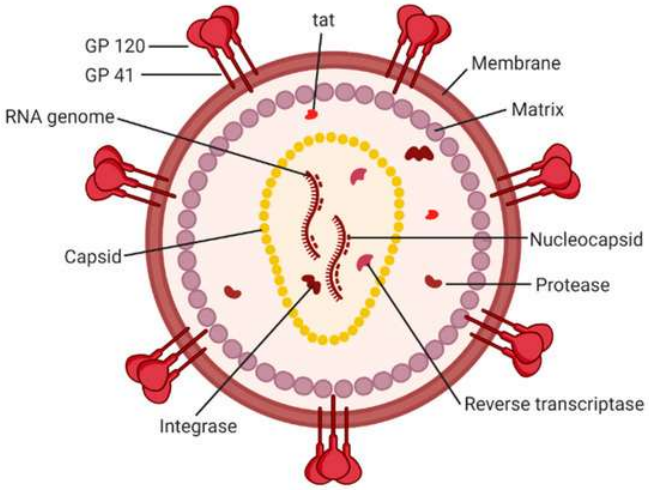


Figure 3. HIV virion. Structure of the HIV-1 virus (reprinted from Rossi *et al.*, 2021 [43]).

Both ends of the HIV DNA genome comprise long terminal repeat (LTR) sequences. The 5'LTR region is followed by the *gag* reading frame, that encodes the structural proteins of the matrix (p17), of the capside (p24), the nucleocapsid (p7) and p6 protein. The *pol* reading frame, that follows the *gag* gene, codes for the enzymes involved on viral replication including the protease (PR), RT and integrase. The following *env* gene encodes the external gp120 glycoprotein and transmembrane gp41 glycoprotein. Moreover, the HIV-1 genome also codes for other proteins crucial HIV gene expression, such as Tat (transactivator protein) and Rev (RNA splicing-regulator) regulatory factors, and other regulatory proteins including the Nef, Vif, Vpr and Vpu, being the last one replaced for Vpx in case of HIV-2 [42,44].

HIV-1 replication cycle begins when HIV gp120 binds with the host CD4 receptors present in the T lymphocytes, but also found in monocytes, dendritic cells and macrophages, followed by an interaction with a second co receptor, usually the CC-chemokine receptor 5 (CCR5) or the CXC-chemokine receptor 4 (CXCR4) (Figure 4). Different HIV-1 variants preferential select one or the other chemokine receptor, although some can use both. Upon the attachment to the host cell, a fusion peptide of gp41 is inserted in the host membrane, which leads to the fusion of the viral and host membranes. The uncoating of the viral core to the host cell cytoplasm occurs, followed by the reversed transcription of the HIV-1 RNA to double stranded DNA. The proviral DNA is then transported to the cell nucleus where is randomly inserted at the host genome by the integrase, establishing the HIV-1 infection (Figure 4) [2,42–46].

The expression of the proviral DNA depends on the host cell activated status, as latent cells function as cellular reservoirs of the virus. RNA polymerase II is responsible for the HIV-1 DNA transcription, resulting, initially, on the synthesis of regulatory proteins including Tat, which maximizes transcription rates, and Rev, that regulates the production of longer RNA transcripts. These longer fragments are transported through the nuclear membrane to the cytoplasm where the viral structural proteins and enzymes are synthesized. The Env proteins migrate and incorporate the host cell membrane, as the viral replication enzymes and two full-size RNA strands form the immature core of the virus, while the core proteins assemble to form the capsid, also migrating to the cell surface. The budding of this complex forms an immature virion that will posteriorly maturate into an infectious particle (Figure 4) [42,44,47,48].

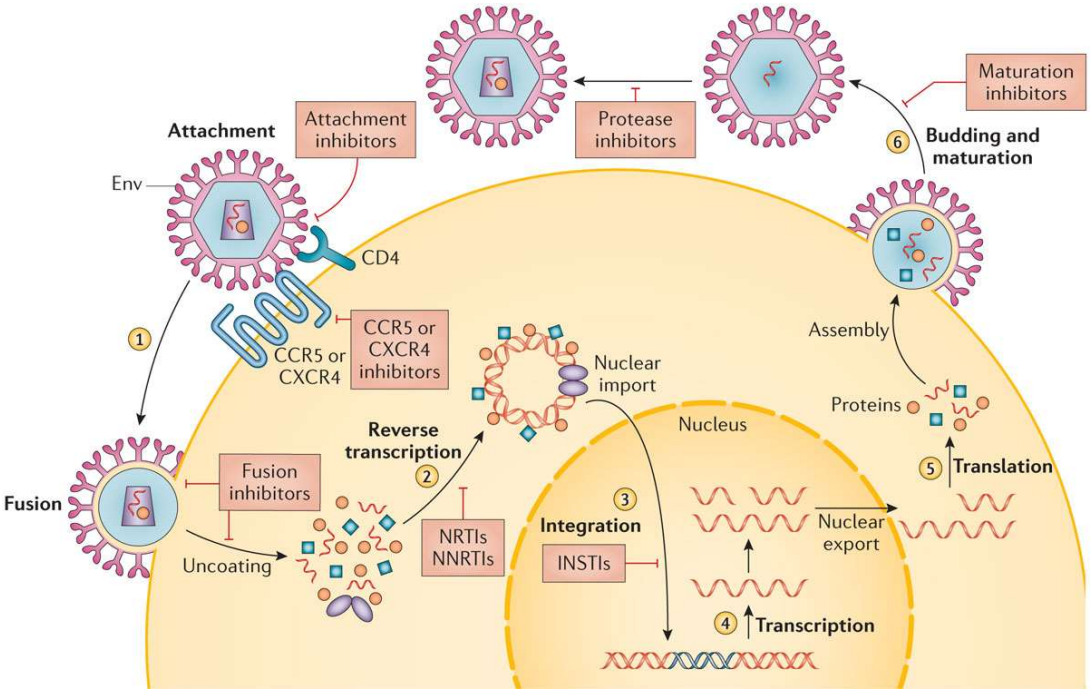


Figure 4. HIV-1 life cycle. HIV-1 enters the cell by interacting with a CD4 receptor and wither either a CCR5 or CXCR4 co-receptor. Viral binding to cell receptor is followed by membrane fusion and uncoating of the HIV-1 RNA and the RT. The double stranded resultant DNA is inserted in the host genome and transcribed to viral mRNA upon cell activation. On the cytoplasm, translation occurs, and new viral

proteins assemble and migrate to the cell surface. With budding and maturation of a new virus, HIV-1 is ready to target other cells (reprinted from Deeks *et al.*, 2015 [2])2.

1.2. HIV-1 transmission, infection and AIDS

1.2.1. HIV-1 transmission and infection

HIV-1 infection can occur following an exposure of mucosal membranes or injured skin to the virus and by parenteral inoculation. The per-act HIV transmission risk varies according to the route of infection (type of parenteral or sexual exposure and mother-to-child vertical transmission), although, in case of sexual transmission, viral loads (VL) in the moment of exposure are the key factor determining the risk of HIV-1 transmission [49–51]. Moreover, different HIV-1 subtypes also seem to present preferential transmission routes, as, for instance, subtype C has been associated with increased vaginal shedding and preferential in-utero transmission, in case of women-to-child transmission, when compared with subtype D and A, being this type of transmission also more likely to occur in case of infection with subtype D when compared with subtype A [34,52–54].

The first phase of HIV-1 infection occurs, approximately, in the first three weeks, before the virus RNA can be detected in the plasma, and it is called the eclipse phase. Being so difficult to detect HIV infection during this phase, most of the knowledge about this period comes from SIV studies in macaques and *in vitro* studies. Thus, in an initial phase, HIV-1 is believed to preferential infect cells harboring CCR5 co-receptor and within 2 days can be detected in the regional lymphatic tissue, followed by the lymph nodes, and, within 10 to 14 days, in the whole body (Figure 5) [2,44,50,55,56] .

The acute phase of infection begins with HIV-1 VL becoming detectable in the plasma, as the virus explosively replicates in the gut-associated lymphoid tissue (GALT) and other lymphoid tissues. During this stage of infection, HIV-1 RNA plasma levels reach their peak at around 10^6–10^7 copies per ml, indicating an increased the risk of transmitting the virus. With the response of the humoral immune system against HIV-1, some flu-like symptoms can occur and others that include headache, skin rash, myalgias and gastrointestinal symptoms, although many people remain asymptomatic. Also, during this phase of infection, CD4+ T cells present a dramatic depletion. which levels posteriorly recover in the

2 Reprinted by permission from Springer Nature: Nature Reviews Disease Primers, HIV infection, Steven G. Deeks et al, [COPYRIGHT] 2015

blood. The activation a virus-specific immune response, particular by action of the $CD8^+$ T-cells, leads to a decrease on VL levels to a steady-state point, called viral set point. It is during the acute phase that HIV-1 specific antibodies arise and evolve (seroconversion) (Figure 3) [2,42,44,50,55,56].

After the viral set point is established, HIV-1 continues to progressively destroy CD4+ T cells and leads to a chronic activation of the immune system, for an asymptomatic period that is known as the chronic phase of infection. In fact, this persistent and massive immune activation, characterized by an accelerated cell turnover, is positively correlated with HIV-1 infection disease progression and patient death (Figure 5) [42,51,57,58]. The virus manages to escape the immune system by the high number of escape mutations, with the continuous production of viral *quasispecies*, that keep occurring within the viral genome and by establishing cellular reservoirs in the lymph nodes, where viral replication remains at slower rates [2,42].

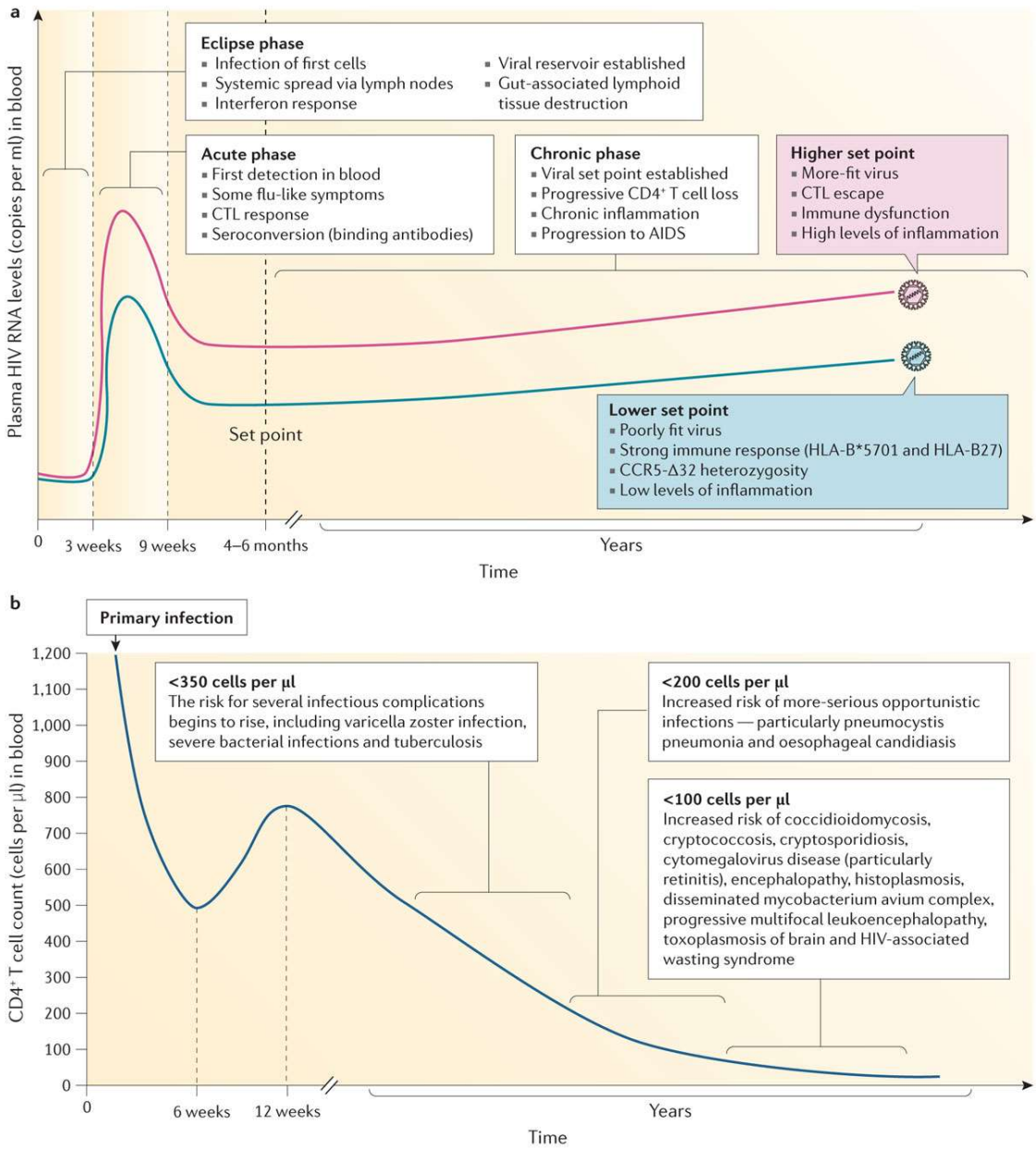


Figure 5. HIV infection and progression to AIDS. On the eclipse phase of infection, HIV-1 spreads through the organism's cells, establishing viral reservoirs. When RNA levels of HIV-1 become detectable in the blood, the acute phase beings. HIV-1 VLs exponential increase till they reach their peak, and some flu-like symptoms can be noticeable. The chronic phase of infection begins with viral RNA levels reaching a steady-state level, and, from that point, CD4 depletions slowly occurs till immunodeficiency is achieved (reprinted from Deeks *et al.*, 2015 [2])3.

3 Reprinted by permission from Springer Nature: Nature Reviews Disease Primers, HIV infection, Steven G. Deeks et al, [COPYRIGHT] 2015

1.2.2. AIDS and death

The ultimate stage of HIV-1 infection is the AIDS phase (Figure 3). According to the World Health Organization (WHO), AIDS can be defined, in adults, when the absolute and percentual (cell percentage among the total lymphocyte count) CD4+ T cell levels are below 200 cells/ μ l and <15%, respectively, or when, upon HIV-1 positive diagnosis and in the absence of CD4 counts, an AIDS related comorbidity is present [1,59]. The time of appearance of apparent symptoms of immunodeficiency can greatly vary, being tightly related with the absence of treatment and the host capacity to contain virus and to reconstitute the pool of CD4+ T cells [42,44]. Likewise, studies have pointed that 5 to 8% of HIV-1 infected people can live with no sign of infection for at least 10 years, including with no alterations in CD4+ T cell counts, and some above 35 years [60]. There is also a strong correlation between different HIV-1 subtypes and the rates of disease progression. In fact, several studies report a faster progression to AIDS and increased CD4+ T cell 4 count decline in case of subtype D infection [61–63]. Similarly, others have report that some subtypes present a reduced viral fitness with longer periods of asymptomatic infection, which is the case of subtype C when compared with D and A, and subtype A1 when compared with subtype B [35,64].

During the AIDS phase, with the severe impairment of the immune system, opportunistic infections by bacteria, viruses, fungi and parasites can develop as well as neoplasms. As CD4+ T cells counts continue to decrease and the organism weakens, AIDS is often accompanied with severe weight loss and anemia, gastrointestinal disfunction, respiratory problems, fever, fatigue, the emergence of tumors and impairment of neurologic functions.

It is estimated that, in the absence of ART, the timeframe between initial HIV-1 infection to AIDS-related death is approximately 11 years [65]. Although the widespread use of ART lead to an important decline in patients' morbidity and mortality, the rate of non-AIDS related deaths among these patients has increased, being the most common causes of non-AIDS deaths the severe infection by other infectious agents, cardiovascular and liver diseases, and cancer [66,67].

Tuberculosis (TB), a disease caused by *Mycobacterium tuberculosis* (*Mtb*) infection, is the leading cause of morbidity and mortality among people living with HIV-1 infection, accounting for 1 in 3 HIV-related deaths [68]. In 2019, the WHO defined the risk of developing TB to be 15 to 21 times higher in people infected with HIV, as studies have pointed that reduced CD4+ T cell counts in case of HIV infection

relate with a higher probability of developing active TB [68,69]. In fact, in case of co-infection, these human pathogens seem to present a synergetic relationship, promoting a faster deterioration of the immune system, leading to reduced life expectancy [68,70]. Moreover, the combination of ART and TB treatment was believed to prevent approximately 11 million deaths between 2000 and 2019 [68].

1.2.3. Antiretroviral therapy

At least 12.1 million deaths due to AIDS related illnesses were believed to be avoided, between 2010 and 2019, because of ART extended use. However, of the 38 million people living with HIV, approximately a third (12.6 million) are still waiting for treatment [9].

The available ART drugs can be classified regarding their target on the HIV-1 replicative cycle. Thus, the mostly used ART drugs can be grouped as nucleoside and non-nucleoside inhibitors of the RT (NRTIs and NNRTIs), protease inhibitors (PIs) or integrase strand-transfer inhibitors (INSTIs) (Figure 2).

Although ART remarkably improved life expectancy and quality of HIV-1 infected patients, there is still no cure for HIV infection. In 1987, Zidovudine (AZT), a NRTI, was approved as the first antiretroviral drug, strongly increasing patient's survival and life quality. However, monotherapy with the early NRTIs failed at suppressing HIV in the long term, resulting in the emergence of drug resistance mutations. With the introduction of NNRTIs and PIs during the 1990s, ART therapy dramatically changed, as combine regimens of two NRTIS with an PI or NNRTI successfully led to viral suppression, reducing morbidity and mortality [71–73] . According to the last published WHO antiretroviral therapy guidelines for adults, the first-line ART usually consists of two NRTIs and a INSTI or, in alternative a NNRTI, while the second-line comprises two NRTIs and a PIs. Currently, ART is recommended to everyone diagnosed with HIV-1 infection [74].

As stated above, ART does not fully restore patients' health and, although it increases life expectancy, this increase is accompanied by a higher rate of non-AIDS comorbidities, conceivably associated with antiretroviral therapy toxicities and with the chronic activation of the immune system resultant from the persisting viral infection [67,75]. Importantly, HIV VL can be completely reestablished upon ART interruption, which is considered to be mainly due to the existence of broad cellular reservoirs of viruses since early phases of infection. ART likely fails to eradicate HIV in all reservoirs, and HIV persists

with low levels of replication in some cells, which can be also a source for the selection of drug-resistant variants [76–78].

Although ART efficiently reduces HIV propagation, it was not capable of stopping the spreading of the HIV epidemic [74,79]. Efforts have been made worldwide towards the prevention HIV infection and, currently, the WHO recommends the use of pre-exposure prophylaxis (PrEP) treatment, containing tenofovir disoproxil fumarate (TDF), for people at substantial risk of HIV infection, reducing the risk of infection by 85% [72,74].

1.3. Drug resistance mutations

The emergence of drug resistance mutations (DRMs) is one of the major threats to an effective outcome of HIV-1 therapy. HIV-1 high recombination and mutation rates result in heterogeneous viral populations and fostering the emergence of drug resistant viral variants with a competitive advantage under ART selective pressure. These DRMs can be classified as primary, that directly lead to a reduction of ART efficacy, or accessory (or secondary) that, alongside primary mutations, help to increase resistance to therapy or compensate the fitness lost caused by primary DRMs, conferring an increased advantage to the mutant variant. However, some drugs used in ART present a higher genetic barrier than others meaning that resistance to these drugs requires a multiple combination of mutations [80–82].

The presence of DRMs can be a consequence of transmitted drug resistance, when the founding virus of infection already harbors the mutation, or acquired drug resistance, which are not present in the beginning of the treatment and emerge upon ART selective pressure. Acquired DRMs can be detected in 50% to 90% of the ART failing patients [83]. Moreover, the genetic variability among HIV-1 subtypes have been associated with different patterns of DRMs [84–86].

As DRMs have been associated with increased mortality, infection rates and treatment costs, the WHO strongly recommends the prevention and monitoring of DRMs at a population level [87,88]. In fact, several high-income countries made HIV DRMs tests prior to treatment initiation the general practice, contrasting with some low- and middle-income countries in which DMRs test is not universally performed. Moreover, The WHO also outlined a list of surveillance drug resistance mutations (SDRMs) with the purpose of facilitating ongoing and future studies of transmitted drug resistance (Table 1) [89].

Table 1. Surveillance Drug Resistance Mutations. (Extracted from Stanford Drug Resistance Database - https://hivdb.stanford.edu/)

NRTI		NNRTI		PI	
M41	L	**L100**	I	**L23**	I
K65	R	**K101**	E, P	**L24**	I
D67	N, G, E	**K103**	N, S	**D30**	N
T69	D, Ins	**V106**	M, A	**V32**	I
K70	R, E	**V179**	F	**M46**	I, L
L74	V, I	**Y181**	C, I, V	**I47**	V, A
V75	M, T, A, S	**Y188**	L, H, C	**G48**	V, M
F77	L	**G190**	A, S, E	**I50**	V, L
Y115	F	**P225**	H	**F53**	L, Y
F116	Y	**M230**	L	**I54**	V, L, M, A, T, S
Q151	M			**G73**	S, T, C, A
M184	V, I			**L76**	V
L210	W			**V82**	A, T, F, S, C, M, L
T215	Y, F, I, S, C, D, V, E			**N83**	D
K219	Q, E, N, R			**I84**	V, A, C
				I85	V
				N88	D, S
				L90	M

1.3.1. NRTI Resistance Mutations

NRTIs represent the foundation of ART, being the most widely used drugs for the treatment of HIV infection. However, all of drugs within this class can select DRMs, that can happen by two main molecular mechanisms: by thymidine analog-associated mutations (TAMs) that increases the removal of the NRTI from the end 3' of the viral DNA, or by increasing natural substrate selective binding preventing NRTI binding [80,90,91]. Although TAMs seem to be exclusively selected by AZT and stavudine (d4T), they also can confer resistance to other NRTI, being part of this group the M41L, D67N, K70R, L210W, T215F/Y and K219Q/E mutations in RT. On the other hand, the most commonly occurring mutations, that are not comprised on the TAM group, include M184V/I, that are selected and promote viral resistance

to lamivudine (3TC) and emtricitabine (FTC) but also confers higher susceptibility to AZT and TDF, and K65R, that reduces susceptivity to most NRTIs but hypersensitizes the virus to AZT [80,90–93].

1.3.2. NNRTI Resistance Mutations

NNRTIs are usually administered in combination with NRTIs. In fact, the WHO recommends the use of efavirenz (EFV) combined with TDF and 3TC as alternative first-line regimen. However, giving the notable increase of pretreatment resistance to NNRTIs in low- and middle-income countries, the WHO also strongly advises a change to non-NNRTI regimens in countries where pretreatment resistance to EFV or nevirapine (NVP) is above 10% [94,95].

NNRTIs typically present a low genetic barrier to resistance and DRM can simply confer resistance to this group of drugs by inhibiting or disrupting binding. Remarkably, nearly 70% of the mutations that occur on RT within the NNRTI binding site can induce resistance [90].

1.3.3. PI Resistance Mutations

Atazanavir (ATV), lopinavir (LPV) and darunavir (DRV), combined with ritonavir, a booster of these drugs' plasma concentrations, are the most widely used PIs, being highly effective at viral suppression and usually presenting a higher genetic barrier when compared with other types of ART. Thus, LPV or ATV in combination with NRTIs are the second-line recommended regiment by the WHO [95]. However, although these agents usually need multiple resistance mutations, drug resistance to all the existent PIs has been observed. DRMs can occur in the active site of PIs binding, being called primary mutations, or outside the binging site, increasing viral resistance, being entitled accessory mutations. Different PIs have been found to promote distinct patterns of mutations and clinical resistance have been associated with more than a hundred protease mutations [80,90,96].

1.3.4. INSTIs Resistance Mutations

In 2018, the WHO changed its preferred first line regimen for adults and adolescents to a combination of TDF, 3TC (or FTC) and dolutegravir (DTG), an INSTI. INSTIs are found be to highly

effective, presenting a low toxicity and low costs. As DTG utilization was only approved in 2013, not many cases of INSTI resistance have been reported in the population. However, many DRMs have been observed to decrease INSTI efficacy, manly reported in case of raltegravir (RAL) (approved in 2007) and elvitegravir (EVG) (approved in 2012). Recently, the first SDRM list of 24 non-polymorphic INSTI-selected mutations was published [96–99].

1.3.5. Entry Inhibitors Resistance Mutations

Maraviroc was the first CCR5 antagonist to be approved for clinical use. This class of entry inhibitors aim at preventing HIV-1 infection by binding to CCR5 receptors and by altering their conformation, avoiding CD4-T cells infection by R5-tropic virus. However, treatment failure can occur if the virus is able to use CXCR4 receptor and in case of emergence of mutations in the V3 loop that allow the virus to bind to the inhibited receptor. Moreover, *in vitro* studies also demonstrated the presence of mutations in viral gp41 upon CCR5 inhibitors selection [100–105]. Currently there is still no CXCR4 inhibitors used for HIV-1 therapy in the clinical practice but resistance mutations for AMD3100 CXCR4 antagonist have been found *in vitro* [106,107].

1.4 HIV epidemics in Brazil

Contrasting with the global trends, Latin America presented an increase of 21% of new HIV-infections, between 2010 and 2019, with Brazil accounting for almost half of Latin America infections. In 2020, it was estimated that 0.9 million people were living with HIV infection in this country, including the approximately 48.000 new reported cases of infections. In the same year, 13.000 AIDS-related deaths were estimated to be reported although 70% of the individuals living with HIV were projected to be on treatment [108]. In fact, Brazil was the first middle-income country to guarantee free access to ART to all HIV infected patients, through the public Unified Health System, being also the only country in South America that provides pre-exposure prophylaxis (PrEP) [9,109]. With a total area as vast as 8.51 million km^2, HIV-1 subtypes and recombinant forms present a heterogeneous distribution across the country, with B subtype dominating the epidemics, followed by subtype C, being the most prevalent subtype in the South region, and subtype F1, accounting for almost 10% of all infections [110,111].

Aims of the thesis

As stated above, HIV-1 pandemics continues to proliferate, being one of the top ten leading causes of death in low-income countries. Until now, there is no effective cure for this viral infection and many cases of treatment failure continue to be reported, mainly due to a deficient adherence to treatment and the emergence of DRMs.

Although new cases of infection are globally decreasing, in Latin America these new infections presented an increase of 21% since 2010. In fact, since 1980, and until June of 2020, more than a million cases of AIDS were reported in Brazil, accounting for almost half of Latin America infections. Only in 2019, 41.909 new cases of HIV infection and a death rate of 4.1 per 100.000 individuals were reported, according to the Epidemiological Bulletin from the Health Surveillance Department.

Since this country comprises the largest number of people living with HIV in Latin America, being one of the 15 countries that represents 75% of global people living with HIV in 2014, we believed that understanding the HIV-1 epidemics in Brazil could give us valuable insights about HIV-1 pandemic worldwide and provide useful information for stablishing new guidelines for treatment according to the genetic profile of the virus. Thus, with this thesis we aimed at:

- **Understanding the dynamics of DRMs in HIV-1 infected individuals under ART in Brazil;**

- **Investigate the evolutionary dynamics of HIV-1 subtypes in the same country.**

This experimental work chapter will be followed by a general discussion highlighting the key points of these studies.

Bibliography

1. Acharya, A.; Vaniawala, S. Human Immunodeficiency Virus: Discovery to Drug Resistance - A Review Update. *Biol. Syst. Open Access* **2015**, *05*, doi:10.4172/2329-6577.1000154.

2. Deeks, S.G.; Overbaugh, J.; Phillips, A.; Buchbinder, S. HIV infection. *Nat. Rev. Dis. Prim.* **2015**, *1*, 1–22, doi:10.1038/nrdp.2015.35.

3. Gottlieb, M.S.; Schroff, R.; Schanker, H.M.; Weisman, J.D.; Fan, P.T.; Wolf, R.A.; Saxon, A. Pneumocystis carinii Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men . *N. Engl. J. Med.* **1981**, *305*, 1425–1431, doi:10.1056/nejm198112103052401.

4. Hymes, K.B.; Greene, J.B.; Marcus, A.; William, D.C.; Cheung, T.; Prose, N.S.; Ballard, H.; Laubenstein, L.J. Kaposi's sarcoma in homosexual men. A report of eight cases. *Lancet* **1981**, *318*, 598–600, doi:10.1016/S0140-6736(81)92740-9.

5. Gallo, R.C.; Sarin, P.S.; Gelmann, E.P.; Robert-Guroff, M.; Richardson, E.; Kalyanaraman, V.S.; Mann, D.; Sidhu, G.D.; Stahl, R.E.; Zolla-Pazner, S.; et al. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science (80-.).* **1983**, *220*, 865–867, doi:10.1126/science.6601823.

6. Barré-Sinoussi, F.; Chermann, J.C.; Rey, F.; Nugeyre, M.T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axler-Blin, C.; Vézinet-Brun, F.; Rouzioux, C.; et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science (80-.).* **1983**, *220*, 868–871, doi:10.1126/science.6189183.

7. J, C.; A, H.; JA, L.; L, M.; S, O.; N, T.; H, T.; K, T.; H, V.; P, V. What to call the AIDS virus? *Nature* **1986**, *321*, 10, doi:10.1038/321010a0.

8. Marx, J.L. Strong new candidate for AIDS agent. *Science (80-.).* **1984**, *224*, 475–477, doi:10.1126/science.6324344.

9. Joint United Nations Programme on HIV/AIDS *UNAIDS Data 2020*; 2020;

10. WHO *Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016. Geneva, World Health Organization, 2018.*; 2018;

11. Peeters, M.; Jung, M.; Ayouba, A. The origin and molecular epidemiology of HIV. *Expert Rev. Anti.*

Infect. Ther. 2013, *11*, 885–896.

12. Hemelaar, J. The origin and diversity of the HIV-1 pandemic. *Trends Mol. Med.* 2012, *18*, 182–192. doi: 10.1586/14787210.2013.825443

13. Sousa, J.D.; Müller, V.; Vandamme, A.M. The epidemic emergence of HIV: What novel enabling factors were involved? *Future Virol.* 2017, *12*, 685–707.

14. Gao, F.; Balles, E.; Robertson, D.L.; Chen, Y.; Rodenburg, C.M.; Michael, S.F.; Cummins, L.B.; Arthur, L.O.; Peeters, M.; Shaw, G.M.; et al. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* **1999**, *397*, 436–441, doi:10.1038/17130.

15. Keele, B.F.; Van Heuverswyn, F.; Li, Y.; Bailes, E.; Takehisa, J.; Santiago, M.L.; Bibollet-Ruche, F.; Chen, Y.; Wain, L. V.; Liegeois, F.; et al. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science (80-.).* **2006**, *313*, 523–526, doi:10.1126/science.1126531.

16. Vallari, A.; Holzmayer, V.; Harris, B.; Yamaguchi, J.; Ngansop, C.; Makamche, F.; Mbanya, D.; Kaptue, L.; Ndembi, N.; Gurtler, L.; et al. Confirmation of Putative HIV-1 Group P in Cameroon. *J. Virol.* **2011**, *85*, 1403–1407, doi:10.1128/jvi.02005-10.

