Characterization of HIV-1 Near Full-Length Proviral Genome Quasispecies from Patients with Undetectable Viral Load Undergoing First-Line HAART Therapy

Brunna M. Alves 1,†, Juliana D. Siqueira 1,†, Marianne M. Garrido 2, Ornella M. Botelho 1, Isabel M. Prellwitz 1, Sayonara R. Ribeiro 2, Esmeralda A. Soares 1 and Marcelo A. Soares 1,3,*

1 Programa de Oncovirologia, Instituto Nacional de Câncer, Rio de Janeiro 20231-050, Brazil; brunna_alves@yahoo.com.br (B.M.A); siduju@hotmail.com (J.D.S); ornellamartins@hotmail.com (O.M.B); imprellwitz@gmail.com (I.M.P); esoares@inca.gov.br (E.A.S)
2 Serviço de Doenças Infecciosas, Hospital Federal de Ipanema, Rio de Janeiro 22411-020, Brazil; marianne.monteiro@inca.gov.br (M.M.G); sayorib@gmail.com (S.R.R.)
3 Departamento de Genética, Universidade Federal do Rio de Janeiro, Rio de Janeiro 21944-970, Brazil
* Correspondence: masoares@inca.gov.br; Tel.: +55-21-3207-6591
† These authors contributed equally to this work.

Received: 1 December 2017; Accepted: 18 December 2017; Published: 19 December 2017

Abstract: Increased access to highly active antiretroviral therapy (HAART) by human immunodeficiency virus positive (HIV+) individuals has become a reality worldwide. In Brazil, HAART currently reaches over half of HIV-infected subjects. In the context of a remarkable HIV-1 genetic variability, highly related variants, called quasispecies, are generated. HIV quasispecies generated during infection can influence virus persistence and pathogenicity, representing a challenge to treatment. However, the clinical relevance of minority quasispecies is still uncertain. In this study, we have determined the archived proviral sequences, viral subtype and drug resistance mutations from a cohort of HIV+ patients with undetectable viral load undergoing HAART as first-line therapy using next-generation sequencing for near full-length virus genome (NFLG) assembly. HIV-1 consensus sequences representing NFLG were obtained for eleven patients, while for another twelve varying genome coverage rates were obtained. Phylogenetic analysis showed the predominance of subtype B (83%; 19/23). Considering the minority variants, 18 patients carried archived virus harboring at least one mutation conferring antiretroviral resistance; for six patients, the mutations correlated with the current ARVs used. These data highlight the importance of monitoring HIV minority drug resistant variants and their clinical impact, to guide future regimen switches and improve HIV treatment success.

Keywords: HIV-1; quasispecies; minority resistance mutations; HAART; drug resistance; undetectable viral load

1. Introduction

According to the Joint United Nations Programme on HIV/AIDS, approximately 36.7 million people were living with the human immunodeficiency virus (HIV) worldwide at the end of 2016, making HIV infection a major public health problem [1]. One of the factors related to the increased number of people living with HIV is the greater access to highly active antiretroviral therapy (HAART), which is strongly associated with an increase in the life expectancy and quality of this population. By 2016, 19.5 million people had access to HAART, an increase of 6% over 2015 [2]. In Brazil, it has been estimated that about 830,000 individuals were living with HIV/AIDS by the end of 2016,
representing an HIV prevalence rate of 0.24%. The access to HAART reached approximately 490,000, which represents more than half of the estimated Brazilian HIV-positive individuals. Of this total, approximately 450,000 had an undetectable viral load at least six months after HAART initiation, one indicator of therapeutic success [1]. According to the Brazilian Ministry of Health, virologic failure occurs when HAART fails to suppress and sustain a person’s undetectable viral load after six months of initiating or modifying treatment, in addition to the detection of viral load in patients who were previously undetectable [3]. Brazil has been cited as a reference in access to HAART since November 1996, when the government guaranteed universal and free access to therapy to the Brazilian HIV+ population. New recommendations that stimulate the initiation of HAART for all HIV-positive individuals, independent of CD4+ T lymphocyte counts, were implemented in October 2013 in order to reduce the transmission of the virus [4].

A remarkable HIV-1 genetic variability results from mutational events associated with the high error-prone rate of the viral reverse transcriptase (RT), high virus replication rates and homologous recombination events. HIV-1 diversity imposes an important clinical challenge, as it allows the virus to adapt to and evade immune responses and antiretroviral therapy [5], therefore influencing diagnosis and treatment [6]. In the context of continuous HIV-1 genetic variability within each individual, several highly related but genetically distinct variants are generated and referred to as viral quasispecies [7]. The quasispecies heterogeneity associated with the selective pressure exerted by the immune system influences virus persistence and pathogenicity, allowing the adaptation of the quasispecies through intrahost persistence and the ability to outgrow other less adapted variants [8].