17. D'Arc, M.; Ayouba, A.; Esteban, A.; Learn, G.H.; Boué, V.; Liegeois, F.; Etienne, L.; Tagg, N.; Leendertz, F.H.; Boesch, C.; et al. Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E1343–E1352, doi:10.1073/pnas.1502022112.

18. Hirsch, V.M.; Olmsted, R.A.; Murphey-Corb, M.; Purcell, R.H.; Johnson, P.R. An African primate lentivirus (SIVsmclosely related to HIV-2. *Nature* **1989**, *339*, 389–392, doi:10.1038/339389a0.

19. Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.; Sharp, P.M.; Shaw, G.M.; Hahn, B.H. Human infection by genetically diverse SIVSM-related HIV-2 in West Africa. *Nature* **1992**, *358*, 495–499, doi:10.1038/358495a0.

20. Tebit, D.M.; Arts, E.J. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. *Lancet Infect. Dis.* 2011, *11*, 45–56. doi: 10.1016/S1473-3099(10)70186-9

21. Gräf, T.; Delatorre, E.; Bello, G. Phylogenetics applied to the human immunodeficiency virus type 1 (HIV-1): From the cross-species transmissions to the contact network inferences. *Mem. Inst. Oswaldo Cruz* **2020**, *115*, doi:10.1590/0074-02760190461.

22. Worobey, M.; Gemmel, M.; Teuwen, D.E.; Haselkorn, T.; Kunstman, K.; Bunce, M.; Muyembe, J.J.; Kabongo, J.M.M.; Kalengayi, R.M.; Van Marck, E.; et al. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* **2008**, *455*, 661–664, doi:10.1038/nature07390.

23. Faria, N.R.; Rambaut, A.; Suchard, M.A.; Baele, G.; Bedford, T.; Ward, M.J.; Tatem, A.J.; Sousa, J.D.; Arinaminpathy, N.; Pepin, J.; et al. The early spread and epidemic ignition of HIV-1 in human populations. *Science (80-.).* **2014**, *346*, 56–61, doi:10.1126/science.1256739.

24. Abfalterer, W.; Giori, E.E.; Fischer, W.; Hraber, P.; Macke, J.; Szinger, J.J.; Wagh, K.; Yoon, H. HIV Sequence Compendium 2018. *Los Alamos Natl. Lab. HIV Seq. Database* **2018**.

25. Yamaguchi, J.; Vallari, A.; McArthur, C.; Sthreshley, L.; Cloherty, G.A.; Berg, M.G.; Rodgers, M.A. Brief Report: Complete Genome Sequence of CG-0018a-01 Establishes HIV-1 Subtype L. *J. Acquir. Immune Defic. Syndr.* **2020**, *83*, 319–322, doi:10.1097/QAI.0000000000002246.

26. Santoro, M.M.; Perno, C.F. HIV-1 Genetic Variability and Clinical Implications. *ISRN Microbiol.* **2013**, *2013*, 1–20, doi:10.1155/2013/481314.

27. Taylor, B.S.; Sobieszczyk, M.E.; McCutchan, F.E.; Hammer, S.M. The Challenge of HIV-1 Subtype Diversity. *N. Engl. J. Med.* **2008**, *358*, 1590–1602, doi:10.1056/NEJMra0706737.

28. Santos-Pereira, A.; Magalhães, C.; Araújo, P.M.M.; Osório, N.S. Evolutionary genetics of mycobacterium tuberculosis and HIV-1: "the tortoise and the hare." *Microorganisms* **2021**, *9*, 1–15, doi:10.3390/microorganisms9010147.

29. Junqueira, D.M.; de Matos Almeida, S.E. HIV-1 subtype B: Traces of a pandemic. *Virology* 2016, *495*, 173–184, doi: 10.1016/j.virol.2016.05.003

30. Hemelaar, J.; Elangovan, R.; Yun, J.; Dickson-Tetteh, L.; Fleminger, I.; Kirtley, S.; Williams, B.; Gouws-Williams, E.; Ghys, P.D.; Abimiku, A.G.; et al. Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *Lancet Infect. Dis.* **2019**, *19*, 143–155, doi:10.1016/S1473-3099(18)30647-9.

31. Kiwanuka, N.; Laeyendecker, O.; Robb, M.; Kigozi, G.; Arroyo, M.; McCutchan, F.; Eller, L.A.; Eller, M.; Makumbi, F.; Birx, D.; et al. Effect of human immunodeficiency virus type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. *J. Infect. Dis.* **2008**, *197*, 707–713, doi:10.1086/527416.

32. Taylor, B.S.; Sobieszczyk, M.E.; McCutchan, F.E.; Hammer, S.M. The Challenge of HIV-1 Subtype Diversity. *N. Engl. J. Med.* **2008**, *358*, 1590–1602, doi:10.1056/nejmra0706737.

33. Sui, H.; Gui, T.; Jia, L.; Guo, W.; Han, J.; Liu, Y.; Bao, Z.; Li, H.; Li, J.; Li, L. Different frequencies of drug resistance mutations among HIV-1 subtypes circulating in China: A comprehensive study. *PLoS One* **2014**, *9*, 91803, doi:10.1371/journal.pone.0091803.

34. Yang, C.; Li, M.; Newman, R.D.; Shi, Y.P.; Ayisi, J.; Van Eijk, A.M.; Otieno, J.; Misore, A.O.; Steketee, R.W.; Nahlen, B.L.; et al. Genetic diversity of HIV-1 in western Kenya: Subtype-specific differences in mother-to-child transmission. *AIDS* **2003**, *17*, 1667–1674, doi:10.1097/00002030-200307250-00011.

35. Araújo, P.M.M.; Carvalho, A.; Pingarilho, M.; Faria, D.; Pinho, R.; Ferreira, J.; Proença, P.; Nunes, S.; Mouro, M.; Teófilo, E.; et al. Characterization of a large cluster of HIV-1 A1 infections detected in Portugal and connected to several Western European countries. *Sci. Rep.* **2019**, *9*, 1–10, doi:10.1038/s41598-019-43420-2.

36. Riemenschneider, M.; Cashin, K.Y.; Budeus, B.; Sierra, S.; Shirvani-Dastgerdi, E.; Bayanolhagh, S.; Kaiser, R.; Gorry, P.R.; Heider, D. Genotypic Prediction of Co-receptor Tropism of HIV-1 Subtypes A and C. *Sci. Rep.* **2016**, *6*, 1–9, doi:10.1038/srep24883.

37. Ciccozzi, M.; Lai, A.; Zehender, G.; Borsetti, A.; Cella, E.; Ciotti, M.; Sagnelli, E.; Sagnelli, C.; Angeletti, S. The phylogenetic approach for viral infectious disease evolution and epidemiology: An updating review. *J. Med. Virol.* **2019**, *91*, 1707–1724, doi:10.1002/JMV.25526.

38. Patiño-Galindo, J.; González-Candelas, F. Molecular evolution methods to study HIV-1 epidemics. *Future Virol.* **2018**, *13*, 399, doi:10.2217/FVL-2017-0159.

39. Holder, M.; Lewis, P.O. Phylogeny estimation: traditional and Bayesian approaches. *Nat. Rev. Genet. 2003 44* **2003**, *4*, 275–284, doi:10.1038/nrg1044.

40. Dellicour, S.; Gill, M.S.; Faria, N.R.; Rambaut, A.; Pybus, O.G.; Suchard, M.A.; Lemey, P. Relax, Keep Walking — A Practical Guide to Continuous Phylogeographic Inference with BEAST. *Mol. Biol. Evol.* **2021**, *38*, 3486–3493, doi:10.1093/MOLBEV/MSAB031.

41. Hong, S.L.; Lemey, P.; Suchard, M.A.; Baele, G. Bayesian Phylogeographic Analysis Incorporating Predictors and Individual Travel Histories in BEAST. *Curr. Protoc.* **2021**, *1*, e98,

doi:10.1002/CPZ1.98.

42. Fanales-Belasio, E.; Raimondo, M.; Suligol., B.; Buttò, S. HIV virology and pathogenetic mechanisms of infection: A brief overview. *Ann. Ist. Super. Sanita* 2010, *46*, 5–14, 10.4415/ANN_10_01_02

43. Rossi, E.; Meuser, M.E.; Cunanan, C.J.; Cocklin, S. Structure, function, and interactions of the hiv-1 capsid protein. *Life* 2021, *11*, 1–25, doi: 10.3390/life11020100

44. Seitz, R. Human Immunodeficiency Virus (HIV). *Transfus. Med. Hemotherapy* **2016**, *43*, 203–222, doi:10.1159/000445852.

45. Engelman, A.; Cherepanov, P. The structural biology of HIV-1: Mechanistic and therapeutic insights. *Nat. Rev. Microbiol.* 2012, *10*, 279–290.

46. Wilen, C.B.; Tilton, J.C.; Doms, R.W. HIV: Cell binding and entry. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, doi:10.1101/cshperspect.a006866.

47. Ferguson, M.R.; Rojo, D.R.; Von Lindern, J.J.; O'Brien, W.A. HIV-1 replication cycle. *Clin. Lab. Med.* 2002, *22*, 611–635, doi: 10.1016/s0272-2712(02)00015-x.

48. Liu, R.D.; Wu, J.; Shao, R.; Xue, Y.H. Mechanism and factors that control HIV-1 transcription and latency activation. *J. Zhejiang Univ. Sci. B* 2014, *15*, 455–465. doi: 10.1631/jzus.B1400059.

49. Patel, P.; Borkowf, C.B.; Brooks, J.T.; Lasry, A.; Lansky, A.; Mermin, J. Estimating per-act HIV transmission risk: A systematic review. *AIDS* **2014**, *28*, 1509–1519, doi:10.1097/QAD.0000000000000298.

50. Shaw, G.M.; Hunter, E. HIV transmission. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, doi:10.1101/cshperspect.a006965.

51. Maartens, G.; Celum, C.; Lewin, S.R. HIV infection: epidemiology, pathogenesis, treatment, and prevention. *Lancet* **2014**, *384*, 258–271, doi:10.1016/S0140-6736(14)60164-1.

52. Renjifo, B.; Gilbert, P.; Chaplin, B.; Msamanga, G.; Mwakagile, D.; Fawzi, W.; Essex, M. Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. *AIDS* **2004**, *18*, 1629–1636, doi:10.1097/01.aids.0000131392.68597.34.

53. Silveira, J.; Santos, A.F.; Martínez, A.M.B.; Góes, L.R.; Mendoza-Sassi, R.; Muniz, C.P.; Tupinambás, U.; Soares, M.A.; Greco, D.B. Heterosexual transmission of human

immunodeficiency virus type 1 subtype C in southern Brazil. *J. Clin. Virol.* **2012**, *54*, 36–41, doi:10.1016/j.jcv.2012.01.017.

54. John-Stewart, G.C.; Nduati, R.W.; Rousseau, C.M.; Mbori-Ngacha, D.A.; Richardson, B.A.; Rainwater, S.; Panteleeff, D.D.; Overbaugh, J. Subtype C is associated with increased vaginal shedding of HIV-1. *J. Infect. Dis.* **2005**, *192*, 492–496, doi:10.1086/431514.

55. Moir, S.; Chun, T.W.; Fauci, A.S. Pathogenic mechanisms of HIV disease. *Annu. Rev. Pathol. Mech. Dis.* **2011**, *6*, 223–248, doi:10.1146/annurev-pathol-011110-130254.

56. McMichael, A.J.; Borrow, P.; Tomaras, G.D.; Goonetilleke, N.; Haynes, B.F. The immune response during acute HIV-1 infection: Clues for vaccine development. *Nat. Rev. Immunol.* 2010, *10*, 11–23, doi: 10.1038/nri2674.

57. Ford, E.S.; Puronen, C.E.; Sereti, I. Immunopathogenesis of asymptomatic chronic HIV Infection: The calm before the storm. *Curr. Opin. HIV AIDS* 2009, *4*, 206–214, doi: 10.1097/COH.0b013e328329c68c.

58. Mogensen, T.H.; Melchjorsen, J.; Larsen, C.S.; Paludan, S.R. Innate immune recognition and activation during HIV infection. *Retrovirology* 2010, *7*, 54.

59. World Health Organization *WHO case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related disease in adults and children*; 2007;

60. Levy, J.A. Dispelling myths and focusing on notable concepts in HIV pathogenesis. *Trends Mol. Med.* 2015, *21*, 341–353, doi: 10.1016/j.molmed.2015.03.004.

61. Ssemwanga, D.; Nsubuga, R.N.; Mayanja, B.N.; Lyagoba, F.; Magambo, B.; Yirrell, D.; van der Paal, L.; Grosskurth, H.; Kaleebu, P. Effect of HIV-1 Subtypes on Disease Progression in Rural Uganda: A Prospective Clinical Cohort Study. *PLoS One* **2013**, *8*, 71768, doi:10.1371/journal.pone.0071768.

62. Easterbrook, P.J.; Smith, M.; Mullen, J.; O'Shea, S.; Chrystie, I.; De Ruiter, A.; Tatt, I.D.; Geretti, A.M.; Zuckerman, M. Impact of HIV-1 viral subtype on disease progression and response to antiretroviral therapy. *J. Int. AIDS Soc.* **2010**, *13*, doi:10.1186/1758-2652-13-4.

63. Leite, T.C.N.F.; Campos, D.P.; Coelho, A.B.; Teixeira, S.L.M.; Veloso, V.; Morgado, M.G.; Guimarães, M.L. Impact of HIV-1 Subtypes on AIDS Progression in a Brazilian Cohort. *AIDS Res.*

Hum. Retroviruses **2017**, *33*, 41–48, doi:10.1089/aid.2016.0126.

64. Venner, C.M.; Nankya, I.; Kyeyune, F.; Demers, K.; Kwok, C.; Chen, P.L.; Rwambuya, S.; Munjoma, M.; Chipato, T.; Byamugisha, J.; et al. Infecting HIV-1 Subtype Predicts Disease Progression in Women of Sub-Saharan Africa. *EBioMedicine* **2016**, *13*, 305–314, doi:10.1016/j.ebiom.2016.10.014.

65. Remien, R.H.; Mellins, C.A. Long-term psychosocial challenges for people living with HIV: Let's not forget the individual in our global response to the pandemic. *AIDS* 2007, *21*, doi: 10.1097/01.aids.0000298104.02356.b3.

66. Todd, J.; Glynn, J.R.; Marston, M.; Lutalo, T.; Biraro, S.; Mwita, W.; Suriyanon, V.; Rangsin, R.; Nelson, K.E.; Sonnenberg, P.; et al. Time from HIV seroconversion to death: a collaborative analysis of eight studies in six low and middle-income countries before highly active antiretroviral therapy. *AIDS* **2007**, *21*, S55–S63, doi:10.1097/01.aids.0000299411.75269.e8.

67. Cowell, A.; Shenoi, S. V.; Kyriakides, T.C.; Friedland, G.; Barakat, L.A. Trends in hospital deaths among human immunodeficiency virus-infected patients during the antiretroviral therapy era, 1995 to 2011. *J. Hosp. Med.* **2015**, *10*, 608–614, doi:10.1002/jhm.2409.

68. World Health Organization *Global Tuberculosis Report*; 2020;

69. Schutz, C.; Meintjes, G.; Almajid, F.; Wilkinson, R.J.; Pozniak, A. Clinical management of tuberculosis and HIV-1 co-infection. *Eur. Respir. J.* **2010**, *36*, 1460–1481, doi:10.1183/09031936.00110210.

70. Pawlowski, A.; Jansson, M.; Sköld, M.; Rottenberg, M.E.; Källenius, G. Tuberculosis and HIV Co-Infection. *PLoS Pathog.* **2012**, *8*, e1002464, doi:10.1371/journal.ppat.1002464.

71. Pau, A.K.; George, J.M. Antiretroviral therapy: Current drugs. *Infect. Dis. Clin. North Am.* 2014, *28*, 371–402, doi: 10.1016/j.idc.2014.06.001.

72. Gulick, R.M.; Flexner, C. Long-acting HIV drugs for treatment and prevention. *Annu. Rev. Med.* 2019, *70*, 137–150, doi: 10.1146/annurev-med-041217-013717.

73. Fischl, M.A.; Richman, D.D.; Grieco, M.H.; Gottlieb, M.S.; Volberding, P.A.; Laskin, O.L.; Leedom, J.M.; Groopman, J.E.; Mildvan, D.; Schooley, R.T.; et al. The Efficacy of Azidothymidine (AZT) in the Treatment of Patients with AIDS and AIDS-Related Complex. *N. Engl. J. Med.* **1987**, *317*,

185–191, doi:10.1056/NEJM198707233170401.

74. World Health Organization (WHO) *Update of recommendations on first- and second-line antiretroviral regimens*; World Health Organization (WHO), Ed.; Geneva, Switzerland, 2019;

75. Lorenc, A.; Ananthavarathan, P.; Lorigan, J.; Banarsee, R.; Jowata, M.; Brook, G. The prevalence of comorbidities among people living with hiv in brent: A diverse London Borough. *London J. Prim. Care (Abingdon).* **2014**, *6*, 84–90, doi:10.1080/17571472.2014.11493422.

76. Churchill, M.J.; Deeks, S.G.; Margolis, D.M.; Siliciano, R.F.; Swanstrom, R. HIV reservoirs: What, where and how to target them. *Nat. Rev. Microbiol.* 2015, *14*, 55–60.

77. Bandera, A.; Gori, A.; Clerici, M.; Sironi, M. Phylogenies in ART: HIV reservoirs, HIV latency and drug resistance. *Curr. Opin. Pharmacol.* 2019, *48*, 24–32.

78. Mzingwane, M.L.; Tiemessen, C.T. Mechanisms of HIV persistence in HIV reservoirs. *Rev. Med. Virol.* 2017, *27*.

79. Melo, M.G. de; Sprinz, E.; Gorbach, P.M.; Santos, B.; Rocha, T. de M.; Simon, M.; Almeida, M.; Lira, R.; Chaves, M.C.; Kerin, T.; et al. HIV-1 heterosexual transmission and association with sexually transmitted infections in the era of treatment as prevention. *Int. J. Infect. Dis.* **2019**, *87*, 128–134, doi:10.1016/j.ijid.2019.08.004.

80. Clutter, D.S.; Jordan, M.R.; Bertagnolio, S.; Shafer, R.W. HIV-1 drug resistance and resistance testing. *Infect. Genet. Evol.* 2016, *46*, 292–307.

81. Van Laethem, K.; Theys, K.; Vandamme, A.M. HIV-1 genotypic drug resistance testing: Digging deep, reaching wide? *Curr. Opin. Virol.* 2015, *14*, 16–23.

82. Iyidogan, P.; Anderson, K.S. Current perspectives on HIV-1 antiretroviral drug resistance. *Viruses* 2014, *6*, 4095–4139.

83. WHO | Global action plan on HIV drug resistance 2017–2021. *WHO* **2017**.

84. de Faria Romero Soldi, G.; Ribeiro, I.C.; Ahagon, C.M.; Coelho, L.P.O.; Cabral, G.B.; Lopes, G.I.S.L.; de Paula Ferreira, J.L.; de Macedo Brígido, L.F. Major drug resistance mutations to HIV-1 protease inhibitors (PI) among patients exposed to PI class failing antiretroviral therapy in São Paulo State, Brazil. *PLoS One* **2019**, *14*, doi:10.1371/journal.pone.0223210.

85. Smit, E.; White, E.; Clark, D.; Churchill, D.; Zhang, H.; Collins, S.; Pillay, D.; Sabin, C.; Nelson, M.; Winston, A.; et al. An association between K65R and HIV-1 subtype C viruses in patients treated with multiple NRTIs. *J. Antimicrob. Chemother.* **2017**, *72*, 2075–2082, doi:10.1093/jac/dkx091.

86. Chaplin, B.; Eisen, G.; Idoko, J.; Onwujekwe, D.; Idigbe, E.; Adewole, I.; Gashau, W.; Meloni, S.; Sarr, A.D.; Sankalé, J.L.; et al. Impact of HIV type 1 subtype on drug resistance mutations in Nigerian patients failing first-line therapy. *AIDS Res. Hum. Retroviruses* **2011**, *27*, 71–80, doi:10.1089/aid.2010.0050.

87. World Health Organization *HIV Drug Resistance Report 2019*; 2019;

88. Phillips, A.N.; Stover, J.; Cambiano, V.; Nakagawa, F.; Jordan, M.R.; Pillay, D.; Doherty, M.; Revill, P.; Bertagnolio, S. Impact of HIV drug resistance on HIV/AIDS-associated mortality, new infections, and antiretroviral therapy program costs in Sub-Saharan Africa. *J. Infect. Dis.* **2017**, *215*, 1362–1365, doi:10.1093/infdis/jix089.

89. Bennett, D.E.; Camacho, R.J.; Otelea, D.; Kuritzkes, D.R.; Fleury, H.; Kiuchi, M.; Heneine, W.; Kantor, R.; Jordan, M.R.; Schapiro, J.M.; et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. *PLoS One* **2009**, *4*, 4724, doi:10.1371/journal.pone.0004724.

90. Weber, I.T.; Harrison, R.W. Decoding HIV resistance: From genotype to therapy. *Future Med. Chem.* 2017, *9*, 1529–1538.

91. Rai, M.A.; Pannek, S.; Fichtenbaum, C.J. Emerging reverse transcriptase inhibitors for HIV-1 infection. *Expert Opin. Emerg. Drugs* 2018, *23*, 149–157.

92. Brenner, B.G.; Coutsinos, D. The K65R mutation in HIV-1 reverse transcriptase: Genetic barriers, resistance profile and clinical implications. *HIV Ther.* 2009, *3*, 583–594.

93. Shafer, R.W.; Schapiro, J.M. HIV-1 drug resistance mutations: An updated framework for the second decade of HAART. *AIDS Rev.* 2008, *10*, 67–84.

94. Rhee, S.Y.; Kassaye, S.G.; Barrow, G.; Sundaramurthi, J.C.; Jordan, M.R.; Shafer, R.W. HIV-1 transmitted drug resistance surveillance: shifting trends in study design and prevalence estimates. *J. Int. AIDS Soc.* 2020, *23*.

95. WHO Updated Recommendations on First-Line and Second-Line Antiretroviral Regimens and Post-Exposure Prophylaxis and Recommendations on Early Infant Diagnosis of HIV. *WHO Guidel.* **2018**.

96. Baxter, J.D.; Chasanov, W.M. An Update on HIV-1 Protease Inhibitor Resistance. *J. AIDS Clin. Res.* **2016**, *7*, doi:10.4172/2155-6113.1000581.

97. Tzou, P.L.; Rhee, S.Y.; Descamps, D.; Clutter, D.S.; Hare, B.; Mor, O.; Grude, M.; Parkin, N.; Jordan, M.R.; Bertagnolio, S.; et al. Integrase strand transfer inhibitor (INSTI)-resistance mutations for the surveillance of transmitted HIV-1 drug resistance. *J. Antimicrob. Chemother.* **2020**, *75*, 170–182, doi:10.1093/jac/dkz417.

98. Rhee, S.Y.; Grant, P.M.; Tzou, P.L.; Barrow, G.; Harrigan, P.R.; Ioannidis, J.P.A.; Shafer, R.W. A systematic review of the genetic mechanisms of dolutegravir resistance. *J. Antimicrob. Chemother.* 2019, *74*, 3135–3149, doi: 10.1093/jac/dkz256.

99. Quashie, P.K.; Mesplède, T.; Wainberg, M.A. Evolution of HIV integrase resistance mutations. *Curr. Opin. Infect. Dis.* 2013, *26*, 43–49, doi: 10.1097/QCO.0b013e32835ba81c

100. Wensing, A.M.; Calvez, V.; Ceccherini-Silberstein, F.; Charpentier, C.; Günthard, H.F.; Paredes, R.; Shafer, R.W.; Richman, D.D. 2019 Update of the Drug Resistance Mutations in HIV-1. *Top. Antivir. Med.* **2019**, *27*, 111.

101. Anastassopoulou, C.G.; Ketas, T.J.; Sanders, R.W.; Johan Klasse, P.; Moore, J.P. Effects of sequence changes in the HIV-1 gp41 fusion peptide on CCR5 inhibitor resistance. *Virology* **2012**, *428*, 86–97, doi:10.1016/J.VIROL.2012.03.008.

102. Woollard, S.M.; Kanmogne, G.D. Maraviroc: a review of its use in HIV infection and beyond. *Drug Des. Devel. Ther.* **2015**, *9*, 5447–5468, doi:10.2147/DDDT.S90580.

103. Latinovic, O.S.; Reitz, M.; Heredia, A. CCR5 Inhibitors and HIV-1 Infection. *J. AIDS HIV Treat.* **2019**, *1*, 1, doi:10.33696/AIDS.1.001.

104. Yuan, Y.; Yokoyama, M.; Maeda, Y.; Terasawa, H.; Harada, S.; Sato, H.; Yusa, K. Structure and Dynamics of the gp120 V3 Loop That Confers Noncompetitive Resistance in R5 HIV-1JR-FL to Maraviroc. *PLoS One* **2013**, *8*, e65115, doi:10.1371/JOURNAL.PONE.0065115.

105. Ratcliff, A.N.; Shi, W.; Arts, E.J. HIV-1 resistance to maraviroc conferred by a CD4 binding site

mutation in the envelope glycoprotein gp120. *J. Virol.* **2013**, *87*, 923–934, doi:10.1128/JVI.01863-12.

106. Hatse, S.; Princen, K.; Vermeire, K.; Gerlach, L.O.; Rosenkilde, M.M.; Schwartz, T.W.; Bridger, G.; De Clercq, E.; Schols, D. Mutations at the CXCR4 interaction sites for AMD3100 influence anti-CXCR4 antibody binding and HIV-1 entry. *FEBS Lett.* **2003**, *546*, 300–306, doi:10.1016/S0014-5793(03)00609-4.

107. Maeda, Y.; Terasawa, H.; Nakano, Y.; Monde, K.; Yusa, K.; Oka, S.; Takiguchi, M.; Harada, S. V3-Independent Competitive Resistance of a Dual-X4 HIV-1 to the CXCR4 Inhibitor AMD3100. *PLoS One* **2014**, *9*, e89515, doi:10.1371/JOURNAL.PONE.0089515.

108. United Nations Programme on HIV/aids. UNAIDS UNAIDS data 2021. **2021**, 4–38.

109. Ministério da Saúde *Clinical Protocol and Therapeutic Guidelines for Handling HIV infection in adults.*; 1st ed.; Brasilia/DF, 2018; ISBN 9788533426436.

110. Arruda, M.B.; Boullosa, L.T.; Cardoso, C.C.; da Costa, C.M.; Brites, C.; de Lima, S.T.S.; Kaminski, H.T.; Aleixo, A.W.; Esposito, A.O.P.; Cavalcanti, A.M.S.; et al. Brazilian network for HIV Drug Resistance Surveillance (HIV-BresNet): a survey of treatment-naive individuals. *J. Int. AIDS Soc.* **2018**, *21*, doi:10.1002/JIA2.25032.

111. Gräf, T.; Bello, G.; Andrade, P.; Arantes, I.; Pereira, J.M.; da Silva, A.B.P.; Veiga, R. V.; Mariani, D.; Boullosa, L.T.; Arruda, M.B.; et al. HIV-1 molecular diversity in Brazil unveiled by 10 years of sampling by the national genotyping network. *Sci. Reports 2021 111* **2021**, *11*, 1–9, doi:10.1038/s41598-021-94542-5.

Chapter 2.1. - Nationwide study of drug resistance mutations in HIV-1 infected individuals under antiretroviral therapy in Brazil

This work was published as:

Santos-Pereira, A.; Triunfante, V.; Araújo, P.M.M.; Martins, J.; Soares, H.; Poveda, E.; Souto, B.; Osório, N.S. Nationwide Study of Drug Resistance Mutations in HIV-1 Infected Individuals under Antiretroviral Therapy in Brazil. Int. J. Mol. Sci. 2021, 22, 5304. *https://doi.org/10.3390/ijms22105304*

Nationwide Study of Drug Resistance Mutations in HIV-1 Infected Individuals under Antiretroviral Therapy in Brazil

1. Abstract

The success of antiretroviral treatment (ART) is threatened by the emergence of drug resistance mutations (DRM). Since Brazil presents the largest number of people living with HIV (PLWH) in South America we aimed at understanding the dynamics of DRM in this country. We analyzed a total of 20,226 HIV-1 sequences collected from PLWH undergoing ART between 2008–2017. Results show a mild decline of DRM over the years but an increase of the K65R reverse transcriptase mutation from 2.23% to 12.11%. This increase gradually occurred following alterations in the ART regimens replacing zidovudine (AZT) with tenofovir (TDF). PLWH harboring the K65R had significantly higher viral loads than those without this mutation ($p < 0.001$). Among the two most prevalent HIV-1 subtypes (B and C) there was a significant ($p < 0.001$) association of K65R with subtype C (11.26%) when compared with subtype B (9.27%). Nonetheless, evidence for K65R transmission in Brazil was found both for C and B subtypes. Additionally, artificial neural network-based immunoinformatic predictions suggest that K65R could enhance viral recognition by HLA-B27 that has relatively low prevalence in the Brazilian population. Overall, the results suggest that tenofovir-based regimens need to be carefully monitored particularly in settings with subtype C and specific HLA profiles.

2. Introduction

According to the Joint United Nations Programme on HIV/AIDS (UNAIDS), 690,000 individuals died in 2019 of human immunodeficiency virus (HIV) infection-related causes. Although the global number of new infections (1.7 million) has been reduced by 23%, since 2010, Latin America has presented a 21% increase. Particularly, Brazil experienced an increase of 17%, with 48,000 new reported infections and 14,000 acquired immunodeficiency syndrome (AIDS)-related deaths, in the same year [1].

Antiretroviral therapy (ART) significantly improves patients' survival, increasing life expectancy and quality [2–4]. Furthermore, it plays a relevant role in the prevention of HIV-1 transmission in the population [5–8]. However, the emergence of drug resistance mutations (DRM) represents a major threat for the continued control of HIV replication, and consequent potential increase in the transmission of viral strains with DRM [9,10]. The prevalence of DRM can greatly vary between different geographical areas, depending on several factors, ranging from study design to the sociodemographic characteristics of the population and most prevalent HIV-1 subtypes [9,11–15]. With a total area as vast as 8.51 million km^2, different DRM rates have also been found in distinct Brazilian geographic regions [9,12,16–20].

In 1996, Brazil became the first middle-income country to ensure free access to ART to all individuals infected with HIV, through the public Unified Health System [21,22]. Currently, 69% of the people living with HIV are under ART [1]. Initial ART regimens in Brazil consist of a combination of three drugs, comprising two NRTIs and a third of different ART class: a non-nucleoside reverse transcriptase inhibitor (NNRTI), a protease inhibitor (PI) or an integrase inhibitor (INSTI) [21]. Although pre-treatment testing for DRM in the reverse transcriptase (RT) and protease (PR) has been widely recommended by the International Antiviral Society–USA, the US Department of Health and Human Services and the European AIDS Clinical Society, among others, viral genotyping and resistance testing are preferably performed in Brazil after 6 months of treatment failure [21,23–25]. After 2013, pre-treatment genotyping was also indicated by the Brazilian Government in the following scenarios: *Mycobacterium tuberculosis* (*Mtb*) co-infection, pregnant woman, children and new diagnoses with a sexual partner under ART [21,22]. Importantly, Brazil is the only country in South America that provides pre-exposure prophylaxis (PrEP) through the public health system [1,21].

Among the drug-resistance mutations (DRM) [26], K65R is commonly associated with tenofovir (TDF) resistance and different prevalence rates of this mutation have been reported in several Brazilian cities [11,17,27]. It was also associated with resistance to other reverse transcriptase inhibitors (NRTIs), but it confers increased susceptibility to zidovudine (AZT) [28–33]. K65R is acquired in B and several other HIV-1 subtypes by the AAA→AGA substitution while in subtype C the involved substitution is AAG→AGG, which has been associated with higher probability of C subtype viruses to acquire this mutation [31,32,34,35].

The relevant reforms made in ART protocols in Brazil during the last decades, promoting the preferential use TDF containing schemes in the absence of universal HIV-1 genotyping, raised the hypothesis that the HIV DRM profile could have changed, bringing novel treatment obstacles and clinical

implications. To test this hypothesis, we retrospectively studied data from 20,226 HIV-1 infected and ART-treated individuals from all regions of Brazil, from 2008 to 2017. We used statistical and genome sequence analysis to understand, at an unprecedented scale in Brazil, the dynamics of evolution of the HIV-1 drug resistance profiles. Additionally, we also relied on immunoinformatics to investigate factors possible underlying treatment failure and transmission of DRM.

3. Results

3.1. Characterization of the Study Population

The study comprised a total of 20,226 HIV-1 infected individuals, with a HIV-1 sequence genotyped between 2008 and 2017, on ART treatment, from which 44.3% (*n* = 8962) were female and 55.7% (*n* = 11,263) male. The main characteristics of the study population are shown in Table 1. The median age was 39.55 (±12.71) years (yrs). Most of the individuals were between 30 and 49 yrs old (61.44%), from the State of São Paulo (21.98%; *n* = 4445), followed by the States of Rio Grande do Sul (11.96%; *n* = 2419), Minas Gerais (9.29%; *n* = 1878), Rio de Janeiro (9.20%; *n* = 1861) and Paraná (7.94%; *n* = 1606). The geographic frequency distribution for PLWH in the study population was in agreement with the official Brazilian HIV-1 prevalence reports [36]. These patients were on ART, in average, for 2.98 (±2.96) yrs and the most common treatment schemes combinations were: lamivudine (3TC)/efavirenz (EFV)/TDF (21.38%; *n* = 4324), 3TC/AZT/EFV (19.79%; *n* = 4003), 3TC/AZT/low-dose ritonavir-boosted lopinavir (LPV) (11.16%; *n* = 2258), 3TC/LPV/TDF (8.24%; *n* = 1666) and 3TC/atazanavir (ATV)/ritonavir (RTV)/TDF (6.70%; *n* = 1355).

Table 1. Characterization of the study population.

	HIV-1⁺ Individuals	**Female**	**Male**
All ages	20,226 (100.00%)	8962 (44.3%)	11,263 (55.7%)
< 1 yrs old	64 (0.32%)	30 (0.15%)	34 (0.17%)
2–9 yrs old	364 (1.80%)	200 (0.99%)	164 (0.81%)
10–17 yrs old	840 (4.15%)	426 (2.11%)	414 (2.05%)
18–30 yrs old	2443 (12.08%)	1201 (5.94%)	1242 (6.14%)
30–49 yrs old	12,426 (61.44%)	5390 (26.65%)	7036 (34.79%)
50–79 yrs old	4074 (20.14%)	1709 (8.45%)	2364 (11.69%)
> 80 yrs old	15 (0.07%)	6 (0.03%)	9 (0.04%)
Age (av. yrs ± std)	39.55 ± 12.71	38.67 ± 13.15	40.25 ± 12.31
Treatment (av. yrs ± std)	2.98 ± 2.96	3.06 ± 2.96	2.92 ± 2.95
Birth Federative Unit	**HIV-1⁺ individuals**	**Treatment Scheme**	**HIV-1⁺ individuals**
São Paulo	4445 (21.98%)	3TC,EFV,TDF	4324 (21.38%)
Rio Grande do Sul	2419 (11.96%)	3TC,AZT,EFV	4003 (19.79%)
Minas Gerais	1878 (9.29%)	3TC,AZT,LPV	2258 (11.16%)
Rio de Janeiro	1861 (9.20%)	3TC,LPV,TDF	1666 (8.24%)
Paraná	1606 (7.94%)	3TC,ATV,RTV,TDF	1355 (6.70%)
Others	8017 (39.64%)	Others	6620 (32.73%)

av. average; std, standard deviation; yrs, years; HIV-1⁺, HIV-1 infected.

3.2. Prevalence of Drug-Resistance Mutations

We evaluated the presence of drug-resistance mutations (DRM) in the study population, including DRM found isolated or in combinations in the same virus, and uncovered an overall prevalence of 84.10% (n = 17,011). DRM were mostly found at the RT (83.24%; n = 16,836) and included 13,845 sequences with NRTI DRM and 11,720 sequences with NNRTI DRM. The most common DRM (Figure 1,

Supplementary table S2 available at https://www.mdpi.com/1422-0067/22/10/5304) were the substitutions in RT amino acids M184V (65.53%, *n* = 13,265), K103N (40.20%, *n* = 8738), and M41L (17.21%, *n* = 3480). DRM in PR were found in 5021 (24.82%) sequences and the more frequent were V82A (9.99%, *n* = 2021), M46I (9.58%, *n* = 1938), and I54V (8.50%, *n* = 1719).

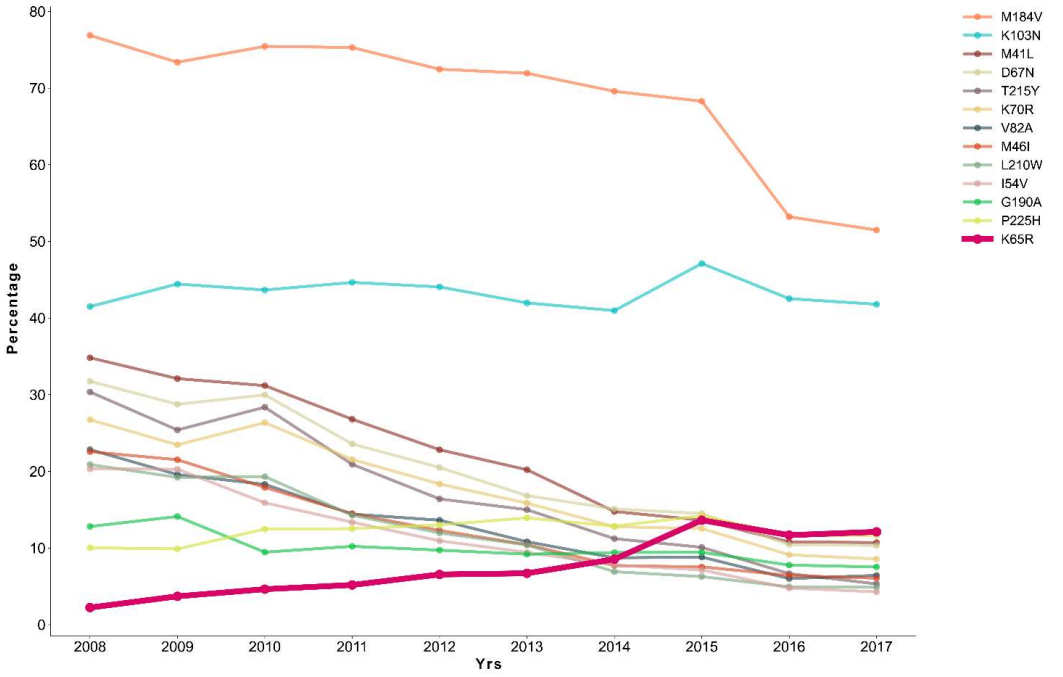


Figure 1. Drug-resistance mutations prevalence in HIV-1 RT and PR through the yrs 2008 to 2017. Percentage of individuals infected with a viral strain that presented one of the most common DRM in the different yrs.

3.3. A Gradual Alteration on the Prevalence of Drug-Resistance Mutations

To investigate the temporal dynamics and evolution of the DRM landscape in the study population we evaluated yearly DRM prevalence from 2008 until 2017. The most common mutation across the years was M184V (Figure 1), notwithstanding a decreasing trend and a significant decrease between 2015 (68.29%, *n* = 1874) and 2016 (53.21%, *n* = 2684) ($p < 0.001$) (Table 2). Similarly, all the other 10 most frequent DRM followed a decreasing trend along the years with the remarkable exception for K65R and K103N (Figure 1). K103N remained stable in the studied period. While the prevalence of viruses harboring K103N increased between 2014 (40.97%, *n* = 1180) and 2015 (47.12%, *n* = 1293; $p < 0.001$) this was

accompanied with a significant decrease in 2016 (42.53%, n = 2145; $p < 0.001$) and no significant difference between 2008 and 2017 ($p = 0.919$). Importantly, K65R showed a clear rise along the years (Figure 1). In 2008, the mutation was relatively rare and only found in 2.23% (n = 8) of the population. Subsequently, this mutation became more prevalent and increased significantly between 2013 and 2015 ($p < 0.001$). Although its prevalence suffered a decrease between 2015 and 2016 ($p = 0.0124$), K65R prevailed the third most common DRM in 2016 and 2017 (Figure 1). Overall, the data strongly supports that, in contrast with other prevalent DRM, K65R prevalence was increasing in Brazil, raising additional concerns about the efficacy of some of the ART combinations used. In line with this hypothesis was the finding that viral loads for PLWH with K65R were significantly higher than for the rest of the cohort (180,739.60 ± 497,000.10 cop/mL vs. 95,358.26 ± 338,930.67 cop/mL, respectively; $p < 0.001$).

Table 2. Yearly distribution of the most prevalent drug-resistance mutations in HIV-1 RT and PR.

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017 **
M184V	276 (76.88)	416 (73.37)	375 (75.45)	1076 (75.3)	1781 (72.46)	2082 (71.94)	2004 (69.58)	1874 (68.29)	2684 * (53.21)	697 (51.48)
K103N	149 (41.5)	252 (44.44)	217 (43.66)	638 (44.65)	1083 (44.06)	1215 (41.98)	1180 (40.97)	1293 * (47.12)	2145 * (42.53)	566 (41.8)
M41L	125 (34.82)	182 (32.1)	155 (31.19)	383 (26.8)	561 * (22.28)	585 * (20.21)	425 * (14.76)	374 (13.63)	545 * (10.8)	145 (10.71)
D67N	114 (31.75)	163 (28.75)	149 (29.98)	337 * (23.58)	504 * (20.5)	487 * (16.83)	435 (15.1)	398 (14.5)	526 * (10.43)	140 (10.34)
T215Y	109 (30.36)	144 (25.4)	141 (28.37)	299 * (20.92)	403 * (16.4)	434 (15)	323 * (11.22)	277 (10.09)	336 * (6.66)	72 (5.32)
K70R	96 (26.74)	133 (23.46)	131 (26.36)	308 * (21.55)	451 * (18.35)	459 * (15.86)	368 * (12.78)	344 (12.54)	460 * (9.12)	116 (8.57)
V82A	82 (22.84)	111 (19.58)	91 (18.31)	206 * (14.42)	335 (13.63)	313 * (10.82)	251 * (8.72)	242 (8.82)	303 * (6.01)	87 (6.43)
M46I	81 (22.56)	122 (21.52)	89 (17.91)	207 (14.49)	302 (12.29)	302 * (10.44)	222 * (7.71)	207 (7.54)	324 (6.42)	82 (6.06)
L210W	75 (20.89)	109 (19.22)	96 (19.32)	204 * (14.28)	294 * (11.96)	301 (10.4)	199 * (6.91)	172 (6.27)	249* (4.94)	66 (4.87)
I54V	73 (20.33)	115 (20.28)	79 (15.9)	191 (13.37)	269 * (10.94)	273 (9.43)	224 * (7.78)	196 (7.14)	241* (4.78)	58 (4.28)
G190A	46 (12.81)	80 (14.11)	47* (9.46)	146 (10.22)	239 (9.72)	266 (9.19)	271 (9.41)	259 (9.44)	392 * (7.77)	102 (7.53)
P225H	36 (10.03)	56 (9.88)	62 (12.47)	179 (12.53)	320 (13.02)	403 (13.93)	370 (12.85)	390 (14.21)	586 * (11.62)	158 (11.67)
K65R	8 (2.23)	21 (3.7)	23 (4.63)	74 (5.18)	161 (6.55)	194 (6.7)	246 * (8.74)	374 * (13.63)	589 * (11.68)	164 (12.11)

Percentage number under parentheses. * A statistical difference was found with data from the previous year ($p < 0.05$). ** Data for 2017 was only available Jan–Apr.

3.4. A Shift on Treatment Scheme during the Years

Having observed the clear increase in K65R along the studied yrs, we then decided to evaluate the ART combination schemes in use during the studied period (Figure 2), aiming at understanding ART usage impact on the DRM. In 2008, from a total of 359 patients, the most prevalent ART combination was 3TC/AZT/EFV (20.06%, *n* = 72), followed by 3TC/AZT/LPV (8.36%, *n* = 30), prevailing the most common treatment till 2014 (25.03%, *n* = 721) among 2880 patients. From 2015 forward, 3TC/EFV/TDF scheme became the most prevalent. In this year (*n* = 2744), 24.85% (682) individuals were under 3TC/EFV/TDF drug combination and 19.97% (*n* = 548) under 3TC/AZT/EFV. In 2017, 3TC/EFV/TDF combination was used on 40.10% (*n* = 543) of the individuals on ART (*n* = 1354).

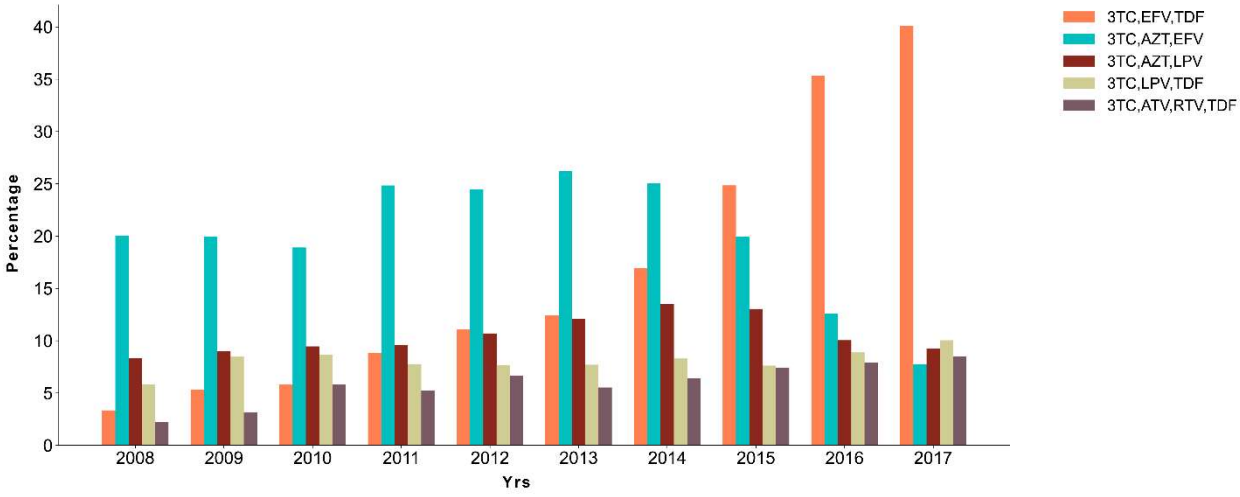


Figure 2. Treatment scheme used at time of viral sequencing between the yrs 2008 and 2017. Percentage of patients (*n* = 20,226) under the five more common ART combinations (3TC/EFV/TDF; 3TC/AZT/EFV; 3TC/AZT/LPV; 3TC/LPV/TDF; 3TC/ATV/RTV/TDF) in the year of the viral sequencing.

As we have found a shift on the ART treatment scheme during the studied years, we decided to evaluate DRM prevalence separating the individuals using the most common drug combinations, 3TC/EFV/TDF (21.38%, *n* = 4324) and 3TC/AZT/EFV (19.79%, *n* = 4003) (Figure 3). The DRM with most similar prevalence when comparing both groups were M46I, V82A, L90M and I54V. In contrast, K65R was only found in 0.37% (*n* = 15) of the viruses isolated from patients treated with 3TC/AZT/EFV

vs. 30.64% (n = 1325) of the viruses from individuals using 3TC/EFV/TDF. This strong association of K65R with TDF containing schemes was found across HIV-1 subtypes (Supplementary Figure S1). Focusing on 2015, the year where the number of patients (n = 2744) using the two schemes was more even (3TC/AZT/EFV: 19.97%, n = 548; vs. 3TC/EFV/TDF: 24.85%, n = 682), 5 patients (0.91%) treated with 3TC/AZT/EFV were infected with a K65R mutant, being this number significantly different (p < 0.001) from the 267 patients (39.15%) treated with 3TC/EFV/TDF combination. Also, comparing all patients revealed a significant association between 3TC/EFV/TDF and K65R (p < 0.001).

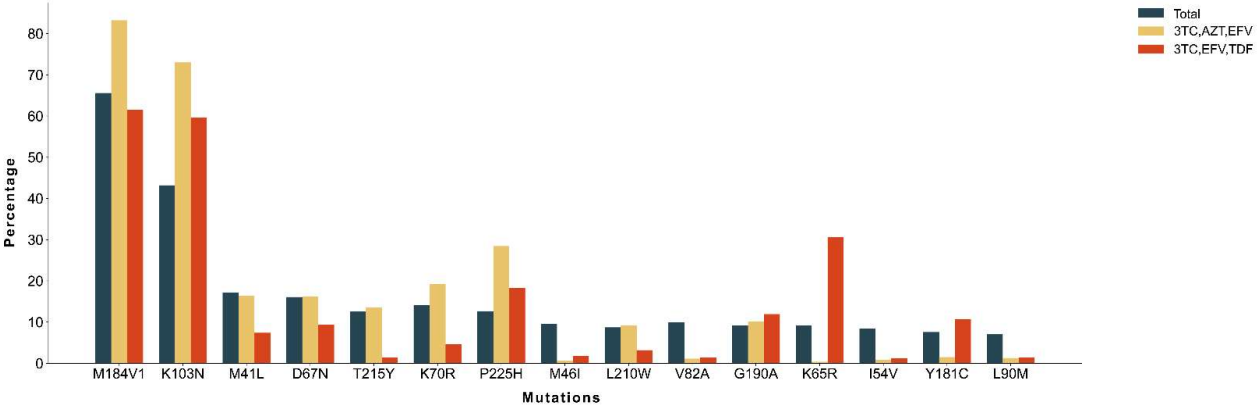


Figure 3. Most found DRM by ART scheme. Percentage of individuals infected with viruses harboring common DRM separated by ART scheme, discriminating 3TC/EFV/TDF (n = 4324, orange), 3TC/AZT/EFV (n = 4003, yellow) and in the total of ART schemes.

3.5. Evidence for Transmission of K65R

To investigate if the transmission of virus harboring K65R could have contributed to the increase in the prevalence of this mutation in Brazil, we performed a phylogenetic analysis on a subset of viral sequences including all the sequences with the K65R mutation and closely related sequences from public databases (Supplementary Figures S2 and S3). The definition of transmission clusters was performed as previously [80] considering tree branch statistical support and mean genetic distance criteria. We found K65R in at least two HIV-1 sequences in 21 well-delimited transmission clusters (16 from subtype B and 5 from subtype C) (Supplementary Table S1). The size of the inferred transmission clusters ranged from

three to 16 individuals. The minimum spanning network analysis of these clusters (Figure 4) supports the occurrence of events of K65R transmission in the study cohort. Furthermore, the analysis of the clinical records from the individuals likely involved in K65R transmission supports this possibility for several cases showing a coincidence of geographic location, proximity in diagnostic dates and similar reported transmission routes (Supplementary Table S1). Despite the significant increased prevalence of the K65R in subtype C (11.26%, n = 347) viruses when compared with B (9.27%, n= 1305) subtype ($p < 0.001$, Supplementary Figure S2) we found probable events of K65R transmission also in subtype B viruses.

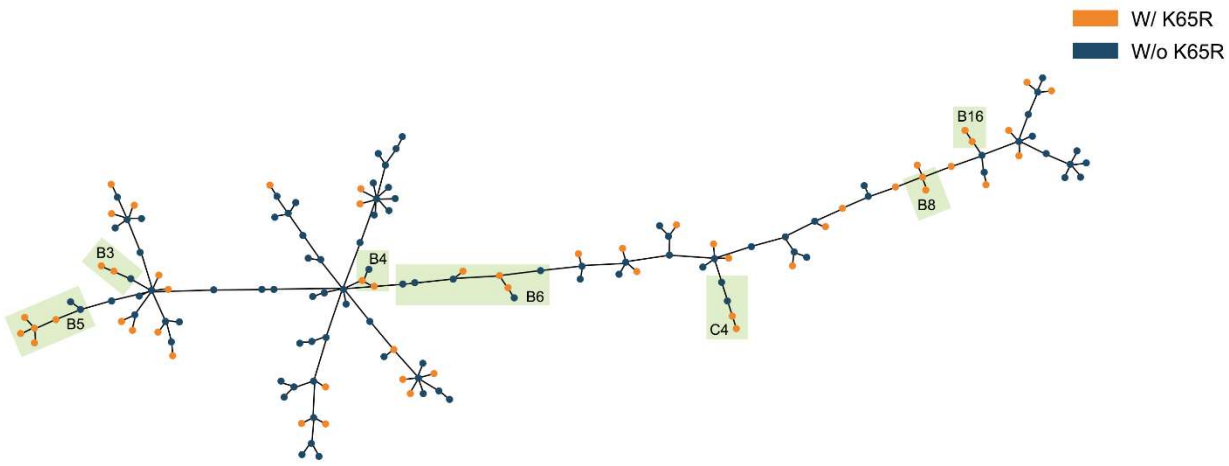


Figure 4. Minimum spanning network analysis of transmission clusters with more than one sequence with K65R mutation. Clusters compatible with K65R transmission are highlighted in green. Sequences without K65R mutation are marked in blue and in orange K65R mutants.

We then used an immunoinformatic approach to investigate the predicted impact of K65R on immune-driven selective pressures. We performed artificial neural network-based predictions of binding to the globally most frequent class I HLAs of all the peptide sequences overlapping K65R and ranging from eight to 12 amino acids with the mutant or the wild-type residue. Interestingly, we found that K65R overlapping peptides were predicted to be recognized by two HLAs (HLA-A03, HLA-B58) in the wildtype version and by one additional HLA (HLA-B27) in the mutant version (Table 3). HLA-B27 has the lowest prevalence (2.23%) in the Brazilian population of all the HLAs tested. Overall, our results suggest that the presence of K65R mutation might contribute for the immune control of the virus in individuals harboring HLA-B27 possibly decreasing its transmission in populations with high prevalence of this HLA.

Table 3. K65R mutant strong binding affinity to most frequent HLAs.

	WT SB Binding	K65R SB Binding	Allele Freq. in BR pop.
HLA-A01	no	no	9.21
HLA-A02	no	no	25.94
HLA-A03	yes	yes	9.26
HLA-A24	no	no	10.00
HLA-A26	no	no	3.35
HLA-B07	no	no	6.92
HLA-B08	no	no	5.12
HLA-B15	no	no	9.08
HLA-B27	no	yes	2.23
HLA-B39	no	no	3.46
HLA-B40	no	no	4.70
HLA-B58	yes	yes	2.65

4. Discussion

Acquired and transmitted DRM remain one of the major obstacles towards ART efficacy and AIDS treatment. In case of Brazil, previous regional studies showed distinct moderate patters of transmitted

DRM prevalence [11,16,20,27,37], and high DRM rates in patients receiving ART [17–19,27,38]. In this study, we aimed at understanding DRM [26] prevalence across the 27 Brazilian federative units focusing on 20,226 HIV-1 infected patients under different ART schemes, with viral sequence genotyped between 2008 and 2017. In accordance to the Brazilian HIV-1 epidemiology report [36], our cohort was composed mainly by male individuals (55.7%) and the majority of the patients' age varied between 30 and 49 years old (61.44%).

A high prevalence of DRM was observed with 84.1% of the infected individuals presenting at least one drug resistance mutation. Nonetheless, 68.5% of the genotyped sequences had a DRM to NRTIs, a value that represents a decrease to what was previously reported in the literature [17–19,38,39], following a trend that has been described by Duani et al. and Diaz et al. in previous years [17,39]. However, NNRTI resistance rates (57.95%) did not follow the increasing trend reported by the same authors, presenting a similar value to what has been found in previous works [18,39]. PI resistance mutations (24.82%) also seem to follow the reported decreasing trend [17,39].

Also in line with previous studies is the finding of M184V (65.53%), K103N (40.20%), and M41L (17.21%) as the most common DRM [17,19,38,39]. M184V, the most prevalent, is a NRTI mutation selected by 3TC and associated with impaired viral fitness and hypersensitization to other NRTI, including AZT and TDF [40–42]. Along with the most common mutations, M184V rates presented a decrease over the yrs (76.88% in 2008 to 51.48% in 2017). This decrease on acquired DRM was also observed in other countries and was hypothesized to be due to an improvement on treatment efficacy and also an increased accessibility to ART which enhances patient adherence to therapy [43–45]. Moreover, in Brazil, pretreatment genotyping has been extended to children living with HIV under 12 years of age, in case of *Mtb* co-infection, to pregnant woman and new diagnoses with a sexual partner on ART, since 2013 [46], which may had an impact the on the DRM general decrease trend.

Nonetheless, it is important to highlight that the percentage of PLWH infected with a virus harboring M148V remained very high and above 50% after 2013. Moreover, K103N and P225H, both frequent NNRTI DRM, remained stable over the yrs. K103N was previously reported as one of the most commonly acquired DRM in Brazil [17,18,38,39] in association with the use of EFV [47]. P225H is also selected by EFV and generally occurs in the presence of K103N, synergically increasing EFV resistance [47–49]. As for K65R, it is selected by TDF and by other NRTI including abacavir (ABC) and 3TC, decreasing viral susceptibility to most of these drugs [29,31,50].

Previous work reported that K65R mutation presented varied prevalence distribution throughout time in different Brazilian regions [17,19,27]. Strikingly, in our nationwide work in Brazil we found that K65R was the only DRM that followed an accentuated increase in prevalence over the studied years (2.23% in 2008 to 12.11% in 2017). The increasing and preferential usage of TDF in the clinical practice [21], including in a context of a failing regimen, could be the primordial reason for the significant expansion of K65R as other studies show a higher prevalence of this mutation in patients failing ART treatment [15,51–54]. However, although an increasing trend was identified in a cohort from India [55], several studies report a decrease on K65R levels over the years in other countries [47,56–59]. Reinheimer and colleagues suggest that the decline on K65R prevalence rates is linked with the increasing usage of single tablet ART regimens and the inclusion of AZT on the used treatment regimen, which has been liked with K65R development suppression [60]. Moreover, Theys et al. work [58] suggests that the found decline of tenofovir K65R selection rate in Portugal, between 2002 and 2010, is mainly caused by changes on treatment guidelines over the years and by the increased usage of combination of TDF and emtricitabine (FTC) [58]. Collectively, these studies suggest that the genetic or sociodemographic characteristics of the population treated with TDF might be influencing the K65R levels.

In 2019, 15.074 individuals were reported to use PrEP in Brazil [1]. Although PrEP consists of a co-formulation of FTC and TDF [61], evidence suggests that the risk of selection for TDF resistance is low and is more likely to occur in cases of undiagnosed HIV infection [62,63]. However, in these cases, high prevalence rates of M184V, which is also selected by FTC, and K65R as found in our study population, might compromise the efficacy of PrEP [64,65].

K65R has been associated with diminished viral fitness and replication, which has been linked with decreased transmission capacity [66,67]. Even though several studies support almost inexistent K65R transmission rates [11,68,69], Rhee and co-works present evidence of higher levels of transmitted K65R, especially in low- and middle-income countries [70]. In our phylogenetic analysis we found 21 transmission clusters containing at least two sequences with the K65R mutation. Moreover, minimum spanning network analysis supported the occurrence of events of K65R transmission in at least seven of these clusters. This might be suggestive that the efficacy of PrEP might be compromised in this setting. Indeed, we found much higher viral loads in patients harboring K65R mutation when compared with the rest of our cohort population. Similar viral loads were observed between individuals with and without tenofovir resistance by The TenoRes Study Group [71], which can be associated with increased mutation transmission. Furthermore, our results suggest that K65R overlapping peptides can be increasingly

recognized by HLA-B27, a HLA that is linked with lower viral loads and slower diseased progression to AIDS [72,73]. The fact that Brazil presents a low prevalence of the HLA-B27, when in comparison with other regions [74–76], might have contributed to K65R expansion in this country. However, it is relevant to point-out that our results are based on in-silico prediction and not taking in consideration the possible impact on immune responses of the combination of different mutations in the same virus. Thus, it is relevant to address this topic in the future by performing functional studies.

Despite of the observed decreasing trend in the prevalence of some major DRM in the studied Brazilian cohort, a high prevalence of these mutations was still verified. Indeed, our results show that the alteration performed in 2013 enlarging the criteria to include more PLWH in the recommendation for baseline DRM testing had a positive impact but was still insufficient. We propose that the lack of universal baseline HIV-1 DRM screening to inform on effective ART regimens resulted in high levels of DRM, such as M184V, K103N, and M41L underlying many cases of treatment failure in Brazil not only from 2008–2012 but also continuing from 2013–2017. Furthermore, we observed a clear increase in K65R reverse transcriptase DRM that is an additional problem in Brazil that could have been aggravated by the circulation of HIV-1 subtype C and the HLA class I makeup of the population. This increase in K65R was mainly found in association with the use of TDF and is particularly relevant in combination with the high M184V levels found in the study population suggesting that the efficacy of PrEP might be compromised.

Overall, our results support that some of the drugs most frequently used in Brazil might be compromised due to the high frequency of DRM and that baseline drug resistance testing should be universal and mandatory as it is the best way to promote personalized selection of the most optimized ART regiment.

5. Materials and Methods

5.1. Study Population

The collection of the patient's data was performed anonymously after approval by the Brazilian national ethic committee through the protocol CAAE 53757016.0.0000.5504. All records from HIV-1-infected individuals followed by the Specialized Assistance Services on Sexually Transmissible Diseases and HIV/AIDS obtained at all 27 Brazilian federative units from 01/01/2008 to 04/30/2017 were

curated in a relational database. A total of 20,226 HIV-1 infected patients were selected for this study according to the following inclusion criteria: (i) be under ART treatment; (ii) have at least one associated partial HIV-1 genome sequence available (complete HIV protease (n = 20,223) and complete reverse transcriptase (n = 20,214)). HIV-1 sequencing was performed as part of the routine clinical testing with commercially available HIV-1 genotyping systems based on Sanger sequencing.

5.2. HIV-1 Subtypes and Drug Resistance Mutations

HIV-1 subtypes were identified by relying on the consensus from least two of the three utilized subtyping tools: SNAPPY [77], jpHMM [78] and Stanford HIV Drug Resistance database (https://hivdb.stanford.edu/, accessed 01/05/2021). All sequences were also analyzed using the Stanford HIVdb Genotypic Resistance Interpretation Algorithm to evaluate and interpret the presence of DRM in each sequence.

5.3. Phylogenetic Analysis of K65R Sequences and Transmission Clusters

All sequences having the K65R were separated by subtype and used to query local and public databases to identify highly related HIV-1 sequences. The resulting sequences from the most common subtypes B and C were aligned with MAFFT v7.309 [79] and used to make a phylogenetic reconstruction using PhyML v3.0 [80]. The best fitting substitution model was GTR + G4 + I, determined by PhyML SMS using AIC [81]. The heuristic trees search was performed using SPR methods. Bayesian evolutionary analyses were performed using BEAST v1.10.4 [82,83] with GTR + G4 + I, as the nucleotide substitution model. This phylogenetic representation was used to infer the transmission clusters as previously [84]. Minimum spanning network analysis of the sequences identified in transmission clusters with more than one K65R mutant was performed with PHYLOViZ [85].

5.4. HLA Binding Affinity Predictions

Predictions of binding for different class 1 human leukocyte antigens (HLA) were performed with NetMHCpan 4.1 [86]. The wildtype sequence (NPYNTPVFAIKKKDSTKWRKLVD) and the sequence with the K65R mutation (NPYNTPVFAIKRKDSTKWRKLVD) were used to generate all possible peptides sequences overlapping position 65 and ranging from eight to 12 amino acids. A group of 11 HLAs were used as supertype representatives: HLA-A01:01; HLA-A02:01; HLA-A03:01; HLA-A24:02; HLA-A26:01; HLA-B07:02; HLA-B08:01; HLA-B27:05; HLA-B39:01; HLA-B40:01; HLA-B58:01; HLA-B15:01. The thresholds for the definition of binding were maintained as the tool defaults, only binding results classified as "strong binding" were considered.

5.5. HLA Prevalence in the Brazilian Population

The Allele Frequencies Database [87] was used as the source of this information for the Brazilian population. To obtain cohesive and yet representative nationwide results, the bone marrow registry (REDOME) data was selected. Since this data is separated by Brazilian state, we calculated the proportional values for a national level estimation.

5.6. Statistical Analysis

Statistical analysis was performed by using Epi Info™ version 7.2.4.0 (https://www.cdc.gov/epiinfo/index.html, accessed 01/05/2021), relying on Mantel-Haenszel test, at significance level of 0.05, to evaluate statistical correlations. SPSS® Statistics version 26.0 (IBM®, Armonk, NY, USA) was also utilized for performing mean comparison with t-test, at a 0.05 significance level.

6. References

1. Joint United Nations Programme on HIV/AIDS *UNAIDS Data 2020*; UNAIDS: Geneva, Switzerland ,2020.
2. Souza Junior, P.R.B.D.; Szwarcwald, C.L.; Castilho, E.A.D. Self-rated health by HIV-infected individuals undergoing antiretroviral therapy in Brazil. *Cad. Saúde Pública/Ministério Saúde,*

Fundação Oswaldo Cruz. Esc. Nac. Saúde Pública **2011**, doi:10.1590/s0102-311x2011001300007.

3. Teeraananchai, S.; Kerr, S.J.; Amin, J.; Ruxrungtham, K.; Law, M.G. Life expectancy of HIV-positive people after starting combination antiretroviral therapy: A meta-analysis. *HIV Med.* **2017**, *18*, 256–266, doi:10.1111/hiv.12421.

4. Trickey, A.; May, M.T.; Vehreschild, J.J.; Obel, N.; Gill, M.J.; Crane, H.M.; Boesecke, C.; Patterson, S.; Grabar, S.; Cazanave, C.; et al. Survival of HIV-positive patients starting antiretroviral therapy between 1996 and 2013: A collaborative analysis of cohort studies. *Lancet HIV* **2017**, *4*, e349–e356, doi:10.1016/S2352-3018(17)30066-8.

5. Saag, M.S.; Gandhi, R.T.; Hoy, J.F.; Landovitz, R.J.; Thompson, M.A.; Sax, P.E.; Smith, D.M.; Benson, C.A.; Buchbinder, S.P.; Del Rio, C.; et al. Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults: 2020 Recommendations of the International Antiviral Society-USA Panel. *JAMA* **2020**, *324*, 1651–1669, doi: 10.1001/jama.2020.17025.

6. Lundgren, J.; Phillips, A. Prevention of HIV transmission by antiretroviral therapy. *Lancet HIV* **2018**, *5*, e108–e109, doi: 10.1016/S2352-3018(17)30204-7

7. Eisinger, R.W.; Dieffenbach, C.W.; Fauci, A.S. HIV viral load and transmissibility of HIV infection undetectable equals untransmittable. *JAMA* **2019**, *321*, 451–452, doi: 10.1001/jama.2018.21167.

8. Cohen, M.S.; Chen, Y.Q.; McCauley, M.; Gamble, T.; Hosseinipour, M.C.; Kumarasamy, N.; Hakim, J.G.; Kumwenda, J.; Grinsztejn, B.; Pilotto, J.H.S.; et al. Antiretroviral Therapy for the Prevention of HIV-1 Transmission. *N. Engl. J. Med.* **2016**, *375*, 830–839, doi:10.1056/nejmoa1600693.

9. Tanaka, T.S.O.; Leite, T.F.; Freitas, S.Z.; Cesar, G.A.; De Rezende, G.R.; Lindenberg, A.D.S.C.; Guimarães, M.L.; Motta-Castro, A.R.C. HIV-1 molecular epidemiology, transmission clusters and transmitted drug resistance mutations in central Brazil. *Front. Microbiol.* **2019**, *10*, doi:10.3389/fmicb.2019.00020.

10. Booth, C.L.; Geretti, A.M. Prevalence and determinants of transmitted antiretroviral drug resistance in HIV-1 infection. *J. Antimicrob. Chemother.* **2007**, *59*, 1047–1056, doi: 10.1093/jac/dkm082.

11. Arruda, M.B.; Boullosa, L.T.; Cardoso, C.C.; da Costa, C.M.; Brites, C.; de Lima, S.T.S.; Kaminski, H.T.; Aleixo, A.W.; Esposito, A.O.P.; Cavalcanti, A.M.S.; et al. Brazilian network for HIV drug resistance surveillance (HIV-BresNet): A survey of treatment-naive individuals. *J. Int. AIDS Soc.* **2018**, *21*, doi:10.1002/jia2.25032.

12. Geretti, A.M. Epidemiology of antiretroviral drug resistance in drug-naïve persons. *Curr. Opin. Infect. Dis.* **2007**, *20*, 22–32, doi:10.1097/QCO.0b013e328013caff.

13. Frentz, D.; Boucher, C.A.B.; Van De Vijver, D.A.M.C. Temporal changes in the epidemiology of transmission of drug-resistant HIV-1 across the world. *AIDS Rev.* **2012**, *14*, 17–27.

14. Sui, H.; Gui, T.; Jia, L.; Guo, W.; Han, J.; Liu, Y.; Bao, Z.; Li, H.; Li, J.; Li, L. Different Frequencies of Drug Resistance Mutations among HIV-1 Subtypes Circulating in China: A Comprehensive Study. *PLoS ONE* **2014**, *9*, e91803, doi:10.1371/journal.pone.0091803.

15. Chaplin, B.; Eisen, G.; Idoko, J.; Onwujekwe, D.; Idigbe, E.; Adewole, I.; Gashau, W.; Meloni, S.; Sarr, A.D.; Sankalé, J.L.; et al. Impact of HIV type 1 subtype on drug resistance mutations in Nigerian patients failing first-line therapy. *AIDS Res. Hum. Retrovir.* **2011**, *27*, 71–80, doi:10.1089/aid.2010.0050.

16. Bahls, L.D.; Canezin, P.H.; Reiche, E.M.V.; Fernandez, J.C.C.; Dias, J.R.C.; Meneguetti, V.A.F.; Ueda, L.T.; Bertolini, D.A. Moderate prevalence of HIV-1 transmitted drug resistance mutations in southern Brazil. *AIDS Res. Ther.* **2019**, doi:10.1186/s12981-019-0219-1.

17. Duani, H.; Aleixo, A.W.; Tupinambás, U. Trends and predictors of HIV-1 acquired drug resistance in Minas Gerais, Brazil: 2002–2012. *Braz. J. Infect. Dis.* **2017**, *21*, 148–154, doi:10.1016/j.bjid.2016.11.009.

18. Cavalcanti, A.M.S.; Lacerda, H.R.; De Brito, A.M.; Pereira, S.; Medeiros, D.; Oliveira, S. Antiretroviral resistance in individuals presenting therapeutic failure and subtypes of the human immunodeficiency virus type 1 in the Northeast Region of Brazil. *Mem. Inst. Oswaldo Cruz* **2007**, *102*, 785–792, doi:10.1590/S0074-02762007005000109.

19. Toledo, P.V.M.; de Carvalho, D.S.; Romagnoli, L.; Marcinko, G.; da Cunha, C.A.; de Souza, M.N.; Brindeiro, R.; de Queiroz-Telles, F. HIV-1 genotypic resistance profile of patients failing antiretroviral therapy in Paraná, Brazil. *Braz. J. Infect. Dis.* **2010**, *14*, 360–371, doi:10.1016/s1413-8670(10)70076-3.

20. Coelho, L.P.O.; Matsuda, E.M.; Nogueira, R.S.; de Moraes, M.J.; Jamal, L.F.; Madruga, J.V.R.; Tancredi, M.V.; de Leão, A.C.Q.; de Faria Romero Soldi, G.; de Macedo Brígido, L.F.; et al. Prevalence of HIV-1 transmitted drug resistance and viral suppression among recently diagnosed adults in São Paulo, Brazil. *Arch. Virol.* **2019**, *164*, 699–706, doi:10.1007/s00705-018-04122-8.

21. Ministério da Saúde *Clinical Protocol and Therapeutic Guidelines for Handling HIV Infection in Adults*, 1st ed.; Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância, Prevenção e Controle das Infecções Sexualmente Transmissíveis, do HIV/Aids e das Hepatites Virais, Brasilia/DF, Brazil, 2018; ISBN 9788533426436.

22. Benzaken, A.S.; Pereira, G.F.M.; Costa, L.; Tanuri, A.; Santos, A.F.; Soares, M.A. Antiretroviral treatment, government policy and economy of HIV/AIDS in Brazil: Is it time for HIV cure in the country? *AIDS Res. Ther.* **2019**, *16*, doi:10.1186/s12981-019-0234-2.

23. European AIDS Clinical Society, E. European AIDS clinical society guidelines. Version 10.0. November 2019. *IEEE Trans. Sonics Ultrason.* **2019,** 6-11.

24. Panel on Antiretroviral Guidelines for Adults and Adolescents Guidelines for the Use of Antiretroviral Agents in Adults and Adolescents with HIV. Department of Health and Human Services. *Dep. Heal. Hum. Serv.* **2019,** C-12-21.

25. Günthard, H.F.; Calvez, V.; Paredes, R.; Pillay, D.; Shafer, R.W.; Wensing, A.M.; Jacobsen, D.M.; Richman, D.D. Human Immunodeficiency Virus Drug Resistance: 2018 Recommendations of the International Antiviral Society-USA Panel. *Clin. Infect. Dis.* **2019**, *68*, 177–187, doi:10.1093/cid/ciy463.

26. Bennett, D.E.; Camacho, R.J.; Otelea, D.; Kuritzkes, D.R.; Fleury, H.; Kiuchi, M.; Heneine, W.; Kantor, R.; Jordan, M.R.; Schapiro, J.M.; et al. Drug Resistance Mutations for Surveillance of Transmitted HIV-1 Drug-Resistance: 2009 Update. *PLoS ONE* **2009**, *4*, e4724, doi:10.1371/journal.pone.0004724.

27. Brites, C.; Pinto-Neto, L.; Medeiros, M.; Nunes, E.; Sprinz, E.; Carvalho, M. Extensive variation in drug-resistance mutational profile of Brazilian patients failing antiretroviral therapy in five large Brazilian cities. *Braz. J. Infect. Dis.* **2016**, *20*, 323–329, doi:10.1016/j.bjid.2016.03.010.

28. Touloumi, G.; Pantazis, N.; Pillay, D.; Paraskevis, D.; Chaix, M.-L.; Bucher, H.C.; Kucherer, C.; Zangerle, R.; Kran, A.-M.B.; Porter, K.; et al. Impact of HIV-1 Subtype on CD4 Count at HIV Seroconversion, Rate of Decline, and Viral Load Set Point in European Seroconverter Cohorts. *Clin. Infect. Dis.* **2013**, *56*, 888–897, doi:10.1093/cid/cis1000.

29. Wensing, A.M.; Calvez, V.; Ceccherini-Silberstein, F.; Charpentier, C.; Günthard, H.F.; Paredes, R.; Shafer, R.W.; Richman, D.D. 2019 update of the drug resistance mutations in HIV-1. *Top. Antivir. Med.* **2019**, *27*, 111–121.

30. White, K.L.; Chen, J.M.; Feng, J.Y.; Margot, N.A.; Ly, J.K.; Ray, A.S.; Macarthur, H.L.; McDermott, M.J.; Swaminathan, S.; Miller, M.D. The K65R reverse transcriptase mutation in HIV-1 reverses the excision phenotype of zidovudine resistance mutations. *Antivir. Ther.* **2006**, *11*, 155–163.

31. Brenner, B.G.; Coutsinos, D. The K65R mutation in HIV-1 reverse transcriptase: Genetic barriers, resistance profile and clinical implications. *HIV Ther.* **2009**, *3*, 583–594, doi:10.2217/hiv.09.40.

32. Smit, E.; White, E.; Clark, D.; Churchill, D.; Zhang, H.; Collins, S.; Pillay, D.; Sabin, C.; Nelson, M.; Winston, A.; et al. An association between K65R and HIV-1 subtype C viruses in patients treated with multiple NRTIs. *J. Antimicrob. Chemother.* **2017**, *72*, 2075–2082, doi:10.1093/jac/dkx091.

33. Inzaule, S.C.; Weidle, P.J.; Yang, C.; Ndiege, K.; Hamers, R.L.; Rinke de Wit, T.F.; Thomas, T.; Zeh, C. Prevalence and dynamics of the K65R drug resistance mutation in HIV-1-infected infants exposed to maternal therapy with lamivudine, zidovudine and either nevirapine or nelfinavir in breast milk. *J. Antimicrob. Chemother.* **2016**, *71*, 1619–1626, doi:10.1093/jac/dkw039.

34. Theys, K.; Vercauteren, J.; Snoeck, J.; Zazzi, M.; Camacho, R.J.; Torti, C.; Schülter, E.; Clotet, B.; Sönnerborg, A.; De Luca, A.; et al. HIV-1 subtype is an independent predictor of reverse transcriptase mutation k65r in HIV-1 patients treated with combination antiretroviral therapy including tenofovir. *Antimicrob. Agents Chemother.* **2013**, *57*, 1053–1056, doi:10.1128/AAC.01668-12.

35. Coutsinos, D.; Invernizzi, C.F.; Moisi, D.; Oliveira, M.; Martinez-Cajas, J.L.; Brenner, B.G.; Wainberg, M.A. A Template-Dependent Dislocation Mechanism Potentiates K65R Reverse Transcriptase Mutation Development in Subtype C Variants of HIV-1. *PLoS ONE* **2011**, *6*, e20208, doi:10.1371/journal.pone.0020208.

36. Ministério da Saúde. *Boletim Epidemiológico de HIV e Aids*; Ministério da Saúde Secretaria de Vigilância em Saúde Departamento de Vigilância, Prevenção e Controle das Infecções Sexualmente Transmissíveis, do HIV/Aids e das Hepatites Virais, Brasilia/DF, Brazil, 2018; ISBN 1517-1159.

37. Sprinz, E.; Netto, E.M.; Lima, M.P.J.S.; Furtado, J.J.D.; Da Eira, M.; Zajdenverg, R.; Madruga, J.V.; Lewi, D.S.; Pedro, R.J.; Soares, M.A. Primary antiretroviral drug resistance among HIV Type 1-infected individuals in Brazil. *AIDS Res. Hum. Retrovir.* **2009**, *25*, 861–867, doi:10.1089/aid.2009.0012.

38. Lopes, C.A.F.; Soares, M.A.; Falci, D.R.; Sprinz, E. The Evolving Genotypic Profile of HIV-1 Mutations Related to Antiretroviral Treatment in the North Region of Brazil. *Biomed Res. Int.* **2015**, *2015*, doi:10.1155/2015/738528.

39. Diaz, R.S.; Inocêncio, L.A.; Sucupira, M.C.A.; Pereira, A.A.; Hunter, J.; Ferreira, J.E.; Araújo, L.V.; Souza, D.F.C.; Sabino, E.C. The Virological and Immunological Characteristics of the HIV-1-Infected Population in Brazil: From Initial Diagnosis to Impact of Antiretroviral Use. *PLoS ONE* **2015**, *10*, e0139677, doi:10.1371/journal.pone.0139677.

40. Wolf, K.; Walter, H.; Beerenwinkel, N.; Keulen, W.; Kaiser, R.; Hoffmann, D.; Lengauer, T.; Selbig, J.; Vandamme, A.M.; Korn, K.; et al. Tenofovir Resistance and Resensitization. *Antimicrob. Agents Chemother.* **2003**, *47*, 3478–3484, doi:10.1128/AAC.47.11.3478-3484.2003.

41. Götte, M.; Arion, D.; Parniak, M.A.; Wainberg, M.A. The M184V Mutation in the Reverse Transcriptase of Human Immunodeficiency Virus Type 1 Impairs Rescue of Chain-Terminated DNA Synthesis. *J. Virol.* **2000**, *74*, 3579–3585, doi:10.1128/jvi.74.8.3579-3585.2000.

42. Turner, D.; Brenner, B.G.; Routy, J.P.; Petrella, M.; Wainberg, M.A. Rationale for maintenance of the M184v resistance mutation in human immunodeficiency virus type 1 reverse transcriptase in treatment experienced patients. *New Microbiol.* **2004**, *27*, 31–39.

43. Pingarilho, M.; Pimentel, V.; Diogo, I.; Fernandes, S.; Miranda, M.; Pineda-Pena, A.; Libin, P.; Theys, K.; Martins, M.R.O.; Vandamme, A.M.; et al. Increasing Prevalence of HIV-1 Transmitted Drug Resistance in Portugal: Implications for First Line Treatment Recommendations. *Viruses* **2020**, *12*, 1238, doi:10.3390/v12111238.

44. Franzetti, M.; Violin, M.; Antinori, A.; De Luca, A.; Ceccherini-Silberstein, F.; Gianotti, N.; Torti, C.; Bonora, S.; Zazzi, M.; Balotta, C. Trends and correlates of HIV-1 resistance among subjects failing an antiretroviral treatment over the 2003-2012 decade in Italy. *BMC Infect. Dis.* **2014**, *14*, 398, doi:10.1186/1471-2334-14-398.

45. Rocheleau, G.; Brumme, C.J.; Shoveller, J.; Lima, V.D.; Harrigan, P.R. Longitudinal trends of HIV drug resistance in a large Canadian cohort, 1996–2016. *Clin. Microbiol. Infect.* **2018**, *24*, 185–191, doi:10.1016/j.cmi.2017.06.014.

46. Ministério da Saúde PROTOCOLO CLÍNICO E DIRETRIZES TERAPÊTUICAS PARA O MANEJO DA INFECÇÃO PELO HIV. Ministério da Saúde Secretaria de Vigilância em Saúde Departamento de Vigilância, Prevenção e Controle das Infecções Sexualmente Transmissíveis, do HIV/Aids e das Hepatites Virais, Brasilia/DF, Brazil, . **2013**.

47. Bacheler, L.T.; Anton, E.D.; Kudish, P.; Baker, D.; Bunville, J.; Krakowski, K.; Bolling, L.; Aujay, M.; Wang, X.V.; Ellis, D.; et al. Human immunodeficiency virus type 1 mutations selected in patients failing efavirenz combination therapy. *Antimicrob. Agents Chemother.* **2000**, *44*, 2475–2484, doi:10.1128/AAC.44.9.2475-2484.2000.

48. Alcaro, S.; Alteri, C.; Artese, A.; Ceccherini-Silberstein, F.; Costa, G.; Ortuso, F.; Bertoli, A.; Forbici, F.; Santoro, M.M.; Parrotta, L.; et al. Docking Analysis and Resistance Evaluation of Clinically Relevant Mutations Associated with the HIV-1 Non-nucleoside Reverse Transcriptase Inhibitors Nevirapine, Efavirenz and Etravirine. *Chem. Med. Chem.* **2011**, *6*, 2203–2213, doi:10.1002/cmdc.201100362.

49. Bacheler, L.; Jeffrey, S.; Hanna, G.; D'Aquila, R.; Wallace, L.; Logue, K.; Cordova, B.; Hertogs, K.; Larder, B.; Buckery, R.; et al. Genotypic Correlates of Phenotypic Resistance to Efavirenz in Virus Isolates from Patients Failing Nonnucleoside Reverse Transcriptase Inhibitor Therapy. *J. Virol.* **2001**, *75*, 4999–5008, doi:10.1128/jvi.75.11.4999-5008.2001.

50. Garforth, S.J.; Lwatula, C.; Prasad, V.R. The lysine 65 residue in HIV-1 reverse transcriptase function and in nucleoside analog drug resistance. *Viruses* **2014**, *6*, 4080–4094, doi:10.3390/v6104080.

51. Ayitewala, A.; Kyeyune, F.; Ainembabazi, P.; Nabulime, E.; Kato, C.D.; Nankya, I. Comparison of HIV drug resistance profiles across HIV-1 subtypes A and D for patients receiving a tenofovir-

based and zidovudine-based first line regimens in Uganda. *AIDS Res. Ther.* **2020**, *17*, doi:10.1186/s12981-020-0258-7.

52. Luo, X.L.; Mo, L.D.; Su, G.S.; Huang, J.P.; Wu, J.Y.; Su, H.Z.; Huang, W.H.; Luo, S.D.; Ni, Z.Y. Incidence and types of HIV-1 drug resistance mutation among patients failing first-line antiretroviral therapy. *J. Pharmacol. Sci.* **2019**, *139*, 275–279, doi:10.1016/j.jphs.2018.11.016.

53. Parker Hudson, F.; Mulenga, L.; Westfall, A.O.; Warrier, R.; Mweemba, A.; Saag, M.S.; Stringer, J.S.A.; Eron, J.J.; Chi, B.H. Evolution of HIV-1 drug resistance after virological failure of first-line antiretroviral therapy in Lusaka, Zambia. *Antivir. Ther.* **2019**, *24*, 291–300, doi:10.3851/IMP3299.

54. Van Zyl, G.U.; Liu, T.F.; Claassen, M.; Engelbrecht, S.; de Oliveira, T.; Preiser, W.; Wood, N.T.; Travers, S.; Shafer, R.W. Trends in Genotypic HIV-1 Antiretroviral Resistance between 2006 and 2012 in South African Patients Receiving First-and Second-Line Antiretroviral Treatment Regimens. *PLoS ONE* **2013**, *8*, doi:10.1371/journal.pone.0067188.

55. Karade, S.; Chaturbhuj, D.N.; Sen, S.; Joshi, R.K.; Kulkarni, S.S.; Shankar, S.; Gangakhedkar, R.R. HIV drug resistance following a decade of the free antiretroviral therapy programme in India: A review. *Int. J. Infect. Dis.* **2018**, *66*, 33–41, doi:10.1016/j.ijid.2017.10.020.

56. Tittle, V.; Boffito, M.; McOwan, A.; Whitlock, G. Antiretroviral resistance and management after pre-exposure to prophylaxis. *Lancet HIV* **2020**, *7*, e84, doi:10.1016/S2352-3018(19)30404-7.

57. Charpentier, C.; Lambert-Niclot, S.; Visseaux, B.; Morand-Joubert, L.; Storto, A.; Larrouy, L.; Landman, R.; Calvez, V.; Marcelin, A.-G.; Descamps, D. Evolution of the K65R, K103N and M184V/I reverse transcriptase mutations in HIV-1-infected patients experiencing virological failure between 2005 and 2010. *J. Antimicrob. Chemother.* **2013**, *68*, 2197–2198, doi:10.1093/jac/dkt184.

58. Theys, K.; Snoeck, J.; Vercauteren, J.; Abecasis, A.B.; Vandamme, A.M.; Camacho, R.J.; Mansinho, K.; Miranda, A.C.; Aldir, I.; Ventura, F.; et al. Decreasing population selection rates of resistance mutation K65R over time in HIV-1 patients receiving combination therapy including tenofovir. *J. Antimicrob. Chemother.* **2013**, *68*, 419–423, doi:10.1093/jac/dks380.

59. Dolling, D.; Sabin, C.; Delpech, V.; Smit, E.; Pozniak, A.; Asboe, D.; Brown, A.L.; Churchill, D.; Williams, I.; Geretti, A.M.; et al. Time trends in drug resistant HIV-1 infections in the United Kingdom up to 2009: Multicentre observational study. *BMJ* **2012**, *345*, doi:10.1136/bmj.e5253.

60. Reinheimer, C.; Wesner, A.; Keppler, O.T.; Doerr, H.W.; Herrmann, E.; Stürmer, M.; Stephan, C. Prevalence of K65R in patients treated with tenofovir disoproxil fumarate: Recommendations based on the Frankfurt HIV Cohort Study Resistance Database (FHCS-RD). *Med. Microbiol. Immunol.* **2016**, *205*, 315–320, doi:10.1007/s00430-015-0448-4.

61. Ministério da Saúde PROTOCOLO CLÍNICO E DIRETRIZES TERAPÊUTICAS PARA PROFILAXIA PRÉ-EXPOSIÇÃO (PrEP) DE RISCO À INFECÇÃO PELO HIV. 1st ed.; Ministério da Saúde Secretaria de Vigilância em Saúde Departamento de Vigilância, Prevenção e Controle das Infecções Sexualmente Transmissíveis, do HIV/Aids e das Hepatites Virais, Brasilia/DF, Brazil,**2018,** 29**,** ISBN 978-85-334-2582-8.

62. Gibas, K.M.; van den Berg, P.; Powell, V.E.; Krakower, D.S. Drug Resistance During HIV Pre-Exposure Prophylaxis. *Drugs* **2019**, *79*, 609–619, doi:10.1007/s40265-019-01108-x.

63. Parikh, U.M.; Mellors, J.W. Should we fear resistance from tenofovir/emtricitabine preexposure prophylaxis? *Curr. Opin. HIV AIDS* **2016**, *11*, 49–55, doi:10.1097/COH.0000000000000209.

64. Margot, N.A.; Waters, J.M.; Miller, M.D. In vitro human immunodeficiency virus type 1 resistance selections with combinations of tenofovir and emtricitabine or abacavir and lamivudine. *Antimicrob. Agents Chemother.* **2006**, *50*, 4087–4095, doi:10.1128/AAC.00816-06.

65. Diallo, K.; Götte, M.; Wainberg, M.A. Molecular Impact of the M184V Mutation in Human Immunodeficiency Virus Type 1 Reverse Transcriptase. *Antimicrob. Agents Chemother.* **2003**, *47*, 3377–3383, doi:10.1128/AAC.47.11.3377-3383.2003.

66. Wagner, B.G.; Garcia-Lerma, J.G.; Blower, S. Factors limiting the transmission of HIV mutations conferring drug resistance: Fitness costs and genetic bottlenecks. *Sci. Rep.* **2012**, *2*, 1–10, doi:10.1038/srep00320.

67. Weber, J.; Chakraborty, B.; Weberova, J.; Miller, M.D.; Quiñones-Mateu, M.E. Diminished replicative fitness of primary human immunodeficiency virus type 1 isolates harboring the K65R mutation. *J. Clin. Microbiol.* **2005**, *43*, 1395–1400, doi:10.1128/JCM.43.3.1395-1400.2005.

68. Chan, P.A.; Huang, A.; Kantor, R. Low prevalence of transmitted K65R and other tenofovir resistance mutations across different HIV-1 subtypes: Implications for pre-exposure prophylaxis. *J. Int. AIDS Soc.* **2012**, *15*, doi:10.7448/IAS.15.2.17701.

69. Margot, N.A.; Wong, P.; Kulkarni, R.; White, K.; Porter, D.; Abram, M.E.; Callebaut, C.; Miller, M.D. Commonly transmitted HIV-1 drug resistance mutations in reverse-transcriptase and protease in antiretroviral treatment-naive patients and response to regimens containing tenofovir disoproxil fumarate or tenofovir alafenamide. *J. Infect. Dis.* **2017**, *215*, 920–927, doi:10.1093/infdis/jix015.

70. Rhee, S.Y.; Jordan, M.R.; Raizes, E.; Chua, A.; Parkin, N.; Kantor, R.; Van Zy, G.U.; Mukui, I.; Hosseinipour, M.C.; Frenkel, L.M.; et al. HIV-1 drug resistance mutations: Potential applications for point-of-care Genotypic resistance testing. *PLoS ONE* **2015**, *10*, 12, doi:10.1371/journal.pone.0145772.

71. Gregson, J.; Tang, M.; Ndembi, N.; Hamers, R.L.; Marconi, V.C.; Brooks, K.; Theys, K.; Arruda, M.; Garcia, F.; Monge, S.; et al. Global epidemiology of drug resistance after failure of WHO recommended first-line regimens for adult HIV-1 infection: A multicentre retrospective cohort study. *Lancet Infect. Dis.* **2016**, *16*, 565–575, doi:10.1016/S1473-3099(15)00536-8.

72. Kløverpris, H.N.; Leslie, A.; Goulder, P. Role of HLA adaptation in HIV evolution. *Front. Immunol.* **2016**, *6*, 1, doi:10.3389/fimmu.2015.00665.

73. Neumann-Haefelin, C. HLA-B27-mediated protection in HIV and hepatitis C virus infection and pathogenesis in spondyloarthritis: Two sides of the same coin? *Curr. Opin. Rheumatol.* **2013**, *25*, 426–433, doi:10.1097/BOR.0b013e328362018f.

74. Roberts, R.L.; Wallace, M.C.; Jones, G.T.; van Rij, A.M.; Merriman, T.R.; Harrison, A.; White, D.; Stamp, L.K.; Ching, D.; Highton, J.; et al. Prevalence of HLA-B27 in the New Zealand population: Effect of age and ethnicity. *Arthritis Res. Ther.* **2013**, *15*, R158, doi:10.1186/ar4341.

75. Reveille, J.D.; Hirsch, R.; Dillon, C.F.; Carroll, M.D.; Weisman, M.H. The prevalence of HLA-B27 in the US: Data from the US National Health and Nutrition Examination Survey, 2009. *Arthritis Rheum.* **2012**, *64*, 1407–1411, doi:10.1002/art.33503.

76. Khan, M.A. HLA-B27 and its subtypes in world populations: Editorial review. *Curr. Opin. Rheumatol.* **1995**, *7*, 263–269, doi:10.1097/00002281-199507000-00001.

77. Araújo, P.M.M.; Martins, J.S.; Osório, N.S. SNAPPy: A snakemake pipeline for scalable HIV-1 subtyping by phylogenetic pairing. *Virus Evol.* **2019**, *5*, doi:10.1093/ve/vez050.

78. Schultz, A.K.; Zhang, M.; Bulla, I.; Leitner, T.; Korber, B.; Morgenstern, B.; Stanke, M. jpHMM: Improving the reliability of recombination prediction in HIV-1. *Nucleic Acids Res.* **2009**, *37*, 647–651, doi:10.1093/nar/gkp371.

79. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, doi:10.1093/molbev/mst010.

80. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321, doi:10.1093/sysbio/syq010.

81. Lefort, V.; Longueville, J.E.; Gascuel, O. SMS: Smart Model Selection in PhyML. *Mol. Biol. Evol.* **2017**, *34*, 2422–2424, doi:10.1093/molbev/msx149.

82. Drummond, A.J.; Suchard, M.A.; Xie, D.; Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **2012**, *29*, 1969–1973, doi:10.1093/molbev/mss075.

83. Ferreira, M.A.R.; Suchard, M.A. Bayesian analysis of elapsed times in continuous-time Markov chains. *Can. J. Stat.* **2008**, *36*, 355–368, doi:10.1002/cjs.5550360302.

84. Araújo, P.M.M.; Carvalho, A.; Pingarilho, M.; Abecasis, A.B.; Osório, N.S. Characterization of a large cluster of HIV-1 A1 infections detected in Portugal and connected to several Western European countries. *Sci. Rep.* **2019**, *9*, 7223, doi:10.1038/s41598-019-43420-2.

85. Ribeiro-Gonçalves, B.; Francisco, A.P.; Vaz, C.; Ramirez, M.; Carriço, J.A. PHYLOViZ Online: Web-based tool for visualization, phylogenetic inference, analysis and sharing of minimum spanning trees. *Nucleic Acids Res.* **2016**, *44*, W246–W251, doi:10.1093/nar/gkw359.

86. Reynisson, B.; Alvarez, B.; Paul, S.; Peters, B.; Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **2020**, *48*, W449–W454, doi:10.1093/nar/gkaa379.

87. Gonzalez-Galarza, F.F.; McCabe, A.; Santos, E.J.M. Dos; Jones, J.; Takeshita, L.; Ortega-Rivera, N.D.; Cid-Pavon, G.M.D.; Ramsbottom, K.; Ghattaoraya, G.; Alfirevic, A.; et al. Allele frequency

net database (AFND) 2020 update: Gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Res.* **2020**, *48*, D783–D788, doi:10.1093/nar/gkz1029

Chapter 2.2. - Evolutionary dynamics of HIV-1 subtype C in Brazil

This work was published as:

Souto, B.*; Triunfante, V.*; Santos-Pereira, A.*; Martins, J.; Araújo, P.M.M.; Osório, N.S. Evolutionary dynamics of HIV-1 subtype C in Brazil. Sci Rep 11, 23060 (2021). https://doi.org/10.1038/s41598-021-02428-3

Evolutionary dynamics of HIV-1 subtype C in Brazil

1. Abstract

The extensive genetic diversity of HIV-1 is a major challenge for the prevention and treatment of HIV-1 infections. Subtype C accounts for most of the HIV-1 infections in the world but has been mainly localized in Southern Africa, Ethiopia and India. For elusive reasons, South Brazil harbors the largest HIV-1 subtype C epidemic in the American continent that is elsewhere dominated by subtype B. To investigate this topic, we collected clinical data and viral sequences from 2611 treatment-naïve patients diagnosed with HIV-1 in Brazil. Molecular epidemiology analysis supported 35 well-delimited transmission clusters of subtype C highlighting transmission within South Brazil but also from the South to all other Brazilian regions and internationally. Individuals infected with subtype C had lower probability to be immunodepressed comparing to subtype B. The HIV-1 epidemics in the South was characterized by high female-to-male infection ratios and women-to-child transmission. Our results suggest that HIV-1 subtype C probably takes advantage of longer asymptomatic periods to maximize transmission and is unlikely to outcompete subtype B in settings where the infection of women is relatively less relevant. This study contributes to elucidate factors possibly underlying the geographical distribution and expansion patterns of the most spread HIV-1 subtypes.

2. Introduction

Retroviruses such as HIV (Human Immunodeficiency Virus) have an extreme capacity to generate genetic diversity [1]. HIV genetic diversity spectrum is divided into types I and II, with HIV-1 comprising the groups M, O, N and P. The pandemic group M is increasingly diversifying and comprises at least 10 subtypes, several sub-subtypes and recombinant forms [2,3]. Interestingly, these HIV-1 clades might be evolving at different rates, to modulate virulence [4,5]. Most accepted theories on virulence evolution postulate that the selection for an optimal virulence level follows a complex trade-off between the factors influencing pathogen induced-host mortality and between-host transmission [6]. In fact, M group subtypes were associated to differences in disease progression [7–11], preferential transmission routes [12,13]

and different capacity to evade the immune system [14,15] or therapy [16–18]. These differences possibly result in subtype-related advantages in different niches contributing for the global subtype spread dynamics [5,11].

Subtype C causes nearly all infections in Southern Africa, Ethiopia and India being responsible for almost half of the HIV-1 infections in the world [19–21]. Despite the increasing amount of evidence that supports the geographic expansion of C subtype and other non-B subtypes in different continents [22–25], globally, in the last decades, subtype C has been shown to have a decreasing profile, along with other subtypes, contrasting with subtype B [20]. In fact, subtype B remains the most geographically spread HIV-1 subtype worldwide. *Ex vivo* evidence following viral infection of peripheral blood mononuclear cells suggests that C subtype might be less cytopathogenic due to a preference for CCR5 co-receptor expressing cells and less fit when compared to B [26–28]. Furthermore, it was shown that HIV-1 subtype C is associated with slower rates of CD4+ T-cell declines and higher frequencies of long-term non-progression when compared to subtype A or D in women from Uganda and Zimbabwe [29]. In cohorts from Kenya [13] or Tanzania [12] it was found that pregnant women infected with subtype C had higher risk of mother-to-child transmission when compared with the ones infected with A or D.

Studies comparing in detail subtype C and B infections in human cohorts are limited by the rarity of informative clinical settings where subtype C and B co-exist in large numbers. In case of Brazil, the HIV-1 epidemics is dominated by B subtype. However, subtype C, that was introduced in the country possibly in 1960-80s [30–32] from East Africa [33], rapidly become the most prevalent subtype in the South region of the country. Reports show that in the early 2000s, C subtype represented around 30% of the HIV-1 infections in several cities in this region and that, after 2005, it became the most prevalent subtype, representing more than 40% of the cases [30,34]. Most strikingly, the spread of C subtype in other regions of Brazil outside of the South were revealed to be slow and modest. Despite intense and regular movement of people between the South and South-East regions, the South-East region of Brazil and other regions bordering South Brazil in Argentina, Paraguay or Uruguay remain with less than 6% of subtype C prevalence [34]. The reasons underlying these regional differences are elusive and gaining insights into the introduction and regional expansion of HIV-1 subtype C in Brazil might give important information about C versus B subtype-related differences in what regards to within-host replication, virulence, transmission, and overall host population infection dynamics. Thus, in the present study, we investigated the phylogeography of HIV-1 sequences and compared clinical and epidemiological information from 2611 Brazilian patients.

3. Results

3.1. Proportion of HIV-1 subtype C infections in Brazil

To investigate the differences in the proportion of cases caused by HIV-1 C and B subtypes in Brazil, we subtyped the sequences from all individuals that were treatment naïve and sampled from 01/2008 to 04/2017 at the National Genotyping Network of Brazil (n=2611; Supplementary Table S2). The region with the higher proportion of naïve HIV-1 infected individuals was the South-East (n=1305; 49.98%) followed by the South (n=507; 19.42%) and North-East (n=487; 18.65%) (Figure 1A). HIV-1 subtype B was the most common at the country level with a total of 1675 cases, representing 64.15% of all infections in the studied population. Combining all regions, the proportion of C subtype among our sample was 13.02% (340 cases of a total of 2611). However, in the South, subtype C represented 50.30% of the cases being the most frequent in the region (Figure 1A). The analysis of the number of cases per year highlights that subtype C was consistently the most abundant in the South during the period under analysis (Figure 1B). In the South-East, the region with most HIV-1 infections, the number of cases with subtype C in the studied population never reached more than 12 cases per year, contrasting with the South, in which the number of C infections was superior to 30 cases per year in the period between 2013 and2016, with a peak of 48 cases in the year of 2014 (Figure 1B). Overall, the South was the region with highest growth in the number of cases caused by subtype C and the only region where this subtype was more frequent than subtype B (Supplementary Table S3).

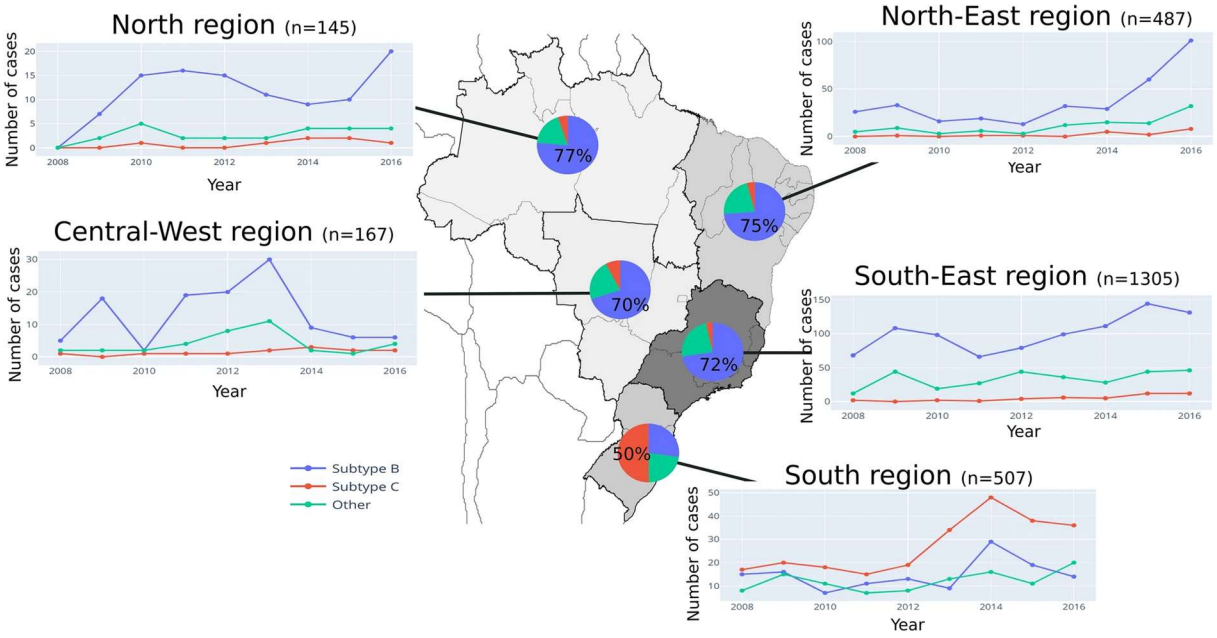


Figure 1. HIV-1 subtype distribution in Brazil nation-wide data obtained from 2008 to 2017 (n=2611). The pie charts represent the prevalence of subtypes B (blue), C (red) and other subtypes and recombinants (green) in the five macro-regions of Brazil. The shade of grey indicates the prevalence of HIV-1 (all subtypes) in each region (raging from the highest prevalence in dark grey to the lowest prevalence in light grey). The line charts represent the absolute number of HIV-1 cases of subtype B, C and other subtypes in all Brazilian regions per year.

3.2. Subtype C associates with lower immunodepression when compared with B

To address subtype-related differences in infection progression outcomes, we compared the viral loads and CD4+ T cell counts between the infections caused by subtypes C or B. For viral load comparisons, we separated individuals with viral loads predictive of slow or moderate disease progression (<=100,000 virus/mL; n=1927) from those with higher viral loads(>100,000 virus/mL; n=535), using a clinical baseline predictor value of treatment failure[35], and found no significant differences between C and B subtypes (p=0.6319; Supplementary Table S4). To investigate the effect of HIV-1 subtype in CD4+ T cell counts we compared individuals with or without immunodepression and, among the immunodepressed, the ones with moderate or severe levels. Considering the age-related differences of CD4+ T cell normality, the classification of each case was done by adjusting the reference values

according to the age of the subject (Supplementary Table S5). Individuals infected with subtype C had a significant lower probability to be immunodepressed (p=0.000) when compared with subtype B (Figure 2; Supplementary Table S6). This association was maintained when dividing the group by age (Supplementary Table S6). Among the individuals with immunodepression, the ones infected with subtype C had significant lower probability of severe immunodepression (p=0.001; Figure 2; Supplementary Table S6). Individuals with less than 18 years infected with subtype C had an even lower probability of severe immunodepression (p=0.008; Supplementary Table S6). Moreover, we decided to evaluate the proportion of ambiguous sites (PAS), a surrogate of age of infection [36–39], on all the viral sequences and no significant difference was found between subtypes (Supplementary Table S7). Overall, these results suggest that C subtype viruses, despite reaching viral loads similar to subtype B, are less able to cause immunodepression, which could lead to longer asymptomatic periods and possibly increase the opportunity for transmission in some settings.

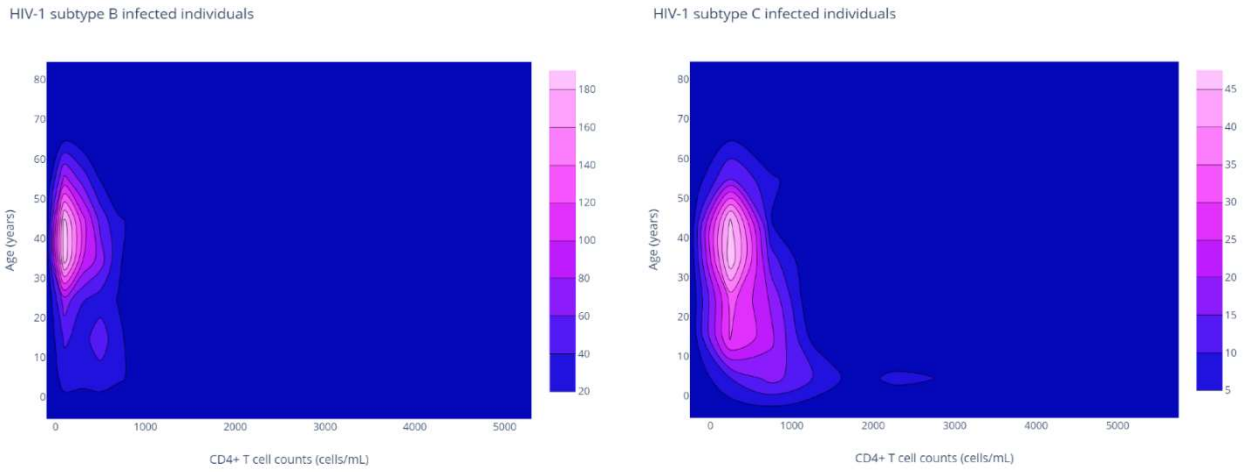


Figure 2. Age and CD4+ T cell counts in individuals infected with HIV-1 subtype B or C. The 2D contour histograms represent the Age (years) in the y axis and CD4+ T cell counts (cells/mL) in the x axis for all studied cases of HIV-1 subtype B infection (n=1675) and subtype C infection (n=340). The color scale represents the frequency of cases from blue (lower) to pink (higher).

3.3. Evidence for interregional and international subtype C transmission

To gain insights into the transmission of subtype C in Brazil, we performed maximum likelihood (ML) and Bayesian phylogenetic analysis of the 340 subtype C sequences described in this study and 854 closely related sequences obtained from public databases (total n=1194). The phylogenetic representation (Figure 3) demonstrated that the vast majority (99.26%; 1076 out of 1084) of the C subtype viruses isolated in Brazil were included in a monophyletic clade that shared ancestors with sequences from the East African region. This large clade also included sequences obtained from public databases and isolated in Asia, Europe, and other American countries. We then performed the characterization of transmission clusters and found 35 well-delimited transmission clusters (TC1 to TC35, Table 1) involving a total of 174 sequences. The average number of sequences per cluster was 4.97. TC24 was the largest cluster including a total of 24 sequences isolated in the South-East, Central-West or North regions of Brazil. Most of the clusters (18 out of 35, 51.43%) were exclusively formed from sequences isolated in the South of Brazil. From the nine clusters spanning more than one Brazilian region (Table 1, interregional), only clusters TC24 and TC35 did not include sequences isolated in the South. Interestingly, we found four clusters that included sequences isolated outside Brazil (Table 1, international). The two largest (TC5, TC16) included sequences from the South, South-East, one other Brazilian region and other countries (USA, Spain, Portugal, or Germany). Most of the sequences in clusters (71.26%, 124 out of 174) lacked information on the reported route of infection. Among those with available data, transmission between heterosexual (28 cases) and MSM (12 cases) were the most reported. The estimates for the time of the MRCA for each cluster ranged from 1992-2009. The cluster depth analysis (obtained for each cluster by subtracting the most recent sampling date minus the time of MRCA) supports that six transmission clusters (TCs 2, 5, 16, 22, 24 and 26) were ongoing for more than 20 years (Table 1).

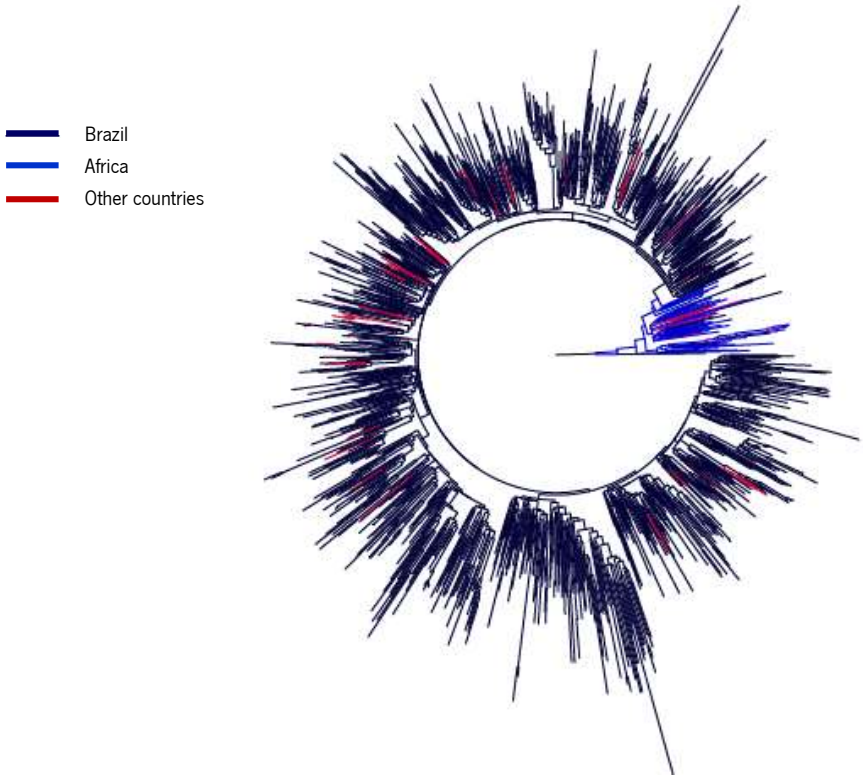


Figure 3. Maximum likelihood tree of the HIV-1 subtype C sequences isolated in Brazil from treatment naïve individuals from 2008-2017 (n=340) and closely related sequences from databases (total n=1194). Branch colors indicate the geographical origin of the sequences. Most subtype C sequences isolated from patients in Brazil are monophyletic suggesting one single major founder event.

Table 1. Characterization of the 35 transmission clusters of HIV-1 subtype C virus identified in this study.

Cluster	Number of individuals				Cluster expansion	Place of sampling (n)							Sampling date range	Time of MRCA (95% HPD, years)	Cluster Depth (years)
						Brazil					Other	Missing			
	Total	Male	Female	N/A		N	S	NE	SE	CW					
TC5	10	6	1	3	international	0	1	1	5	0	3[a]	0	2009.0-2017.2	1992.7 (1984.8-1999.6)	24.5
TC16	9	5	0	4		0	4	0	2	1	2[c]	0	2008.0-2017.1	1996.3 (1990.6-2001.4)	20.8
TC19	6	0	0	6		0	0	0	0	0	6[b]	0	N/A	N/A	N/A
TC17	3	1	1	1		0	2	0	0	0	1[d]	0	2007.0-2017.2	1998.5 (1993.1-2003.3)	18.7
TC24	24	16	4	4	interregional	3	0	0	13	8	0	0	2010.0-2017.2	1992.9 (1985.7-1999.7)	24.3
TC26	9	4	1	4		0	5	0	3	0	0	1	2006.0-2017.0	1995.9 (1989.4-2001.8)	21.1
TC9	6	2	4	0		1	1	0	1	3	0	0	2010.0-2017.0	1997.1 (1991.4-2002.4)	19.9
TC29	5	2	2	1		0	4	0	0	1	0	0	2008.0-2014.8	1999.7 (1994.7-2004.1)	15.1
TC34	4	1	2	1		0	1	0	3	0	0	0	2013.0-2016.8	2003.2 (1997.7-2008.5)	13.6
TC31	3	0	3	0		0	1	2	0	0	0	0	2016.7-2016.8	2000.9 (1994.9-2006.7)	15.9
TC27	3	0	3	0		0	2	0	1	0	0	0	2016.1-2016.2	2008.5 (2002.2-2013.5)	7.7
TC32	3	0	3	0		0	2	0	1	0	0	0	2016.2-2016.8	2002.5	14.3

														(1997-2008.3)	
TC35	3	0	1	2		0	0	1	1	1	0	0	2006.0-2016.7	1999.5 (1995.3-2003.6)	17.2
TC2	7	2	1	4		5	0	0	0	0	0	2	2004.0-2016.8	1992.2 (1982.3-2000.6)	24.6
TC6	6	1	2	3		0	6	0	0	0	0	0	2012.0-2015.5	1996.3 (1990.1-2002)	19.2
TC12	5	1	0	4		0	5	0	0	0	0	0	2009.8-2015.0	2006.9 (2003.1-2009.8)	8.1
TC22	5	2	2	1		0	5	0	0	0	0	0	2008.8-2015.9	1995.5 (1988.6-2001.7)	20.4
TC28	5	5	0	0		0	5	0	0	0	0	0	2013.2-2017.2	2004.9 (1999.8-2009.7)	12.3
TC7	4	0	4	0		0	4	0	0	0	0	0	2013.3-2017.1	2000.9 (1995.3-2006.1)	16.2
TC1	4	3	0	1	regional	0	4	0	0	0	0	0	2008.0-2016.9	2000.2 (1995.3-2004.8)	16.7
TC13	4	0	4	0		0	4	0	0	0	0	0	2012.6-2014.5	1999.8 (1993.9-2005.3)	14.7
TC14	4	2	2	0		0	4	0	0	0	0	0	2011.3-2015.5	1999.8 (1993.7-2005.2)	15.7
TC11	4	1	3	0		0	4	0	0	0	0	0	2015.9-2017.3	1999.2 (1993.2-2005.2)	18.1
TC30	4	1	1	2		0	4	0	0	0	0	0	2008.0-2014.4	1999.3 (1994.3-2003.8)	15.1
TC20	4	0	4	0		0	4	0	0	0	0	0	2008.4-2015.3	1998.1 (1992.3-2003.5)	17.2
TC21	3	1	2	0		0	3	0	0	0	0	0	2012.2-2013.2	2000.9 (1995.1-2006.5)	12.3
TC33	3	2	1	0		0	3	0	0	0	0	0	2014.4-2015.6	2001.3	14.3

													(1995.6-2007.1)	
TC4	3	1	2	0		0	3	0	0	0	0	0	2013.8-2016.7	2009.5 (2004.8-2013.2)	7.2
TC15	3	1	1	1		0	3	0	0	0	0	0	2008.0-2012.2	1998.8 (1993.4-2004.1)	13.4
TC18	3	0	3	0		0	3	0	0	0	0	0	2011.9-2014.2	2009.1 (2005.4-2011.8)	5.1
TC10	3	1	2	0		0	0	0	3	0	0	0	2010.4-2014.1	2006.7 (2002.6-2010)	7.4
TC8	3	2	1	0		0	3	0	0	0	0	0	2015.8-2017.2	1998.8 (1992-2005.5)	18.4
TC23	3	0	3	0		0	3	0	0	0	0	0	2013.4-2014.8	2001.3 (1995.7-2006.5)	13.5
TC3	3	0	0	3		0	0	0	2	0	0	1	2007.0	2002.1 (1998.4-2005.4)	4.9
TC25	3	0	2	1		2	0	0	0	0	0	1	2010.0-2010.4	1998.8 (1992.4-2004)	11.6

a) USA (1), Spain (1), Portugal (1); b) UK (6); c) Germany(1), Spain(1); d) Germany(1).

3.4. Relevant role of the South-East in the transmission of subtype C

A phylogeographic analysis was performed using sample collection location (the Brazilian state or the country outside Brazil) as discrete traits (Figure 4). The root of subtype C epidemic in Brazil was inferred to Santa Catarina (SC, pp=0.97) in the South. Transmission to outside of the South was estimated to be as early as 1992 (1985.7-1999.7) with the states of São Paulo (SP) and Manaus/Amazonas (AM), having the highest posterior probability (pp=0.9702 and pp=0.9609, respectively) of being the first points of introduction in other regions. A total of 13 pairwise rates of diffusion between locations were found to have a strong support value (Bayes factor >10, Supplementary Table S8). These include the transmission among Southern states (SC-RS, SC-PR), between the South and the South-East (SC-SP, SC-RJ, SC-MG) or the Central-West (RS-GO). Importantly, transmission relating the South-Eastern states (SP, RJ) and the

Central-West (MS, DF), North (PA, RO) or North Eastern (BA) states were also strongly supported. This shows that, at a given point in the transmission history, the South-East not only received C viruses from the South, but was also a relevant transmission donor to all the other Brazilian regions. The well supported diffusion rates also include international transmission of the Brazilian C subtype clade relating the Southern state of RS with Germany and Spain with the United States of America.

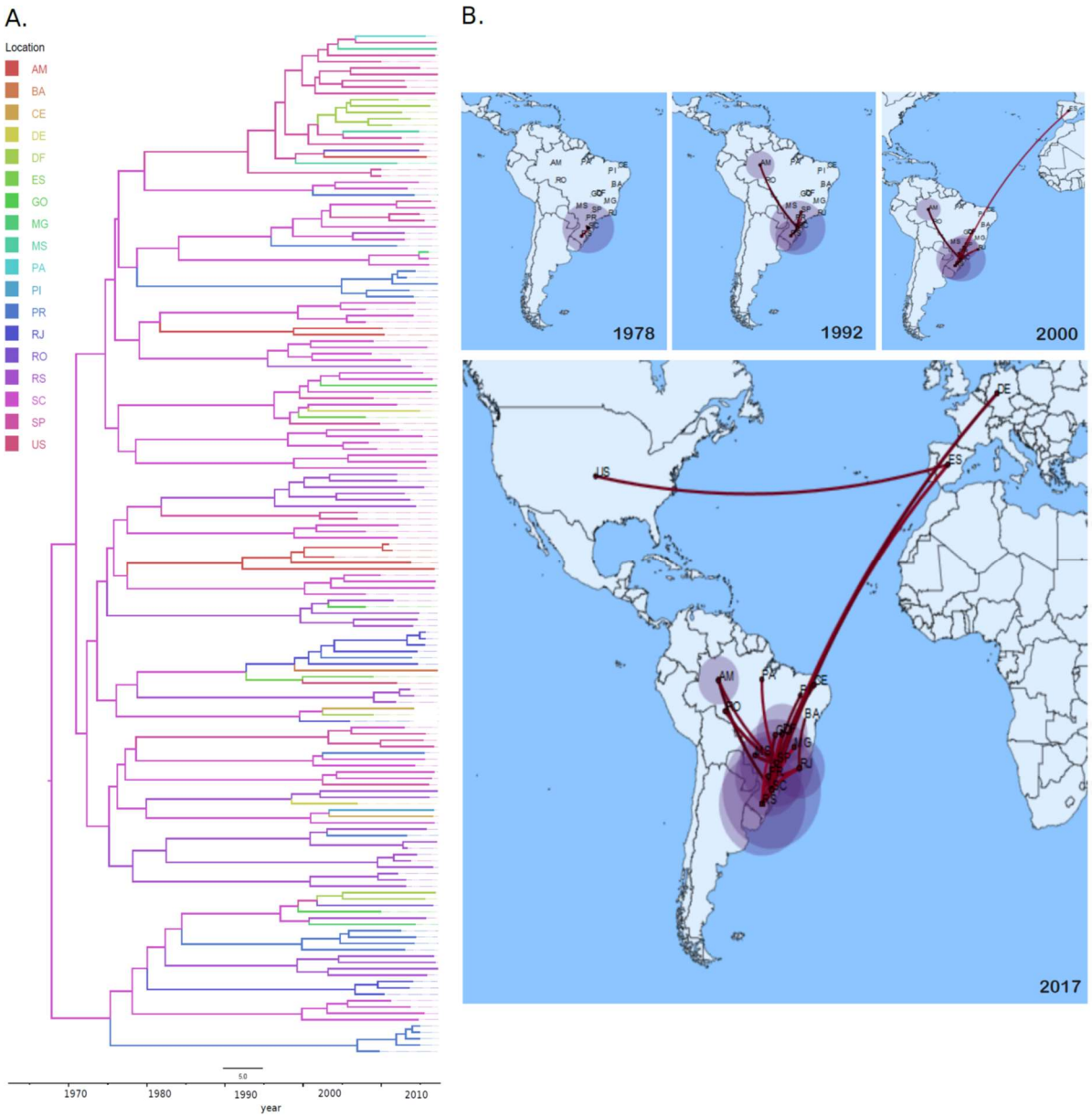


Figure 4. Phylogeographic analysis of the evolution of HIV-1 subtype C transmission clusters. (a) Bayesian MCC time scaled discrete phylogeographic tree built using BEAST with the Brazilian state or Country outside Brazil where the viral sequence was isolated as the discrete trait. (b)Geographical representation of this transmission history. Acronyms: AM (Amazonas), BA (Bahia), CE (Ceará), DF (Distrito Federal), GO (Goiás), MG (Minas Gerais), MS (Mato Grosso do Sul), PA (Pará), PI (Piauí), PR (Paraná), RJ (Rio de Janeiro), RO (Rondônia), RS (Rio Grande do Sul), SC (Santa Catarina), SP (São Paulo), DE (Germany), ES (Spain) and US (USA).

3.5. Demographic differences in subtype C infections

Having established evidence for intense South to South-East transmission of subtype C, we then explored the demographic and epidemiologic characteristic of the HIV-1 epidemics in these regions to investigate possible reasons for the inferior capacity of C subtype to become dominant outside the South. In total, 60.18% (204 of 339) of the infections by subtype C in Brazil were in women and only 39.82% (135 of 339) in men (OR=1.64; CI=1.30-2.08; p=0.000). In accordance, we found significant differences (p=0.0160) in the distribution of the sex of the HIV-1 infected individuals in the South when compared with the South-East (Table 2). In the South, HIV-1 affected more females (55.82%) than males (female-to-male ratio=1.27) while in the South-East most of the infections were in male (50.04%; female-to-male ratio=0.98). Despite the missing data, within our study population mother-to-child transmission was significantly more likely (CI =1.76-6.96; p=0.0002) to occur in the South than in the South-East. Moreover, the number of infected individuals with less than 18 years of age infected with HIV-1 in the South was also significantly (p=0.000) higher than in the South-East. Transmission between men that have sex with men (MSM) was significantly more associated with the South-East (OR=3.72; CI=1.22-15.13; p=0.0218). These findings highlight clear demographic and epidemiological differences between these two neighboring Brazilian regions.

Table 2. Epidemiological comparison of HIV-1 epidemics between South and South-East Brazil.

		Brazil	South	South-East	OR*	CI	p value
Sex	Male	1314 (50.33%)	223 (43.98%)	653 (50.04%)	0.78	0.63-0.95	0.016
	Female	1286 (49.25%)	283 (55.82%)	643 (49.27%)	1.29	1.05-1.58	
	Missing	11 (0.42%)	1 (0.20%)	9 (0.69%)	0.28	0.01-2.06	0.359
Age	< 18 yr	702 (26.89%)	165 (32.54%)	276 (21.15%)	1.80	1.43-2.26	0.000
	>= 18 yr	1903 (72.88%)	341 (67.26%)	1027 (78.70%)	0,55	0.44-0.70	
	Missing	6 (0.23%)	1 (0.20%)	2 (0.15%)	1.29	0.02-23.78	1.000
Transmission route	MSM	50 (1.91%)	4 (0.79%)	32 (2.45%)	0.27	0.07-0.82	0.022
	Women-to-child	82 (3.14%)	24 (4.73%)	25 (1.92%)	3.49	1.76-6.96	0.000
	Heterosexual	165 (6.32%)	27 (5.33%)	82 (6.28%)	0.70	0.38-1.30	0.260
	Other	12 (0.46%)	2 (0.39%)	7 (0.54%)	0.72	0.07-3.96	0.984
	Missing	2302 (88.17%)	450 (88.76%)	1159 (88.81%)	0.99	0.72-1.38	0.974
	total	2611	507	1305			

* The OR, CI and p value are presented for the South vs. South-East comparison using two-tailed Fisher Exact Test or Corrected Mantel-Haenszel chi-square test were applicable.

4. Discussion

The large HIV-1 subtype diversity adds a considerable layer of complexity for the control of the HIV-1 pandemics. HIV-1 has an extreme capacity to generate diversity within the host and takes advantage of a highly diverse population-level genetic pool for efficient adaptation. HIV-1 subtypes C and B can be considered the evolutionarily most successful subtypes. C is the most prevalent being responsible for more than 50% of the HIV infections in word but it is mainly localized to Southern Africa, Ethiopia and India [19–21]. B is the most spread worldwide. Given the differences in geographic distribution between C and B subtypes it is reasonable to assume that there are particularities in these viruses possibly conferring subtype-specific advantages in different settings. Brazil has the ideal characteristics to study

the factors underlying the capacity of C or B subtypes to outcompete each other in a population. This large American country has bordering regions dominated in prevalence by subtype C or B and more than two decades of detailed and comprehensive clinical records on HIV infections [40]. In this study, country level data was collected from 2008-2017 and 2611 cases met the inclusion criteria. The observed proportion of HIV-1 infections by Brazilian region in the study population was in accordance with the official HIV-1 prevalence reports [40]. Regarding HIV-1 subtype distribution in Brazil, our results update and expand to the country-level previous literature that reported only state or regional data [41–43].

Our results, clearly show that, during the period under analysis, the only region where subtype C dominated in proportion was the South. Most interestingly, despite intense and regular movement of people between the South and South-East regions, the lowest overall subtype C proportion of cases in the studied population was found in the South-East (3.52%; 44 cases out of 1248). Moreover, the proportion of C subtype infections increased in all Brazilian regions at an average rate of 1.16% per year. The highest growth rate was found in the South (1.73% per year) while the bordering South-East region presented one of the lowest subtype C proportion growth rates in the country (0.67% per year).

Subtype C was previously associated with higher CD4+ T cell counts in African cohorts when compared with subtypes A and D [29]. In the comparison with B subtype, our analysis in the Brazilian cohort suggests that Subtype C, despite reaching similar viral loads than subtype B, often leads to more moderate rates of destruction of CD4+ T cells. In fact, among people infected with subtype C there were significantly less individuals with immunodepression when compared with the ones infected with subtype B, which was not due to differences on the age of the infected individuals or in the time since infection, as no significant differences were observed on the statistical analysis of the PAS [36–39]. This could lead to longer asymptomatic periods in subtype C infections and possibly increased opportunities for transmission. To investigate the C subtype transmission, we performed a molecular epidemiology and phylogeographical analysis using the 340 C subtype sequences obtained in this study and the closest related sequences from databases. This was performed to enrich the information that could be obtained in what regards to the transmission outside Brazil. Our analysis generated information on the origin and probable place of introduction of C subtype in Brazil. In accordance with the previous studies [30,32], we found strong evidence supporting one major founding event of introduction of subtype C in Brazil originating from Middle East African countries. We found no evidence supporting the introduction from UK to Brazil as suggested in one study [31]. We did find a transmission cluster (TC19) with sequences isolated in the UK that likely originated in Brazil and was transmitted to the UK. We found strong statistical

support for international transmission from the Southern Brazil state of Rio Grande do Sul (RS) to Germany. This link is possibly explained by the known migratory fluxes between these two geographic locations.

In our analysis, the state with the highest probability for the place of entrance of subtype C in Brazil was Santa Catarina (SC). The characterization of transmission clusters and phygeographic dynamics suggests that the inferior capacity of C subtype to thrive outside the South was not due to absence of cross-regional transmission. In fact, we found that more than 20% of the C subtype transmission events bridged, in the last decades, the South and at least one other Brazilian region with emphasis on the South-East. We found strong statistical support indicating that the states of São Paulo and Rio de Janeiro were not only recipients but also donors in interregional transmission clusters of subtype C viruses. This suggests that, although the South-East have among the lowest overall proportion and annual growth rate of subtype C in the country it played a role in disseminating C subtype virus to other Brazilian regions. Considering our results, it is tempting to speculate that for HIV-1 subtype C to thrive in a population it relies on its high within-host replicative capacity (like that of B subtype) but possibly also takes advantage of longer asymptomatic periods that might increase its opportunities to transmit. The epidemiological comparison between the South and South-East Brazil suggests that C subtype capacity to outcompete B might be facilitated in settings with higher female-to-male infection ratios and women-to-child transmission. However, these conclusions are limited by the presence of missing data on the reported route of infection and to what is possible by means of a cross-sectional study. Notwithstanding, this data finds parallels in previous studies in African cohorts showing that C subtype was more adapted to women-to-child transmission than A or D subtypes [12,13]. In a Kenyan cohort, it was found that pregnant women infected with subtype C were significantly more likely to shed HIV-1–infected vaginal cells than were those infected with subtype A or D [13]. Whether C subtype virus are present in higher levels in cells from the vaginal mucosa or even breast milk when compared to B subtype virus has not, to our best knowledge, been studied, being a matter for future investigation. On the other hand, the distribution of HIV infection among men, women and children is also influenced by sociocultural factors such as breast feeding and other gender equality-related factors. It is relevant to point out that the practice of cross-breastfeeding was a culturally established and accepted behavior in Brazil [44,45]. It was initially provided by lactating slaves mainly originating from the same African regions that are the most probable point of origin of the HIV-1 subtype C introduced in Brazil. Long after slavery was abolished and at least until the first half of the XX century, it was frequent that lactating Afro-Brazilian women were

paid to cross-breastfeed [45]. It is possible that sociocultural heritages from this past influenced the introduction and transmission of subtype C and, consequently, its distribution in the Brazilian territory. The South has the highest prevalence in Brazil of AIDS in pregnant women and children and the higher female-to-male infection ratio [40]. Apparently, subtype C might have found in the Southern region of Brazil, sociocultural and behavioral conditions favorable to its dissemination with similarities to those found in other African and Asian regions, where it is also the most prevalent HIV-1 subtype [20,46].

Overall, this study opens lines of research on the differences between the two most prevalent HIV-1 subtypes and, at the same time, it is useful for the management of the health care and public HIV-1 control policies. Regarding the dynamics between B and C subtypes it is possible that C subtype outcompetes B only in settings with sizable infection of women and women-to-child transmission. Thus, it is suggested that, where the prevalence of subtype C is higher, care professionals and public policies define specific strategies for the protection of women and the pregnancy-puerperal cycle against HIV infection. Targeting this group by close surveillance to make the diagnosis and treatment as close as possible to the time of infection is likely to reduce the epidemiological burden of subtype C HIV-1 infections.

5. Materials and Methods

5.1. Study Population

Data was collected from the HIV-1 infected patients followed at the Specialized Assistance Services on Sexually Transmissible Diseases and HIV/Aids located in all 26 Brazilian states and Federal District from 01/01/2008 to 04/30/2017 that met the inclusion criteria for this study (n=2611, Supplementary Table S2). The inclusion criteria were availability of partial HIV-1 genome sequence, and absence of previous antiretroviral treatment upon sampling. For all cases matching the criteria the following patient secondary data was collected: self-reported transmission route; sex; birth year; date of the viral sample collection for sequencing; CD4+ T-cell count at sampling; viral load at sampling; geographical origin of the sample; and pregnancy. The collection of the patient's data was performed anonymously after approval by the Brazilian national ethic committee through the protocol CAAE 53757016.0.0000.5504.

5.2. Sequencing

DNA sequencing, from the reverse transcriptase PCR amplicons was performed with commercially available HIV-1 genotyping systems based on Sanger sequencing and performed using standardized protocols in the National Genotyping Network of Brazil. The HIV-1 genome sequence portion used in this study corresponds to the pol region. The HIV-1 positions in this study refer to the HXB2 HIV-1 reference genome (GenBank: K03455.1). All multiple sequence alignments were performed using MAFFT v7.309[47]removing columns containing at least 10% gaps. The HIV-1 subtype was assigned using Rega HIV-1 Subtyping Tool v3.0[48], Comet HIV-1 v2.3[49], jpHMM[50], RIP v3.0[51], SCUEAL[52], and SNAPPY[53]. The results of the different tools were compared, and subtype was classified based on the agreement between the used tools and manual inspection of the results from phylogenetic and recombination analysis. The 2611 sequences selected for this study were made available in GenBank (accession numbers pending).

5.3. Phylogenetic analysis

To obtain additional sequences from outside the National Genotyping Network of Brazil we queried the HIV reference sequence database (http://www.hiv.lanl.gov/) using BLAST[54]. For each of the 340 subtype C sequences described in this study the 10 most closely related generated outputs were selected. We excluded duplicates or sequences from the same patient and sequences showing evidence of recombination - all sequences were analysed by the subtyping tools as mentioned above - Rega HIV-1 Subtyping Tool v3.0[48], Comet HIV-1 v2.3[49], jpHMM[50], RIP v3.0[51], SCUEAL[52] and SNAPPY[53] - and only sequences classified as subtype C, based on the agreement between the different tools, were selected. Applying these criteria 854 database sequences were added to this study for phylogenetic analysis. An alignment of 1194 sequences was used to make a phylogenetic reconstruction using PhyML v3.0[55]. The best fitting substitution model was GTR + G4 + I, determined by PhyML SMS(Smart Model Selection) using AIC (Akaike Information Criterion)[56].The heuristic trees search was performed using SPR and NNI methods. The branch support was calculated with the approximate likelihood-ratio (aLRT) SH-like test. The tree with the best likelihood value was performed using SPR with 3 random starting trees (Fig. 2).Bayesian evolutionary and phylogeographic analyses were performed using BEAST

v1.10.4[57][58], with GTR + G4 + I for two different codon partitions (1 + 2, 3), as nucleotide substitution model, coalescent Skygrid model and uncorrelated relaxed clock. The site model GTR + G4 + I corresponding to the best model selected by jModelTest program[59]. Two different runs (random seeds) of 320 million generations, converged to similar values. Outputs were analysed with Tracer v1.7.1[60] to ensure all parameters had an effective sampling size (ESS) superior to 300. The two multiple tree output files were combined, using LogCombiner v1.10.4[57], excluding 10% burn-in, and used to build the maximum clade credibility tree with mean heights using TreeAnnotator v1.10.4[57]. The Skygrid plot was also created (Supplementary Fig. S1). In the phylogeographic analysis the sampling Brazilian state or country outside Brazil were used as a discrete trait, with a total of 18 different locations (Fig. 3).

5.4. Definition of transmission cluster and tree visualization

The criteria for the definition of a clade as a transmission cluster were likelihood ratio test (aLRT) SH-like branch support ≥0.95 (estimated with PhyML v3) with GTR + G4 + I as the beast fitting substitution (determined by PhyML SMS using AIC) and with heuristic trees search performed by using SPR and NNI methods. The tree with the best likelihood value was performed using SPR with 3 random starting tree; branch posterior probability ≥0.99 (estimated with BEAST v1.10.4), with GTR + G4 + I for two different codon partitions (1 + 2, 3), as nucleotide substitution model, coalescent Skygrid model and uncorrelated relaxed clock; mean cluster genetic distance <0.003 substitutions per site; and maximum genetic distance <0.05 substitutions per site. MEGA X v10.05 [61] was used for genetic distance calculation. Only clusters with more than 2 sequences were included. For the visualization and manipulation of the trees the software FigTree v1.4.4 were used. The phylogeographic representation was created with SpreaD3 [62]. For database sequences from outside Brazil, the country's locations were plotted as their geographic centre.

5.5. Statistical analysis

After verifying and optimizing the quality of epidemiological data (transmission route; sex; birth year; date of the viral sample collection for sequencing; CD4+ T-cell count at sampling; viral load at sampling; geographical origin of the sample), they were organized into spreadsheets and processed by

the software Epi Info, from the Center for Disease Control and Prevention (United States). For statistical analysis, the Mantel-Haenszel chi-square test was used when the minimum sample size in all variables was greater than or equal to 5. When sample size was less than 5 units in at least one of the variables, the Fisher exact test was used for calculating the Odds Ratio and the corrected Mantel-Haenszel chi-square test for calculating the p-value. In all cases, the tests were two-tailed, and the level of significance considered was 5%.

6. References

1. Peeters, M.; D'Arc, M.; Delaporte, E. Origin and diversity of human retroviruses. *AIDS Rev.* 2014, *16*, 23–34, PMCID: PMC4289907.

2. Robertson, D.L.; Anderson, J.P.; Bradac, J.A.; Carr, J.K.; Foley, B.; Funkhouser, R.K.; Gao, F.; Hahn, B.H.; Kalish, M.L.; Kuiken, C.; et al. HIV-1 nomenclature proposal. *Science* **2000**, *288*, 55–65, doi:10.1126/science.288.5463.55d.

3. Yamaguchi, J.; Vallari, A.; McArthur, C.; Sthreshley, L.; Cloherty, G.A.; Berg, M.G.; Rodgers, M.A. Brief Report: Complete Genome Sequence of CG-0018a-01 Establishes HIV-1 Subtype L. *J. Acquir. Immune Defic. Syndr.* **2020**, *83*, 319–322, doi:10.1097/QAI.0000000000002246.

4. Jefferys, R.J. Evidence for HIV weakening over time. *Proc. Natl. Acad. Sci. U. S. A.* 2015, doi: 10.1073/pnas.1502380112

5. Ariën, K.K.; Vanham, G.; Arts, E.J. Is HIV-1 evolving to a less virulent form in humans? *Nat. Rev. Microbiol.* **2007**, doi:10.1038/nrmicro1594.

6. Alizon, S.; Hurford, A.; Mideo, N.; Van Baalen, M. Virulence evolution and the trade-off hypothesis: History, current state of affairs and the future. *J. Evol. Biol.* 2009, doi: 10.1111/j.1420-9101.2008.01658.x.

7. Kaleebu, P.; French, N.; Mahe, C.; Yirrell, D.; Watera, C.; Lyagoba, F.; Nakiyingi, J.; Rutebemberwa, A.; Morgan, D.; Weber, J.; et al. Effect of Human Immunodeficiency Virus (HIV) Type 1 Envelope Subtypes A and D on Disease Progression in a Large Cohort of HIV-1–Positive Persons in Uganda. *J. Infect. Dis.* **2002**, doi:10.1086/340130.

8. Easterbrook, P.J.; Smith, M.; Mullen, J.; Shea, S.O.; Chrystie, I.; Ruiter, A. De; Tatt, I.D.; Geretti, A.M.; Zuckerman, M. Impact of HIV-1 viral subtype on disease progression and response to antiretroviral therapy. *J. Int. AIDS Soc.* **2010**, doi: 10.1186/1758-2652-13-4.

9. Kiwanuka, N.; Laeyendecker, O.; Robb, M.; Kigozi, G.; Arroyo, M.; McCutchan, F.; Eller, L.A.; Eller, M.; Makumbi, F.; Birx, D.; et al. Effect of human immunodeficiency virus Type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. *J. Infect. Dis.* **2008**, *197*, 707–713, doi:10.1086/527416.

10. Baeten, J.M.; Chohan, B.; Lavreys, L.; Chohan, V.; McClelland, R.S.; Certain, L.; Mandaliya, K.; Jaoko, W.; Overbaugh, J. HIV-1 Subtype D Infection Is Associated with Faster Disease Progression than Subtype A in Spite of Similar Plasma HIV-1 Loads. *J. Infect. Dis.* **2007**, *195*, 1177–1180, doi:10.1086/512682.

11. Araújo, P.M.M.; Carvalho, A.; Pingarilho, M.; Faria, D.; Pinho, R.; Ferreira, J.; Proença, P.; Nunes, S.; Mouro, M.; Teófilo, E.; et al. Characterization of a large cluster of HIV-1 A1 infections detected in Portugal and connected to several Western European countries. *Sci. Rep.* **2019**, doi:10.1038/s41598-019-43420-2.

12. Renjifo, B.; Gilbert, P.; Chaplin, B.; Msamanga, G.; Mwakagile, D.; Fawzi, W.; Essex, M. Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. *AIDS* **2004**, *18*, 1629–1636, doi:10.1097/01.aids.0000131392.68597.34.

13. John-Stewart, G.C.; Nduati, R.W.; Rousseau, C.M.; Mbori-Ngacha, D.A.; Richardson, B.A.; Rainwater, S.; Panteleeff, D.D.; Overbaugh, J. Subtype C Is associated with increased vaginal shedding of HIV-1. *J. Infect. Dis.* **2005**, *192*, 492–6, doi:10.1086/431514.

14. Serwanga, J.; Nakiboneka, R.; Mugaba, S.; Magambo, B.; Ndembi, N.; Gotch, F.; Kaleebu, P. Frequencies of Gag-restricted T-cell escape "footprints" differ across HIV-1 clades A1 and D chronically infected Ugandans irrespective of host HLA B alleles. *Vaccine* **2015**, doi:10.1016/j.vaccine.2015.02.037.

15. Kinloch, N.N.; Lee, G.Q.; Carlson, J.M.; Jin, S.W.; Brumme, C.J.; Byakwaga, H.; Muzoora, C.; Bwana, M.B.; Cobarrubias, K.D.; Hunt, P.W.; et al. Genotypic and Mechanistic Characterization of Subtype-Specific HIV Adaptation to Host Cellular Immunity. *J. Virol.* **2018**, *93*, 1502–1520, doi:10.1128/jvi.01502-18.

16. Brenner, B.; Turner, D.; Oliveira, M.; Moisi, D.; Detorio, M.; Carobene, M.; Marlink, R.G.; Schapiro, J.; Roger, M.; Wainberg, M.A. A V106M mutation in HIV-1 clade C viruses exposed to efavirenz confers cross-resistance to non-nucleoside reverse transcriptase inhibitors. *AIDS* **2003**, *17*, F1-5, doi:10.1097/01.aids.0000042957.95433.6c.

17. Abecasis, A.B.; Deforche, K.; Bacheler, L.T.; McKenna, P.; Carvalho, A.P.; Gomes, P.; Vandamme, A.M.; Camacho, R.J. Investigation of baseline susceptibility to protease inhibitors in HIV-1 subtypes C, F, G and CRF02_AG. *Antivir. Ther.* **2006**.

18. Abecasis, A.B.; Deforche, K.; Snoeck, J.; Bacheler, L.T.; McKenna, P.; Carvalho, A.P.; Gomes, P.; Camacho, R.J.; Vandamme, A.-M. Protease mutation M89I/V is linked to therapy failure in patients infected with the HIV-1 non-B subtypes C, F or G. *AIDS* **2005**, *19*, 1799–806.

19. Gartner, M.J.; Roche, M.; Churchill, M.J.; Gorry, P.R.; Flynn, J.K. Understanding the mechanisms driving the spread of subtype C HIV-1. *EBioMedicine* 2020, *53*, 102682, doi: 10.1016/j.ebiom.2020.102682.

20. Hemelaar, J.; Elangovan, R.; Yun, J.; Dickson-Tetteh, L.; Fleminger, I.; Kirtley, S.; Williams, B.; Gouws-Williams, E.; Ghys, P.D.; Abimiku, A.G.; et al. Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *Lancet Infect. Dis.* **2019**, doi:10.1016/S1473-3099(18)30647-9.

21. Santos-Pereira, A.; Magalhães, C.; Araújo, P.M.M.; Osório, N.S. Evolutionary Genetics of Mycobacterium Tuberculosis and HIV-1: "The Tortoise and the Hare." *Microorganisms* **2021**, *9*, 147, doi:10.3390/microorganisms9010147.

22. Soares, E.A.; Martinez, A.M.; Souza, T.M.; Santos, A.F.; Da Hora, V.; Silveira, J.; Bastos, F.I.; Tanuri, A.; Soares, M.A. HIV-1 subtype C dissemination in southern Brazil. *Aids* **2005**, *19 Suppl 4*, S81–6, doi:00002030-200510004-00014 [pii].

23. Vidal, N.; Mulanga, C.; Bazepeo, S.E.; Mwamba, J.K.; Tshimpaka, J.W.; Kashi, M.; Mama, N.; Laurent, C.; Lepira, F.; Delaporte, E.; et al. Distribution of HIV-1 variants in the Democratic Republic of Congo suggests increase of subtype C in Kinshasa between 1997 and 2002. *J. Acquir. Immune Defic. Syndr.* **2005**, doi:10.1097/01.qai.0000159670.18326.94.

24. Li, D.Q.; Zheng, X.W.; Zhang, G.Y. Study on the distribution HIV-1 C subtype in Ruili and other counties, Yunnan, China. *Zhonghua Liu Xing Bing Xue Za Zhi* **1996**, *17*, 337–339.

25. Carvalho, A.; Costa, P.; Triunfante, V.; Branca, F.; Rodrigues, F.; Santos, C.L.; Correia-Neves, M.; Saraiva, M.; Lecour, H.; Castro, A.G.; et al. Analysis of a local HIV-1 epidemic in portugal highlights established transmission of non-B and non-G subtypes. *J. Clin. Microbiol.* **2015**, *53*, 1506–14, doi:10.1128/JCM.03611-14.

26. Ball, S.C.; Abraha, A.; Collins, K.R.; Marozsan, A.J.; Baird, H.; Quinones-Mateu, M.E.; Penn-Nicholson, A.; Murray, M.; Richard, N.; Lobritz, M.; et al. Comparing the Ex Vivo Fitness of CCR5-Tropic Human Immunodeficiency Virus Type 1 Isolates of Subtypes B and C. *J. Virol.* **2003**, doi:10.1128/jvi.77.2.1021-1038.2003.

27. Abraha, A.; Nankya, I.L.; Gibson, R.; Demers, K.; Tebit, D.M.; Johnston, E.; Katzenstein, D.; Siddiqui, A.; Herrera, C.; Fischetti, L.; et al. CCR5- and CXCR4-Tropic Subtype C Human Immunodeficiency Virus Type 1 Isolates Have a Lower Level of Pathogenic Fitness than Other Dominant Group M Subtypes: Implications for the Epidemic. *J. Virol.* **2009**, doi:10.1128/jvi.02051-08.

28. Ndung'u, T.; Sepako, E.; McLane, M.F.; Chand, F.; Bedi, K.; Gaseitsiwe, S.; Doualla-Bell, F.; Peter, T.; Thior, I.; Moyo, S.M.; et al. HIV-1 subtype C in vitro growth and coreceptor utilization. *Virology* **2006**, doi:10.1016/j.virol.2005.11.047.

29. Venner, C.M.; Nankya, I.; Kyeyune, F.; Demers, K.; Kwok, C.; Chen, P.L.; Rwambuya, S.; Munjoma, M.; Chipato, T.; Byamugisha, J.; et al. Infecting HIV-1 Subtype Predicts Disease Progression in Women of Sub-Saharan Africa. *EBioMedicine* **2016**, doi:10.1016/j.ebiom.2016.10.014.

30. Bello, G.; Passaes, C.P.; Guimarães, M.L.; Lorete, R.S.; Matos Almeida, S.E.; Medeiros, R.M.; Alencastro, P.R.; Morgado, M.G. Origin and evolutionary history of HIV-1 subtype C in Brazil. *AIDS* **2008**, *22*, 1993–2000, doi:10.1097/QAD.0b013e328315e0aa.

31. De Oliveira, T.; Pillay, D.; Gifford, R.J. The HIV-1 subtype C epidemic in South America is linked to the United Kingdom. *PLoS One* **2010**, *5*, doi:10.1371/journal.pone.0009311.

32. Véras, N.M.C.; Gray, R.R.; Brígido, de L.F.M.; Rodrigues, R.; Salemi, M. High-resolution phylogenetics and phylogeography of human immunodeficiency virus type 1 subtype C epidemic in South America. *J. Gen. Virol.* **2011**, doi:10.1099/vir.0.028951-0.

33. Delatorre, E.; Couto-Fernandez, J.C.; Guimarães, M.L.; Vaz Cardoso, L.P.; de Alcantara, K.C.;

Martins de Araújo Stefani, M.; Romero, H.; Freire, C.C.M.; Iamarino, A.; de A Zanotto, P.M.; et al. Tracing the Origin and Northward Dissemination Dynamics of HIV-1 Subtype C in Brazil. *PLoS One* **2013**, *8*, doi:10.1371/journal.pone.0074072.

34. Raboni, S.M.; de Almeida, S.M.; Rotta, I.; Ribeiro, C.E.L.; Rosario, D.; Vidal, L.R.; Nogueira, M.B.; Riedel, M.; da Graça Winhescki, M.; Ferreira, K.A.; et al. Molecular epidemiology of HIV-1 clades in Southern Brazil. *Mem. Inst. Oswaldo Cruz* 2010, doi: 10.1590/S0074-02762010000800015

35. Wood, E.; Hogg, R.S.; Yip, B.; Harrigan, P.R.; Montaner, J.S.G. Why are baseline HIV RNA levels 100,000 copies/mL or greater associated with mortality after the initiation of antiretroviral therapy? *J. Acquir. Immune Defic. Syndr.* **2005**, *38*, 289–295, doi:10.1097/01.qai.0000149792.14334.25.

36. Shankarappa, R.; Margolick, J.B.; Gange, S.J.; Rodrigo, A.G.; Upchurch, D.; Farzadegan, H.; Gupta, P.; Rinaldo, C.R.; Learn, G.H.; He, X.; et al. Consistent Viral Evolutionary Changes Associated with the Progression of Human Immunodeficiency Virus Type 1 Infection. *J. Virol.* **1999**, *73*, 10489–10502, doi:10.1128/jvi.73.12.10489-10502.1999.

37. Ragonnet-Cronin, M.; Aris-Brosou, S.; Joanisse, I.; Merks, H.; Vallée, D.; Caminiti, K.; Rekart, M.; Krajden, M.; Cook, D.; Kim, J.; et al. Genetic diversity as a marker for timing infection in HIV-infected patients: Evaluation of a 6-month window and comparison with BED. *J. Infect. Dis.* **2012**, *206*, 756–764, doi:10.1093/infdis/jis411.

38. Kouyos, R.D.; Von Wyl, V.; Yerly, S.; Böni, J.; Rieder, P.; Joos, B.; Taffé, P.; Shah, C.; Bürgisser, P.; Klimkait, T.; et al. Ambiguous nucleotide calls from population-based sequencing of HIV-1 are a marker for viral diversity and the age of infection. *Clin. Infect. Dis.* **2011**, *52*, 532–539, doi:10.1093/cid/ciq164.

39. Andersson, E.; Shao, W.; Bontell, I.; Cham, F.; Cuong, D.D.; Wondwossen, A.; Morris, L.; Hunt, G.; Sönnerborg, A.; Bertagnolio, S.; et al. Evaluation of sequence ambiguities of the HIV-1 pol gene as a method to identify recent HIV-1 infection in transmitted drug resistance surveys. *Infect. Genet. Evol.* **2013**, *18*, 125–131, doi:10.1016/j.meegid.2013.03.050.

40. Ministério da Saúde (BR); Secretaria de Vigilância em Saúde *Boletim Epidemiológico HIV/Aids*; 2018;

41. Gräf, T.; Passaes, C.P.B.; Ferreira, L.G.E.; Grisard, E.C.; Morgado, M.G.; Bello, G.; Pinto, A.R. HIV-

1 genetic diversity and drug resistance among treatment naïve patients from Southern Brazil: An association of HIV-1 subtypes with exposure categories. *J. Clin. Virol.* **2011**, doi:10.1016/j.jcv.2011.04.011.

42. de Medeiros, R.M.; Junqueira, D.M.; Matte, M.C.C.; Barcellos, N.T.; Chies, J.A.B.; Matos Almeida, S.E. Co-circulation HIV-1 subtypes B, C, and CRF31-BC in a drug-naïve population from Southernmost Brazil: Analysis of primary resistance mutations. *J. Med. Virol.* **2011**, doi:10.1002/jmv.22188.

43. Silva, M.M.G. Da; Telles, F.Q.; da Cunha, C.A.; Rhame, F.S. HIV subtype, epidemiological and mutational correlations in patients from Paraná, Brazil. *Braz. J. Infect. Dis.* **2010**, *14*, 495–501, doi:10.1590/S1413-86702010000500012.

44. Barbieri, C.L.A.; Couto, M.T. As amas de leite e a regulamentação biomédica do aleitamento cruzado: contribuições da socioantropologia e da história. *Cad. História da Ciência - Inst. Butantan* **2012**, *VIII*, 61–76, doi: 10.47692/cadhistcienc.2012.v8.35821

45. Langland, V. Expressing Motherhood: Wet Nursing and Human Milk Banking in Brazil. *J. Hum. Lact.* **2019**, *35*, 354–361, doi:10.1177/0890334419830987.

46. Essex, M. Human Immunodeficiency Viruses in The Developing World. *Adv. Virus Res.* **1999**, doi:10.1016/S0065-3527(08)60343-7.

47. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, doi:10.1093/molbev/mst010.

48. Pineda-Peña, A.C.; Faria, N.R.; Imbrechts, S.; Libin, P.; Abecasis, A.B.; Deforche, K.; Gómez-López, A.; Camacho, R.J.; De Oliveira, T.; Vandamme, A.M. Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: Performance evaluation of the new REGA version 3 and seven other tools. *Infect. Genet. Evol.* **2013**, doi:10.1016/j.meegid.2013.04.032.

49. Struck, D.; Lawyer, G.; Ternes, A.M.; Schmit, J.C.; Perez Bercoff, D. COMET: Adaptive context-based modeling for ultrafast HIV-1 subtype identification. *Nucleic Acids Res.* **2014**, *42*, e144, doi:10.1093/nar/gku739.

50. Schultz, A.K.; Zhang, M.; Bulla, I.; Leitner, T.; Korber, B.; Morgenstern, B.; Stanke, M. jpHMM: Improving the reliability of recombination prediction in HIV-1. *Nucleic Acids Res.* **2009**, *37*, 647–651, doi:10.1093/nar/gkp371.

51. Siepel, A.C.; Halpern, A.L.; Macken, C.; Korber, B.T.M. A Computer Program Designed to Screen Rapidly for HIV Type 1 Intersubtype Recombinant Sequences. *AIDS Res. Hum. Retroviruses* **1995**, doi:10.1089/aid.1995.11.1413.

52. Pond, S.L.K.; Posada, D.; Stawiski, E.; Chappey, C.; Poon, A.F.Y.; Hughes, G.; Fearnhill, E.; Gravenor, M.B.; Brown, A.J.L.; Frost, S.D.W. An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. *PLoS Comput. Biol.* **2009**, doi:10.1371/journal.pcbi.1000581.

53. Araújo, P.M.M.; Martins, J.S.; Osório, N.S. SNAPPy: A snakemake pipeline for scalable HIV-1 subtyping by phylogenetic pairing. *Virus Evol.* **2019**, *5*, doi:10.1093/ve/vez050.

54. Karlin, S.; Altschul, S.F. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, doi:10.1073/pnas.87.6.2264.

55. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321, doi:10.1093/sysbio/syq010.

56. Lefort, V.; Longueville, J.E.; Gascuel, O. SMS: Smart Model Selection in PhyML. *Mol. Biol. Evol.* **2017**, *34*, 2422–2424, doi:10.1093/molbev/msx149.

57. Drummond, A.J.; Suchard, M.A.; Xie, D.; Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **2012**, *29*, 1969–1973, doi:10.1093/molbev/mss075.

58. Ferreira, M.A.R.; Suchard, M.A. Bayesian analysis of elapsed times in continuous-time Markov chains. *Can. J. Stat.* **2008**, doi:10.1002/cjs.5550360302.

59. Posada, D. jModelTest: Phylogenetic Model Averaging. *Mol. Biol. Evol.* **2008**, *25*, 1253–1256, doi:10.1093/molbev/msn083.

60. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **2018**, doi:10.1093/sysbio/syy032.

61. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, doi:10.1093/molbev/msy096.

62. Bielejec, F.; Baele, G.; Vrancken, B.; Suchard, M.A.; Rambaut, A.; Lemey, P. SpreaD3: Interactive

Visualization of Spatiotemporal History and Trait Evolutionary Processes. *Mol. Biol. Evol.* **2016**, doi:10.1093/molbev/msw082.

Chapter 3 | General Discussion and Conclusions

1. General discussion and conclusions

Brazil comprises the largest number of people living with HIV in Latin America, as more than one million cases of AIDS have been reported in this country. According to the Epidemiological Bulletin from the Health Surveillance Department, in 2019, at least 41.000 individuals were diagnosed with HIV infection and more than 10.000 (14.000 estimated by UNAIDS) die from AIDS-related causes [1,2]. HIV-1 subtypes present differences on distribution across the different Brazilian geographic regions, with subtype B dominating the epidemics countrywide, and variations on DRM rates have also been observed [3,4].

HIV-1 genetic diversity among subtypes and recombinant forms has been associated with differences on disease progression and dynamics, preferential transmission routes and response to treatment, which is mainly due to differences on the prevalence of DRMs [4–20]. These differences can provide a certain advantage regarding host genetics and social factors, contributing to the broad HIV-1 subtype dispersion worldwide [21–23]. Likewise, prevalence of DRMs also presents different patterns according to the sociodemographic characteristics of the population [24–27].

In the first part of our experimental work, as DRM represent one of the biggest obstacles regarding effective therapy, we decided to study the prevalence of DRMs across the 27 Brazilian federative units focusing on 20,226 HIV-1 infected patients, under different treatment schemes. Our results corroborated the high frequency of acquired DRMs reported in Brazil in previous studies, with 84.1% of the infected individuals presenting at least one DRM [28–31]. From the overall genotyped sequences, 68.5% presented at least a DRM to NRTIs, 57.95% to NNRTIs and 24.82% presented a PI resistance mutation.

Overall, most of the top commonly found mutations, being the three most common M184V, K103N, and M41L, presented a decrease on their prevalence or remained stable over the years. This decrease on acquired DRM was also observed in other countries, giving raise to the hypothesis that an improvement on treatment efficacy and an increased accessibility to ART could be the reason behind this decline on DRM presence [32–34]. In case of Brazil, it is possible that the extension of pretreatment genotyping to children and pregnant woman, in case of *Mtb* co-infection and new infections from a sexual partner under ART, might have influenced the decreasing trend [35].

Contrasting with these group of mutations, K65R, that was present on 2.23% of the genotyped sequences in 2008 and in 12.11% in 2017, became the third most common mutation of our cohort in

2016. This mutation was mainly found in association with the use of TDF, which is included in the first-line recommended ART regimen in Brazil, corroborating other studies that pointed a higher prevalence of this mutation in patients failing ART treatment containing TDF [36–40]. However, different patterns on K65R rates have been reported in several countries in which TDF is commonly used, which could indicate that this mutation frequency can be influenced by the populations' genetic or sociodemographic profiles [41–45]. In fact, we found, by relying on artificial neural network-based immunoinformatic predictions, that K65R could increase HIV recognition by HLA-B27, that presents low prevalence levels in Brazil. However, it is important to reinforce that more studies should be carried to better understand the impact of DRM on HLAs recognition, as our results are only based on in-silico predictions and should be interesting to perform functional studies to address this topic.

Moreover, evidence of K65R transmission was also observed, as 21 clusters containing at least two sequences with this mutation were found. Although K65R has been associated with diminished viral fitness and replication, which has been linked with decreased transmission capacity, other studies indicate the presence of similar viral loads in individuals with and without the tenofovir resistance, which highlights the importance of pre-treatment genotyping for this mutation, which transmission can compromise the efficacy of PrEP [46–51]. In future studies, it would be interesting to perform the same studies in a treatment naïve population, to access the levels of K65R transmission.

In line with what has been observed in this study, K65R has been associated with HIV-1 subtype C [52,53]. Subtype C is the most prevalent worldwide, accounting for more than 50% of the HIV global infections. However, it is mainly localized to Southern Africa, Ethiopia and India [54,55]. In case of Brazil, subtype B is the most prevalent subtype at a national level [56]. Studies have pointed that subtype C dominates in proportion in the Brazilian Southern area, which was corroborated by the results of the second part of the experimental work of this thesis [56]. Remarkably, according to our study, all geographic regions presented an increase on subtype C growth rate, which was dominated by the South (1.73% per year). Most interestingly, despite intense and regular movement of people between the South and South-East regions, the lowest overall subtype C proportion of cases in the studied population was found in the South-East, with the lowest subtype C proportion growth rates in the country.

Previous work indicated that subtype C is associated with higher CD4+ T cell counts in African cohorts when compared with subtypes A and D [57]. Accordingly, we observed that subtype C infected patients presented lower levels of immunodepression when compared with individuals infected with subtype B, despite reaching similar viral loads, and that this difference was not related with patients' age

or time of infection, as no significant differences were observed on the statistical analysis of the PAS. In fact, a previous study from Sub-Saharan Africa also demonstrated that women infected with HIV-1 subtype C presented a slower CD4 decline and, consequently, disease progression, when compared with subtype A and D [57]. Moreover, when compared with other variants, subtype C also presented higher CD4+ T-cell counts in a Brazilian cohort [10]. Longer asymptomatic periods associated with slower disease progression might be related with higher transmission rates which could be one of the reasons behind the increase growth of this subtype in our cohort.

We also investigated C subtype transmission by performing molecular epidemiology and phylogeographical analysis and found that the state with the highest probability for the place of entrance of subtype C in Brazil was Santa Catarina (SC). Moreover, the characterization of transmission clusters and phygeographic dynamics suggest that the inferior capacity of C subtype to thrive outside the South was not due to absence of cross-regional transmission, as we found more than 20% of the C subtype transmission events between the South and other Brazilian regions, particularly in the South-East, although small proportions of this subtype were observed in these areas. These observations corroborate the hypothesis that viral subtypes might take advantage of the sociodemographic characteristics of population to become dominant in a particular setting. Future studies, relying on a bigger data set of patients with complete information regarding, for instance, the route of transmission, should be carried to better understand the epidemiology of HIV-1 subtypes across the country.

We also compared the HIV-1 epidemics in the South and South-East region, which suggests that C subtype capacity to outcompete B might be facilitated in settings with higher female-to-male infection ratios and women-to-child transmission, although information regarding the transmission route of infection was missing in several patients. Subtype C has been previously associated with preferential in-utero transmission and increased vaginal shedding, in case of mother-to-child transmission, when compared with subtype D and A, and presented an association with heterosexual transmission in comparison with other subtypes [7,9,10]. To your knowledge, there are still no studies comparing subtype C and B presence in cells from the vaginal mucosa or even breast milk, being important to address this question in the future Moreover, cross-breastfeeding was a culturally established and accepted behavior in Brazil, and at least until the first half of the XX century, Afro-Brazilian women were frequently paid to cross-breastfeed [58]. These facts raise the possibility that sociocultural heritages could have influenced the introduction and transmission of subtype C and its distribution in Brazil.

Overall, this work reinforced the importance of HIV-1 genotyping and resistance testing prior to treatment as we found an heterogenous distribution of viral subtypes across Brazil and high levels of DRMs to drugs that are included the commonly used treatment regimens. The increasing prevalence of K65R, which is significantly more present in subtype C, that was also found to be proliferating across the country, could become a major obstacle to the standard first-line treatment scheme containing TDF and compromise PrEP efficacy in case of transmission. Our results also corroborated the differences on the transmission dynamics between B and C subtypes, study opening lines of research on the differences between the two most prevalent HIV-1 subtypes. This work also highlights the importance of defining public control strategics focusing on specific social groups that are more prone to be infected by a dominating subtype, aiming at reducing its epidemiological burden.

Bibliography

1. Secretaria de Vigilância em Saúde, M. da S. *Boletim Epidemiológico Especial HIV/Aids 2020*; Brasília/DF, 2020;

2. Joint United Nations Programme on HIV/AIDS *UNAIDS Data 2020*; 2020;

3. Librelotto, C.S.; Gräf, T.; Simon, D.; de Almeida, S.E.M.; Lunge, V.R. HIV-1 epidemiology and circulating subtypes in the countryside of South Brazil. *Rev. Soc. Bras. Med. Trop.* **2015**, *48*, 249–257, doi:10.1590/0037-8682-0083-2015.

4. Leite, T.C.N.F.; Campos, D.P.; Coelho, A.B.; Teixeira, S.L.M.; Veloso, V.; Morgado, M.G.; Guimarães, M.L. Impact of HIV-1 Subtypes on AIDS Progression in a Brazilian Cohort. *AIDS Res. Hum. Retroviruses* **2017**, *33*, 41–48, doi:10.1089/aid.2016.0126.

5. Désiré, N.; Cerutti, L.; Le Hingrat, Q.; Perrier, M.; Emler, S.; Calvez, V.; Descamps, D.; Marcelin, A.G.; Hué, S.; Visseaux, B. Characterization update of HIV-1 M subtypes diversity and proposal for subtypes A and D sub-subtypes reclassification. *Retrovirology* **2018**, *15*, 80, doi:10.1186/s12977-018-0461-y.

6. Chu, M.; Zhang, W.; Zhang, X.; Jiang, W.; Huan, X.; Meng, X.; Zhu, B.; Yang, Y.; Tao, Y.; Tian, T.; et al. HIV-1 CRF01-AE strain is associated with faster HIV/AIDS progression in Jiangsu Province, China. *Sci. Rep.* **2017**, *7*, 1–8, doi:10.1038/s41598-017-01858-2.

7. Renjifo, B.; Gilbert, P.; Chaplin, B.; Msamanga, G.; Mwakagile, D.; Fawzi, W.; Essex, M. Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. *AIDS* **2004**, *18*, 1629–1636, doi:10.1097/01.aids.0000131392.68597.34.

8. Geretti, A.M.; Harrison, L.; Green, H.; Sabin, C.; Hill, T.; Fearnhill, E.; Pillay, D.; Dunn, D. Effect of HIV-1 subtype on virologic and immunologic response to starting highly active antiretroviral therapy. *Clin. Infect. Dis.* **2009**, *48*, 1296–1305, doi:10.1086/598502.

9. John-Stewart, G.C.; Nduati, R.W.; Rousseau, C.M.; Mbori-Ngacha, D.A.; Richardson, B.A.; Rainwater, S.; Panteleeff, D.D.; Overbaugh, J. Subtype C is associated with increased vaginal shedding of HIV-1. *J. Infect. Dis.* **2005**, *192*, 492–496, doi:10.1086/431514.

10. Silveira, J.; Santos, A.F.; Martínez, A.M.B.; Góes, L.R.; Mendoza-Sassi, R.; Muniz, C.P.;

Tupinambás, U.; Soares, M.A.; Greco, D.B. Heterosexual transmission of human immunodeficiency virus type 1 subtype C in southern Brazil. *J. Clin. Virol.* **2012**, *54*, 36–41, doi:10.1016/j.jcv.2012.01.017.

11. Yuan, R.; Cheng, H.; Chen, L.S.; Zhang, X.; Wang, B. Prevalence of different HIV-1 subtypes in sexual transmission in China: A systematic review and meta-analysis. *Epidemiol. Infect.* 2016, *144*, 2144–2153, doi: 10.1017/S0950268816000212.

12. Easterbrook, P.J.; Smith, M.; Mullen, J.; O'Shea, S.; Chrystie, I.; De Ruiter, A.; Tatt, I.D.; Geretti, A.M.; Zuckerman, M. Impact of HIV-1 viral subtype on disease progression and response to antiretroviral therapy. *J. Int. AIDS Soc.* **2010**, *13*, doi:10.1186/1758-2652-13-4.

13. Kiwanuka, N.; Laeyendecker, O.; Robb, M.; Kigozi, G.; Arroyo, M.; McCutchan, F.; Eller, L.A.; Eller, M.; Makumbi, F.; Birx, D.; et al. Effect of human immunodeficiency virus type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. *J. Infect. Dis.* **2008**, *197*, 707–713, doi:10.1086/527416.

14. Taylor, B.S.; Sobieszczyk, M.E.; McCutchan, F.E.; Hammer, S.M. The Challenge of HIV-1 Subtype Diversity. *N. Engl. J. Med.* **2008**, *358*, 1590–1602, doi:10.1056/nejmra0706737.

15. Sui, H.; Gui, T.; Jia, L.; Guo, W.; Han, J.; Liu, Y.; Bao, Z.; Li, H.; Li, J.; Li, L. Different frequencies of drug resistance mutations among HIV-1 subtypes circulating in China: A comprehensive study. *PLoS One* **2014**, *9*, 91803, doi:10.1371/journal.pone.0091803.

16. Yang, C.; Li, M.; Newman, R.D.; Shi, Y.P.; Ayisi, J.; Van Eijk, A.M.; Otieno, J.; Misore, A.O.; Steketee, R.W.; Nahlen, B.L.; et al. Genetic diversity of HIV-1 in western Kenya: Subtype-specific differences in mother-to-child transmission. *AIDS* **2003**, *17*, 1667–1674, doi:10.1097/00002030-200307250-00011.

17. Araújo, P.M.M.; Carvalho, A.; Pingarilho, M.; Faria, D.; Pinho, R.; Ferreira, J.; Proença, P.; Nunes, S.; Mouro, M.; Teófilo, E.; et al. Characterization of a large cluster of HIV-1 A1 infections detected in Portugal and connected to several Western European countries. *Sci. Rep.* **2019**, *9*, 1–10, doi:10.1038/s41598-019-43420-2.

18. Riemenschneider, M.; Cashin, K.Y.; Budeus, B.; Sierra, S.; Shirvani-Dastgerdi, E.; Bayanolhagh, S.; Kaiser, R.; Gorry, P.R.; Heider, D. Genotypic Prediction of Co-receptor Tropism of HIV-1 Subtypes A and C. *Sci. Rep.* **2016**, *6*, 1–9, doi:10.1038/srep24883.

19. Vasan, A.; Renjifo, B.; Hertzmark, E.; Chaplin, B.; Msamanga, G.; Essex, M.; Fawzi, W.; Hunter, D. Different rates of disease progression of HIV type 1 infection in Tanzania based on infecting subtype. *Clin. Infect. Dis.* **2006**, *42*, 843–852, doi:10.1086/499952.

20. Amornkul, P.N.; Karita, E.; Kamali, A.; Rida, W.N.; Sanders, E.J.; Lakhi, S.; Price, M.A.; Kilembe, W.; Cormier, E.; Anzala, O.; et al. Disease progression by infecting HIV-1 subtype in a seroconverter cohort in sub-Saharan Africa. *AIDS* **2013**, *27*, 2775–2786, doi:10.1097/QAD.0000000000000012.

21. Poon, A.F.Y.; Joy, J.B.; Woods, C.K.; Shurgold, S.; Colley, G.; Brumme, C.J.; Hogg, R.S.; Montaner, J.S.G.; Harrigan, P.R. The impact of clinical, demographic and risk factors on rates of HIV transmission: A population-based phylogenetic analysis in British Columbia, Canada. *J. Infect. Dis.* **2015**, *211*, 926–935, doi:10.1093/infdis/jiu560.

22. Filho, A.W. de O.; Brites, C. Geolocalization of HIV-1 subtypes and resistance mutations of patients failing antiretroviral therapy in Salvador – Brazil. *Brazilian J. Infect. Dis.* **2017**, *21*, 234–239, doi:10.1016/j.bjid.2017.02.006.

23. Abecasis, A.B.; Wensing, A.M.J.; Paraskevis, D.; Vercauteren, J.; Theys, K.; Van de Vijver, D.A.M.C.; Albert, J.; Asjö, B.; Balotta, C.; Beshkov, D.; et al. HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. *Retrovirology* **2013**, *10*, 1–13, doi:10.1186/1742-4690-10-7.

24. Tanaka, T.S.O.; Leite, T.F.; Freitas, S.Z.; Cesar, G.A.; De Rezende, G.R.; Lindenberg, A.D.S.C.; Guimarães, M.L.; Motta-Castro, A.R.C. HIV-1 molecular epidemiology, transmission clusters and transmitted drug resistance mutations in central Brazil. *Front. Microbiol.* **2019**, *10*, doi:10.3389/fmicb.2019.00020.

25. Arruda, M.B.; Boullosa, L.T.; Cardoso, C.C.; da Costa, C.M.; Brites, C.; de Lima, S.T.S.; Kaminski, H.T.; Aleixo, A.W.; Esposito, A.O.P.; Cavalcanti, A.M.S.; et al. Brazilian network for HIV drug resistance surveillance (HIV-BresNet): A survey of treatment-naive individuals. *J. Int. AIDS Soc.* **2018**, *21*, doi:10.1002/jia2.25032.

26. Geretti, A.M. Epidemiology of antiretroviral drug resistance in drug-naïve persons. *Curr. Opin. Infect. Dis.* **2007**, *20*, 22–32, doi:10.1097/QCO.0b013e328013caff.

27. Frentz, D.; Boucher, C.A.B.; Van De Vijver, D.A.M.C. Temporal changes in the epidemiology of

transmission of drug-resistant HIV-1 across the world. *AIDS Rev.* 2012, *14*, 17–27.

28. Lopes, C.A.F.; Soares, M.A.; Falci, D.R.; Sprinz, E. The Evolving Genotypic Profile of HIV-1 Mutations Related to Antiretroviral Treatment in the North Region of Brazil. *Biomed Res. Int.* **2015**, *2015*, doi:10.1155/2015/738528.

29. Brites, C.; Pinto-Neto, L.; Medeiros, M.; Nunes, E.; Sprinz, E.; Carvalho, M. Extensive variation in drug-resistance mutational profile of Brazilian patients failing antiretroviral therapy in five large Brazilian cities. *Brazilian J. Infect. Dis.* **2016**, *20*, 323–329, doi:10.1016/j.bjid.2016.03.010.

30. Duani, H.; Aleixo, A.W.; Tupinambás, U. Trends and predictors of HIV-1 acquired drug resistance in Minas Gerais, Brazil: 2002–2012. *Brazilian J. Infect. Dis.* **2017**, *21*, 148–154, doi:10.1016/j.bjid.2016.11.009.

31. Toledo, P.V.M.; de Carvalho, D.S.; Romagnoli, L.; Marcinko, G.; da Cunha, C.A.; de Souza, M.N.; Brindeiro, R.; de Queiroz-Telles, F. HIV-1 genotypic resistance profile of patients failing antiretroviral therapy in Paraná, Brazil. *Brazilian J. Infect. Dis.* **2010**, *14*, 360–371, doi:10.1016/s1413-8670(10)70076-3.

32. Pingarilho, M.; Pimentel, V.; Diogo, I.; Fernandes, S.; Miranda, M.; Pineda-Pena, A.; Libin, P.; Theys, K.; O. Martins, M.; Vandamme, A.-M.; et al. Increasing Prevalence of HIV-1 Transmitted Drug Resistance in Portugal: Implications for First Line Treatment Recommendations. *Viruses* **2020**, *12*, 1238, doi:10.3390/v12111238.

33. Rocheleau, G.; Brumme, C.J.; Shoveller, J.; Lima, V.D.; Harrigan, P.R. Longitudinal trends of HIV drug resistance in a large Canadian cohort, 1996–2016. *Clin. Microbiol. Infect.* **2018**, *24*, 185–191, doi:10.1016/j.cmi.2017.06.014.

34. Franzetti, M.; Violin, M.; Antinori, A.; De Luca, A.; Ceccherini-Silberstein, F.; Gianotti, N.; Torti, C.; Bonora, S.; Zazzi, M.; Balotta, C. Trends and correlates of HIV-1 resistance among subjects failing an antiretroviral treatment over the 2003-2012 decade in Italy. *BMC Infect. Dis.* **2014**, *14*, 398, doi:10.1186/1471-2334-14-398.

35. Ministério da Saúde PROTOCOLO CLÍNICO E DIRETRIZES TERAPÊTUICAS PARA O MANEJO DA INFECÇÃO PELO HIV. *Ministério da Saúde. Secr. Vigilância em Saúde. Dep. DST, Aids e Hepatites Virais.* **2013**.

36. Ministério da Saúde *Clinical Protocol and Therapeutic Guidelines for Handling HIV infection in adults.*; 1st ed.; Brasilia/DF, 2018; ISBN 9788533426436.

37. Brenner, B.G.; Coutsinos, D. The K65R mutation in HIV-1 reverse transcriptase: Genetic barriers, resistance profile and clinical implications. *HIV Ther.* **2009**, *3*, 583–594, doi:10.2217/hiv.09.40.

38. Chaplin, B.; Eisen, G.; Idoko, J.; Onwujekwe, D.; Idigbe, E.; Adewole, I.; Gashau, W.; Meloni, S.; Sarr, A.D.; Sankalé, J.L.; et al. Impact of HIV type 1 subtype on drug resistance mutations in Nigerian patients failing first-line therapy. *AIDS Res. Hum. Retroviruses* **2011**, *27*, 71–80, doi:10.1089/aid.2010.0050.

39. Ayitewala, A.; Kyeyune, F.; Ainembabazi, P.; Nabulime, E.; Kato, C.D.; Nankya, I. Comparison of HIV drug resistance profiles across HIV-1 subtypes A and D for patients receiving a tenofovir-based and zidovudine-based first line regimens in Uganda. *AIDS Res. Ther.* **2020**, *17*, doi:10.1186/s12981-020-0258-7.

40. Luo, X. lu; Mo, L. da; Su, G. sheng; Huang, J. ping; Wu, J. yu; Su, H. zhen; Huang, W. hong; Luo, S. da; Ni, Z. yan Incidence and types of HIV-1 drug resistance mutation among patients failing first-line antiretroviral therapy. *J. Pharmacol. Sci.* **2019**, *139*, 275–279, doi:10.1016/j.jphs.2018.11.016.

41. Karade, S.; Chaturbhuj, D.N.; Sen, S.; Joshi, R.K.; Kulkarni, S.S.; Shankar, S.; Gangakhedkar, R.R. HIV drug resistance following a decade of the free antiretroviral therapy programme in India: A review. *Int. J. Infect. Dis.* **2018**, *66*, 33–41, doi:10.1016/j.ijid.2017.10.020.

42. Charpentier, C.; Lambert-Niclot, S.; Visseaux, B.; Morand-Joubert, L.; Storto, A.; Larrouy, L.; Landman, R.; Calvez, V.; Marcelin, A.-G.; Descamps, D. Evolution of the K65R, K103N and M184V/I reverse transcriptase mutations in HIV-1-infected patients experiencing virological failure between 2005 and 2010. *J. Antimicrob. Chemother.* **2013**, *68*, 2197–2198, doi:10.1093/jac/dkt184.

43. Theys, K.; Snoeck, J.; Vercauteren, J.; Abecasis, A.B.; Vandamme, A.M.; Camacho, R.J.; Mansinho, K.; Miranda, A.C.; Aldir, I.; Ventura, F.; et al. Decreasing population selection rates of resistance mutation K65R over time in HIV-1 patients receiving combination therapy including tenofovir. *J. Antimicrob. Chemother.* **2013**, *68*, 419–423, doi:10.1093/jac/dks380.

44. Dolling, D.; Sabin, C.; Delpech, V.; Smit, E.; Pozniak, A.; Asboe, D.; Brown, A.L.; Churchill, D.;

Williams, I.; Geretti, A.M.; et al. Time trends in drug resistant HIV-1 infections in the United Kingdom up to 2009: Multicentre observational study. *BMJ* **2012**, *345*, doi:10.1136/bmj.e5253.

45. Reinheimer, C.; Wesner, A.; Keppler, O.T.; Doerr, H.W.; Herrmann, E.; Stürmer, M.; Stephan, C. Prevalence of K65R in patients treated with tenofovir disoproxil fumarate: recommendations based on the Frankfurt HIV Cohort Study Resistance Database (FHCS-RD). *Med. Microbiol. Immunol.* **2016**, *205*, 315–320, doi:10.1007/s00430-015-0448-4.

46. Chan, P.A.; Huang, A.; Kantor, R. Low prevalence of transmitted K65R and other tenofovir resistance mutations across different HIV-1 subtypes: Implications for pre-exposure prophylaxis. *J. Int. AIDS Soc.* **2012**, *15*, doi:10.7448/IAS.15.2.17701.

47. Margot, N.A.; Wong, P.; Kulkarni, R.; White, K.; Porter, D.; Abram, M.E.; Callebaut, C.; Miller, M.D. Commonly transmitted HIV-1 drug resistance mutations in reverse-transcriptase and protease in antiretroviral treatment-naive patients and response to regimens containing tenofovir disoproxil fumarate or tenofovir alafenamide. *J. Infect. Dis.* **2017**, *215*, 920–927, doi:10.1093/infdis/jix015.

48. Rhee, S.Y.; Jordan, M.R.; Raizes, E.; Chua, A.; Parkin, N.; Kantor, R.; Van Zy, G.U.; Mukui, I.; Hosseinipour, M.C.; Frenkel, L.M.; et al. HIV-1 drug resistance mutations: Potential applications for point-of-care Genotypic resistance testing. *PLoS One* **2015**, *10*, 12, doi:10.1371/journal.pone.0145772.

49. Wagner, B.G.; Garcia-Lerma, J.G.; Blower, S. Factors limiting the transmission of HIV mutations conferring drug resistance: Fitness costs and genetic bottlenecks. *Sci. Rep.* **2012**, *2*, 1–10, doi:10.1038/srep00320.

50. Weber, J.; Chakraborty, B.; Weberova, J.; Miller, M.D.; Quiñones-Mateu, M.E. Diminished replicative fitness of primary human immunodeficiency virus type 1 isolates harboring the K65R mutation. *J. Clin. Microbiol.* **2005**, *43*, 1395–1400, doi:10.1128/JCM.43.3.1395-1400.2005.

51. Gregson, J.; Tang, M.; Ndembi, N.; Hamers, R.L.; Marconi, V.C.; Brooks, K.; Theys, K.; Arruda, M.; Garcia, F.; Monge, S.; et al. Global epidemiology of drug resistance after failure of WHO recommended first-line regimens for adult HIV-1 infection: A multicentre retrospective cohort study. *Lancet Infect. Dis.* **2016**, *16*, 565–575, doi:10.1016/S1473-3099(15)00536-8.

52. Smit, E.; White, E.; Clark, D.; Churchill, D.; Zhang, H.; Collins, S.; Pillay, D.; Sabin, C.; Nelson, M.;

Winston, A.; et al. An association between K65R and HIV-1 subtype C viruses in patients treated with multiple NRTIs. *J. Antimicrob. Chemother.* **2017**, *72*, 2075–2082, doi:10.1093/jac/dkx091.

53. Theys, K.; Vercauteren, J.; Snoeck, J.; Zazzi, M.; Camacho, R.J.; Torti, C.; Schülter, E.; Clotet, B.; Sönnerborg, A.; De Luca, A.; et al. HIV-1 subtype is an independent predictor of reverse transcriptase mutation k65r in HIV-1 patients treated with combination antiretroviral therapy including tenofovir. *Antimicrob. Agents Chemother.* **2013**, *57*, 1053–1056, doi:10.1128/AAC.01668-12.

54. Hemelaar, J.; Elangovan, R.; Yun, J.; Dickson-Tetteh, L.; Fleminger, I.; Kirtley, S.; Williams, B.; Gouws-Williams, E.; Ghys, P.D.; Abimiku, A.G.; et al. Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *Lancet Infect. Dis.* **2019**, *19*, 143–155, doi:10.1016/S1473-3099(18)30647-9.

55. Santos-Pereira, A.; Magalhães, C.; Araújo, P.M.M.; Osório, N.S. Evolutionary Genetics of Mycobacterium Tuberculosis and HIV-1: "The Tortoise and the Hare." *Microorganisms* **2021**, *9*, 147, doi:10.3390/microorganisms9010147.

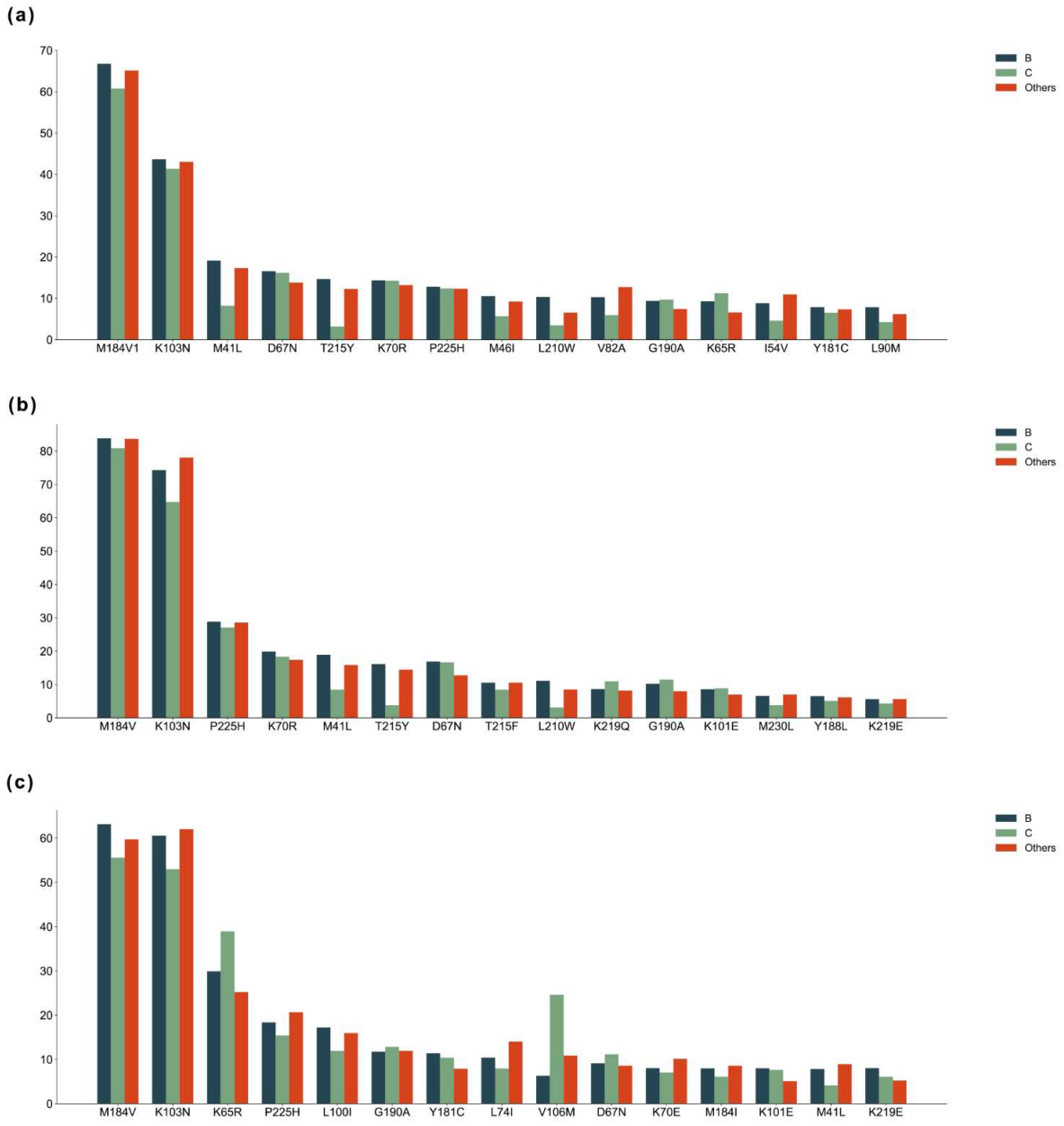
56. Gräf, T.; Bello, G.; Andrade, P.; Arantes, I.; Pereira, J.M.; da Silva, A.B.P.; Veiga, R. V.; Mariani, D.; Boullosa, L.T.; Arruda, M.B.; et al. HIV-1 molecular diversity in Brazil unveiled by 10 years of sampling by the national genotyping network. *Sci. Reports 2021 111* **2021**, *11*, 1–9, doi:10.1038/s41598-021-94542-5.

57. Venner, C.M.; Nankya, I.; Kyeyune, F.; Demers, K.; Kwok, C.; Chen, P.L.; Rwambuya, S.; Munjoma, M.; Chipato, T.; Byamugisha, J.; et al. Infecting HIV-1 Subtype Predicts Disease Progression in Women of Sub-Saharan Africa. *EBioMedicine* **2016**, *13*, 305–314, doi:10.1016/j.ebiom.2016.10.014.

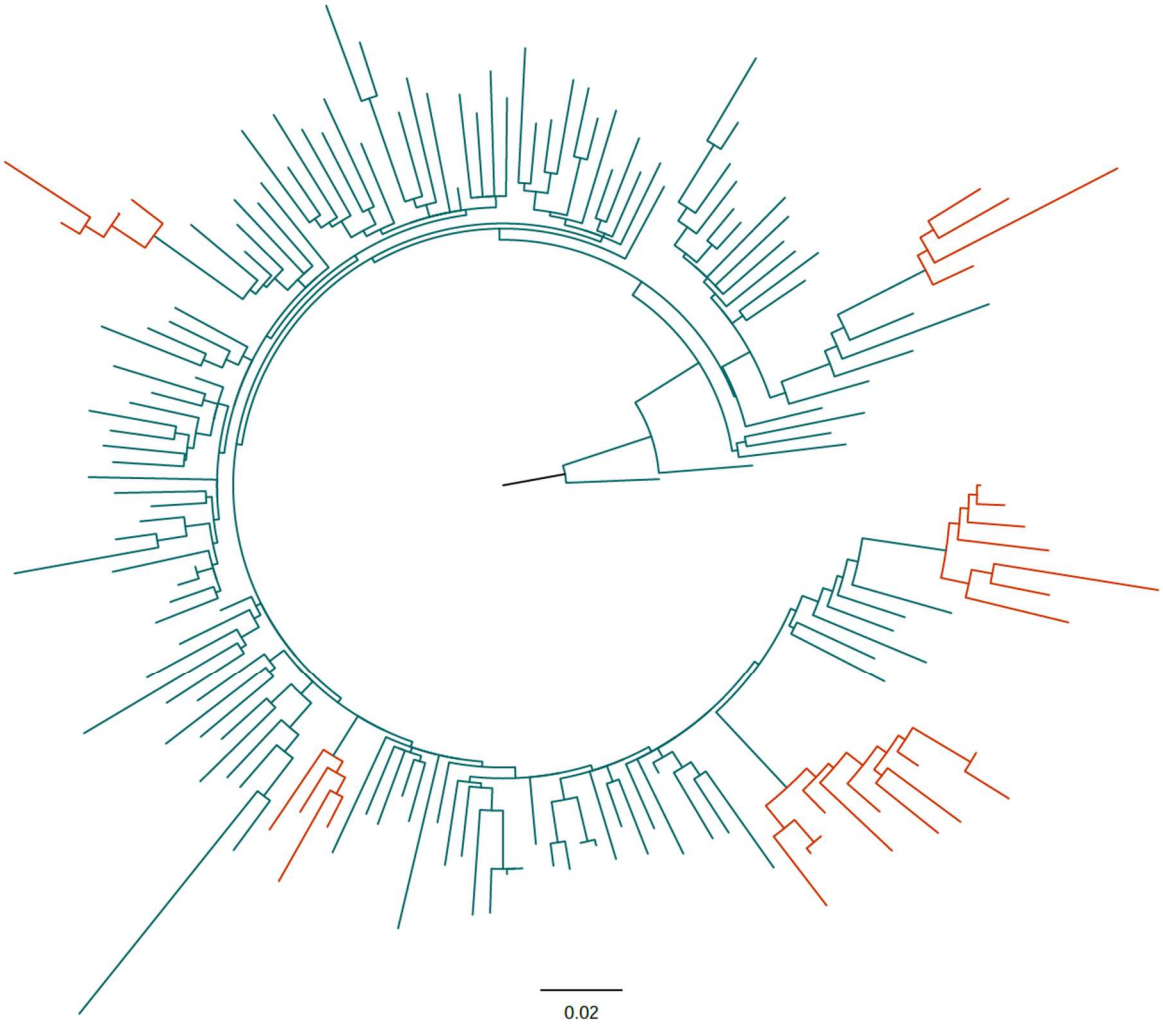
58. Langland, V. Expressing Motherhood: Wet Nursing and Human Milk Banking in Brazil. *J. Hum. Lact.* **2019**, *35*, 354–361, doi:10.1177/0890334419830987.

Supplementary Data

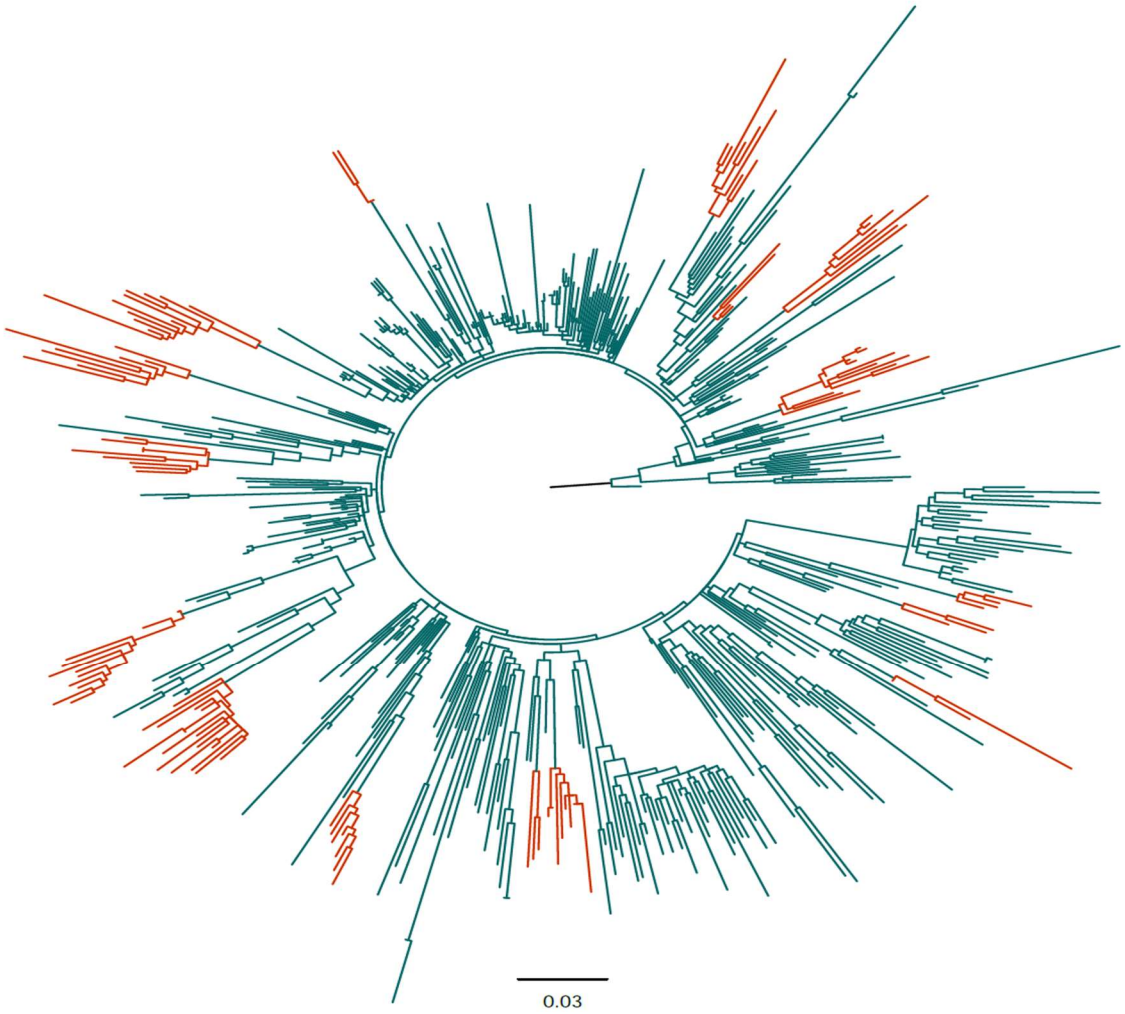
Supplementary Data



Sup Figure 1. Presence of DMR among subtypes. Top 15 most prevalent SMDR among different HIV-1 subtypes was evaluated in all the individuals of the cohort (a), in patients receiving a treatment scheme including AZT (b) or TDF (c).



Sup Figure 2. Phylogenetic representation of a subset of HIV-1 subtype C sequences isolated in Brazil with the K65R mutation and closely related sequences from databases. Circular cladogram representation of the Maximum likelihood tree. Branch colored in red indicate inferred transmission clusters. The tree was rooted considering references of subtype A1. To improve tree visualization a separate analysis excluding the outgroup sequences is shown.



Sup Figure 3. Phylogenetic representation of a subset of HIV-1 subtype B sequences isolated in Brazil with the K65R mutation and closely related sequences from databases. Circular cladogram representation of the Maximum likelihood tree. Branch colored in red indicate inferred transmission clusters. The tree was rooted considering references of subtype A1. To improve tree visualization a separate analysis excluding the outgroup sequences is shown.

Sup Table 1. Phylogeny inferred HIV-1 transmission clusters with more than one sequences harboring the K65R mutation.

Sequence	Origin	Cluster	Sample Year	Country	Federative Unit*	Transmission route	Gender	Presence of K65R mutation
18SC110074	This study	B1	2011	Brazil	SC	Homosexual	Male	No
21PR130051	This study	B1	2013	Brazil	PR	Bisexual	Male	No
21PR140146	This study	B1	2014	Brazil	PR		Male	Yes
21PR140647	This study	B1	2014	Brazil	PR		Male	Yes
26PR150111	This study	B1	2015	Brazil	PR	Homosexual	Male	No
29PR1604451	This study	B1	2016	Brazil	PR	Homosexual	Male	No
29PR1610702	This study	B1	2016	Brazil	PR		Male	No
24PI150732	This study	B2	2015	Brazil	PI		Male	Yes
29PI1700952	This study	B2	2017	Brazil	PI		Male	Yes
KF782694	Public Databases	B2	2012	Brazil	MA	Homosexual	Male	No
29SP1606275	This study	B3	2016	Brazil	SP	Heterosexual, Drugs	Male	No
29SP1608861	This study	B3	2016	Brazil	SP	Homosexual	Male	Yes
29SP1610014	This study	B3	2016	Brazil	SP	Heterosexual	Female	Yes
11SP151909	This study	B4	2015	Brazil	SP		Male	Yes
29ES1700443	This study	B4	2017	Brazil	ES	Bisexual	Male	No
29MG1601542	This study	B4	2016	Brazil	MG		Male	Yes
29DF1701982	This study	B5	2017	Brazil	DF	Homosexual	Male	Yes
29PA1605620	This study	B5	2016	Brazil	PA		Male	Yes
29PA1608064	This study	B5	2016	Brazil	PA		Male	Yes
29PA1608071	This study	B5	2016	Brazil	PA	Heterosexual	Male	Yes
29PA1702877	This study	B5	2017	Brazil	PA		Female	Yes
KX887674	Public Databases	B5	2015	Brazil	PA			No
KX887675	Public Databases	B5	2015	Brazil	PA			No
11SP140792	This study	B6	2014	Brazil	SP		Female	Yes
29RJ1608399	This study	B6	2016	Brazil	RJ		Male	No
29SP1609861	This study	B6	2016	Brazil	SP	Heterosexual	Male	Yes
29SP1610943	This study	B6	2016	Brazil	SP	Homosexual	Male	Yes
29SP1701160	This study	B6	2017	Brazil	AL		Male	No
JN100838	Public Databases	B6	2003	United Kingdom		Homosexual	Male	No
JN195956	Public Databases	B6	2008	Brazil	SP			No
KT427797	Public Databases	B6	2010	Brazil	SP			No
04DF150339	This study	B7	2015	Brazil	PB		Male	No
29RS1600171	This study	B7	2016	Brazil	RS		Male	Yes
EU248364	Public Databases	B7	2003	Belgium				No
FJ228082	Public Databases	B7	2006	Italy				No
FJ228130	Public Databases	B7	2004	Italy				No
GQ399355	Public Databases	B7	2005	Portugal				No

GU969527	Public Databases	B7	2008	Italy			Male	No
JF487842	Public Databases	B7	2008	Brazil	RS			No
JX299665	Public Databases	B7	2007	Germany				No
MN972080	Public Databases	B7	2015	Brazil	MA			Yes
MT570426	Public Databases	B7	2015	United Kingdom				No
MT787777	Public Databases	B7	2008	Italy				No
MT787784	Public Databases	B7	2007	Italy				No
09CE140280	This study	B8	2014	Brazil	CE	Heterosexual	Female	Yes
21PR120215	This study	B8	2012	Brazil	PR		Male	Yes
EU340735	Public Databases	B8	2006	Brazil	PR			No
FJ591477	Public Databases	B8	2006	Brazil	PR			No
11PB150082	This study	B9	2015	Brazil	PB	Homosexual	Male	Yes
29MG1602837	This study	B9	2016	Brazil	MG		Male	No
29PE1601370	This study	B9	2016	Brazil	PE		Male	No
29PE1602220	This study	B9	2016	Brazil	PE	Homosexual	Male	No
29PE1605080	This study	B9	2016	Brazil	PE		Male	Yes
29PE1609931	This study	B9	2016	Brazil	PE	Heterosexual	Female	Yes
29PE1610633	This study	B9	2016	Brazil	PE		Male	Yes
29PE1610635	This study	B9	2016	Brazil	PE	Homosexual	Male	No
29PE1610642	This study	B9	2016	Brazil	PE	Bisexual	Male	No
29RJ1606703	This study	B9	2016	Brazil	RJ		Male	No
KJ849784	Public Databases	B9	2010	Brazil	PE			No
06MG140069	This study	B10	2014	Brazil	MG		Male	No
09CE140292	This study	B10	2014	Brazil	CE	Homosexual	Male	Yes
11SP140667	This study	B10	2014	Brazil	SP		Male	Yes
11SP140670	This study	B10	2014	Brazil	SP	Homosexual	Male	No
29CE1605374	This study	B10	2016	Brazil	CE		Male	Yes
29DF1605721	This study	B10	2016	Brazil	DF	Heterosexual	Male	No
29RJ1606078	This study	B10	2016	Brazil	RJ		Male	Yes
29SP1610234	This study	B10	2016	Brazil	SP	Homosexual	Male	No
29SP1702293	This study	B10	2017	Brazil	CE	Homosexual	Male	No
KP115562	Public Databases	B10	2014	Brazil	SP			No
KX661656	Public Databases	B10	2011	United Kingdom				No
KX888117	Public Databases	B10	2015	Brazil	GO			No
KX888368	Public Databases	B10	2015	Brazil	MG			No
KX888759	Public Databases	B10	2015	Brazil	SP			No
KX888774	Public Databases	B10	2015	Brazil	RJ			No
KX889064	Public Databases	B10	2015	Brazil	SC			No
06BA150083	This study	B11	2014	Brazil	BA		Male	No
06BA150203	This study	B11	2015	Brazil	SE	Homosexual	Male	Yes

06BA150391	This study	B11	2015	Brazil	BA		Male	No
29BA1603668	This study	B11	2016	Brazil	BA	Heterosexual	Female	No
29BA1605576	This study	B11	2016	Brazil	BA	Bisexual	Male	Yes
29BA1605660	This study	B11	2016	Brazil	BA		Male	No
29BA1701551	This study	B11	2017	Brazil	BA		Male	No
29SP1702999	This study	B11	2017	Brazil	BA	Heterosexual	Male	Yes
22MS140168	This study	B12	2014	Brazil	MS		Male	Yes
29MS1604163	This study	B12	2016	Brazil	MS		Male	Yes
29PR1610894	This study	B12	2016	Brazil	PR		Male	Yes
JF342308	Public Databases	B12	2008	Brazil	MS			No
JF342321	Public Databases	B12	2010	Brazil	MS			No
JN196004	Public Databases	B12	2010	Brazil	SP			No
KJ849805	Public Databases	B12	2010	Brazil	SP			No
MN528456	Public Databases	B12	2017	Brazil	AM			No
MN528463	Public Databases	B12	2017	Brazil	AM			No
10SP151332	This study	B13	2015	Brazil	SP		Male	Yes
29SP1603165	This study	B13	2016	Brazil	SP	Bisexual	Male	Yes
29SP1700487	This study	B13	2017	Brazil	SP		Male	No
06MG130344	This study	B14	2013	Brazil	RJ		Male	No
09RJ130607	This study	B14	2013	Brazil	RJ		Male	Yes
11SP138750	This study	B14	2013	Brazil	SP		Male	No
11SP151590	This study	B14	2015	Brazil	SP	Homosexual	Male	No
29MG1611880	This study	B14	2016	Brazil	MG		Male	Yes
29RJ1610193	This study	B14	2016	Brazil	RJ	Homosexual	Male	No
29SP1601805	This study	B14	2016	Brazil	SP		Male	No
KX662590	Public Databases	B14	2010	United Kingdom				No
LC162092	Public Databases	B14	2013	Japan				No
LC162094	Public Databases	B14	2013	Japan				No
MN163692	Public Databases	B14	2014	Croatia				No
04DF130122	This study	B15	2013	Brazil	DF		Male	Yes
29DF1600435	This study	B15	2016	Brazil	DF		Male	No
29DF1701062	This study	B15	2017	Brazil	DF	Homosexual	Male	Yes
29RS1703532	This study	B15	2017	Brazil	RS		Male	No
KX888083	Public Databases	B15	2015	Brazil	DF			No
KX888084	Public Databases	B15	2015	Brazil	DF			No
KX888087	Public Databases	B15	2015	Brazil	DF			No
LC526723	Public Databases	B15	2017	Japan				No
MT570579	Public Databases	B15	2016	United Kingdom				No
12SP091875	This study	B16	2009	Brazil	SP		Male	Yes
23SP110075	This study	B16	2011	Brazil	SP	Homosexual	Male	Yes

HM534048	Public Databases	B16	2009	Brazil	SP			No
HM534096	Public Databases	B16	2009	Brazil	SP			No
11SP139055	This study	C1	2013	Brazil	SP		Male	Yes
21PR110049	This study	C1	2011	Brazil	PR	Homosexual	Male	No
21PR130023	This study	C1	2013	Brazil	PR	Bisexual	Male	No
29PR150058	This study	C1	2015	Brazil	PR		Male	Yes
15SP131030	This study	C2	2013	Brazil	SP	Bisexual	Male	Yes
29TO1610854	This study	C2	2016	Brazil	TO	Homosexual	Male	Yes
FJ784207	Public Databases	C2	2003	Brazil	GO		Female	No
KT744431	Public Databases	C2	2009					No
29TO1611348	This study	C2	2016	Brazil	TO		Male	No
29MG1609380	This study	C2	2016	Brazil	MG	Heterosexual	Male	No
JN114145	Public Databases	C2	2009	Brazil	GO		Female	No
07SC140751	This study	C3	2014	Brazil	SC	Bisexual	Male	No
21PR130350	This study	C3	2013	Brazil	PR	Heterosexual	Male	Yes
29SP1701779	This study	C3	2017	Brazil	SP		Male	Yes
EU340722	Public Databases	C3	2005	Brazil	PR			No
29SP150078	This study	C4	2015	Brazil	SP		Male	No
29SP1701889	This study	C4	2017	Brazil	SP		Male	Yes
29SP1702463	This study	C4	2017	Brazil	SP	Heterosexual	Male	Yes
KX888620	Public Databases	C4	2015	Brazil	SP			No
26PR150119	This study	C5	2015	Brazil	PR		Male	Yes
29MS1701288	This study	C5	2017	Brazil	MS	Heterosexual	Male	No
29SP1611091	This study	C5	2016	Brazil	SP	Bisexual	Male	No
29SP1700076	This study	C5	2017	Brazil	SP	Heterosexual	Male	Yes
29SP1700481	This study	C5	2017	Brazil	SP		Male	No
29SP1702372	This study	C5	2017	Brazil	SP		Female	Yes
JN196005	Public Databases	C5	2010	Brazil	SP			No
JN196007	Public Databases	C5	2010	Brazil	SP			No
KT427806	Public Databases	C5	2010	Brazil	SP			No
KY640158	Public Databases	C5	2015	Brazil	SP			Yes
MT787796	Public Databases	C5	2008	Italy				No

***** SC – Santa Catarina, PR – Paraná, PI – Piauí, MA – Maranhão, SP – São Paulo, ES – Espírito Santo, MG – Minas Gerais, DF – Distrito Federal, PA – Pará, RJ – Rio de Janeiro, AL – Alagoas, PB – Paraíba, RS – Rio Grande do Sul, CE – Ceará, PE – Pernambuco, GO – Goiás, BA – Bahia, SE – Sergipe, MS – Mato Grosso do Sul, AM – Amazonas, TO – Tocantins.

Sup Table 2. Distribution by subtype and region of the studied 2611 samples.

Region	Subtype			Total
	B	C	Other	
Central-West	117 70.06%	13 7.78%	37 22.16%	167 6.40%
North-East	364 74.74%	18 3.70%	105 21.56%	487 18.65%
North	111 76.55%	7 4.83%	27 18.62%	145 5.55%
South-East	946 72.49%	47 3.60%	312 23.91%	1305 49.98%
South	137 27.02%	255 50.30%	115 22.68%	507 19.42%
Total	1675 64.15%	340 13.02%	596 22.83%	2611

Sup Table 3. Annual growth rate of HIV-1 infections in the Brazilian cohort under study by subtype and region.

Subtype	Region	Annual growth rate (%)
C	South-East	0.67
	South	1.73
	North-East	0.66
	North	1.29
	Central-West	1.44
B	South-East	-1.04
	South	-1.23
	North-East	-1.54
	North	0.03
	Central-West	-1.35

Sup Table 4. Viral load comparison between individuals infected with subtypes C vs. B or C vs.other non-B and non-C subtypes.

Viral load (copies/mL)	Subtype			N	C vs. B			C vs. other		
	B	C	Other		OR	CI	p value	OR	CI	p value
<=100,000	1270 65.91%	251 13.03%	406 21.07%	1927	0.93	0.70-1.25	0.6319	1.26	0.91-1.76	0.0162
>100,000	325 60.75%	69 12.90%	141 26.36%	535	1.07	0.80-1.43		0.79	0.57-1.10	
Missing	80 53.69%	20 13.42%	49 32.89%	149	1.24	0.74-2.04	0.3919	0.70	0.40-1.18	0.1880

Sup Table 5. Cut-off age points63 used for the classification of immunological status.

Immunological status	Age group		
	<1	1-5	>5
	CD4+ counts (cell/mm3)		
Without immunodepression	>1500	>1000	>500
Moderate immunodepression	750-1499	500-999	200-499
Severe immunodepression	<750	<500	<200

Sup Table 6. Comparison of the number of infections with subtype C, B and other classified as leading to immunodepression with moderate or severe levels.

Age group (years old)	Immunodepression	Subtype			N*	C vs. B**			C vs. Other**		
		B	C	Other		OR	CI	p value	OR	CI	p value
All	No	562 60.30%	150 16.09%	220 23.61%	932	1.71	1.33-2.19	0.000	1.36	1.02-1.81	0.034
	Yes	979 68.13%	153 10.65%	305 21.22%	1437	0.58	0.46-0.75		0.74	0.55-0.98	
< 18	No	178 51.30%	66 19.02%	103 29.68%	347	1.68	1.05-2.70	0.029	1.02	0.60-1.74	0.948
	Yes	154 63.64%	34 14.05%	54 22.31%	242	0.60	0.37-0.95		0.98	0.57-1.67	
≥ 18	No	384 65.64%	84 14.36%	117 20.00%	585	1.52	1.11-2.05	0.007	1.51	1.06-2.16	0.022
	Yes	825 69.04%	119 9.96%	251 20.00%	1195	0.66	0.49-0.90		0.66	0.46-0.94	
All	Yes. Moderate	320 62.99%	71 13.98%	117 23.03%	508	1.78	1.26-2.52	0.001	1.40	0.94-2.06	0.099
	Yes. Severe	659 70.94%	82 8.83%	188 20.24%	929	0.56	0.40-0.79		0.72	0.48-1.07	
< 18	Yes. Moderate	61 56.48%	22 20.37%	25 23.15%	108	2.78	1.29-6.20	0.008	2.11	0.87-5.24	0.094
	Yes. Severe	93	12	29	134	0.36	0.16-0.78		0.47	0.19-1.15	

		69.40%	8.96%	21.64%							
≥ 18	Yes. Moderate	259 64.75%	49 12.25%	92 23.00%	400	1.53	1.03- 2.26	0.033	1.21	0.77- 1.89	0.403
	Yes. severe	566 71.19%	70 8.81%	159 20.00%	795	0.65	0.44- 0.97		0.83	0.53- 1.30	

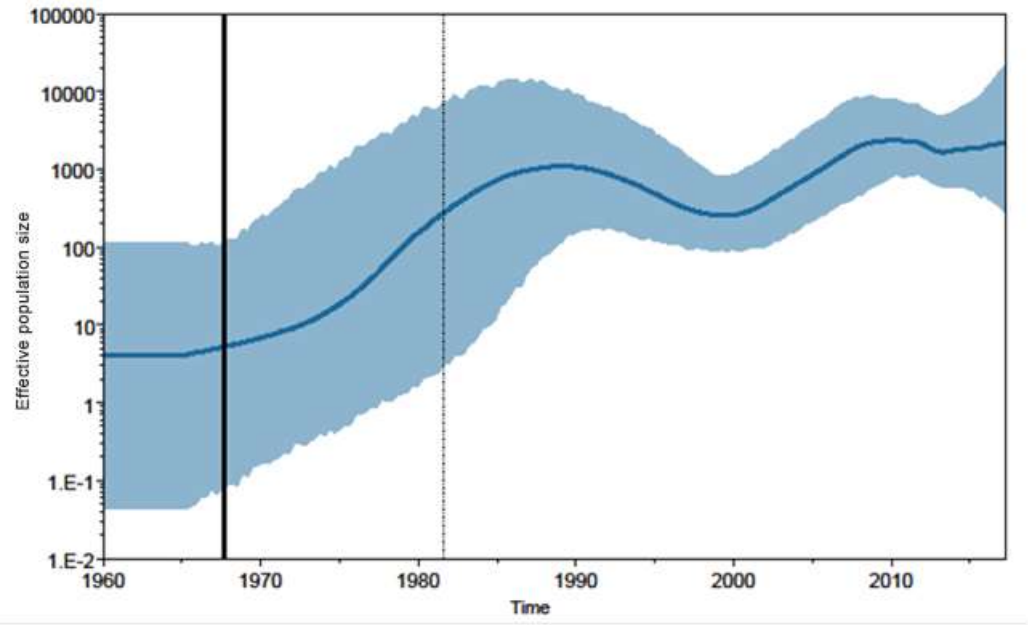
* From the 2611 cases selected for this study 245 lacked information on age, or CD4+ counts and were not included in this calculation. The missing values had a random distribution between the groups.
**Statistical comparisons were performed with corrected Mantel-Haenszel chi-square test.

Sup Table 7. Comparison of proportion of ambiguous sites (PAS) between B and C subtypes.

		B		**C**		**C vs. B**		
		n	%	n	%	OR	IC	P
Proportion of ambiguous sites (PAS) (percentile)	<25	399	23.82	87	25.59	1.10	0.84-1.43	0.487
	25 a <50	422	25.19	76	22.35	0.85	0.64-1.12	0.268
	50 a <75	395	23.58	81	23.82	1.01	0.77-1.33	0.924
	75-100	459	27.40	96	28.24	1.04	0.80-1.35	0.754

Sup Table 8. Pairwise rates of diffusion between geographic locations (discrete traits).

Location trait pair		Regions involved	BAYES FACTOR	POSTERIOR PROBABILITY
SC	RS	South	123895.7829	1
SC	SP	South. South-East	123895.7829	1
SC	PR	South	6513.5858	0.9988
SP	MS	South-East. Central-West	1129.0813	0.9933
SP	DF	South-East. Central-West	251.5648	0.9705
SC	MG	South. South-East	237.2210	0.9688
SP	PA	South-East. North	44.8986	0.8545
RJ	BA	South-East. North-East	42.9874	0.8490
SP	RO	South-East. Central-West	30.6890	0.8005
SC	RJ	South. Central-West	27.0303	0.7795
ES	US	International	16.1299	0.6784
RS	GO	South. Central-West	12.9997	0.6296
RS	DE	South. International	10.8927	0.5875



Sup Figure 4. Skygrid plot inferred with BEAST v1.10.4 and Tracer v1.7.1 using the subset of 174 sequences that make up the transmission clusters. Effective population (y-axis) size by chronological time (x-axis). The mean and the 95% Bayesian highest posterior density (HPD) interval is represented by the solid line and shaded blue area, respectively. The solid black vertical line represents the estimated root of the tree and the second line the highest age value for the root.