The HIV evolutionary dynamics associated with ART pressure allows the appearance of drug-resistant variants [9,10]. Antiretroviral resistance is an important public health concern since it limits therapeutic options, causes treatment failure and can be transmitted, compromising future treatment option in untreated individuals [11,12]. Rates ranging from 4 to 14% of newly diagnosed patients have been reported in different parts of the world to be infected with strains that have at least one transmitted drug resistance mutation (TDRM) [13–16].

Aiming to establish more efficient HAART regimens, genotypic assays to detect drug resistance-associated mutations have been widely used in the clinical setting [17]. This practice has lately been greatly benefited by the use of Next-Generation Sequencing (NGS), which provides a large data volume in a cost-effective and highly sensitive way. Of note, NGS allows the detection of HIV minority variants, previously undetectable by Sanger sequencing, which has a detection limit of 10–25% frequency in the viral population [18–20]. Recent studies using NGS for HIV-1 sequencing were able to detect minority variants below 1% in frequency in the viral population, allowing the identification of drug-resistant minority variants, the study of transmitted resistant viruses and the impact of those minority variants on treatment efficacy [21–25]. Moreover, HIV-1 sequencing by NGS significantly contributes to the analysis of viral genetic diversity, evolutionary and epidemic processes, since near full-length genomes (NFLG) can be obtained [26]. These genomes contribute to the growing interest in tests that simultaneously probe multiple genomic regions that are targetted by antiretroviral drugs acting on different steps of the virus replicative cycle [27].

HIV minor drug resistant variants may persist in the infected individual [28,29]. Despite their low replicative capacity due to the presence of drug-resistance mutations, they may persist as archived proviruses in peripheral blood mononuclear cells (PBMCs) for several years and may have a long-term impact on therapeutic response [30–32]. However, the clinical significance of these minority drug resistant variants is still uncertain, as their role in the future response to treatment is not fully understood. Some studies have shown the association of such variants with increased risk of treatment failure in treated patients, as well as in patients with no previous ART history [21,33–38]. On the other hand, different studies did not find any influence of these minority resistant variants on treatment response [39–45]. The association between HIV minor non-nucleoside reverse transcriptase inhibitor (NNRTI) resistant variants and a worse treatment prognosis has been described [25,29,46–48]. Patients carrying those variants had a three-fold higher risk of treatment
failure compared to patients without NNRTI-resistance mutations when subjected to therapeutic regimens based on this antiretroviral (ARV) class [49,50]. With respect to patients under suppressive therapy, a single study was conducted that showed the presence of minority drug-resistance mutations in five out of eleven patients and a large variability of the archived proviral epitopes [51]. This study highlights the importance of more sensitive genotypic resistance tests and further studies about the influence of these mutations on HAART outcomes, especially in patients under successful therapy, since this cohort is still scarcely discussed.

In the present study, we have assessed the presence of HIV minority drug-resistant variants archived in PBMC samples in a Brazilian cohort of chronic HIV patients undergoing HAART as first-line therapy and with undetectable viral load. Upon analyzing the HIV antiretroviral resistance profiles of the patients by NGS, our study was able to evidence a high prevalence of drug resistance mutations in this cohort, despite the therapeutic success achieved by their carriers. This is the first study investigating the presence of HIV minority drug resistance mutations among Brazilian patients under virologic control.

2. Materials and Methods

2.1. Study Population and Sample Collection

A cross-sectional study was carried out among patients attending a sexually transmitted diseases/HIV ambulatory at Hospital Federal de Ipanema, Rio de Janeiro, Brazil, between February and July 2016. These patients were recruited during the clinical follow-up routine and had a 10 mL sample of whole peripheral blood collected. A questionnaire was applied to collect epidemiological (date of birth, sex, risk behavior) and clinical (date of HIV diagnosis, HIV-1 viral load, CD4+ T-cell counts, CD8+ T-cell counts, treatment history) data. Written informed consent was obtained from all participants and data were processed using unique identifiers to ensure confidentiality.

Inclusion criteria included age ≥ 18 years, being under first-line HAART and having undetectable HIV viral load for at least 12 months prior to collection date. Patients who had a history of previous virological failure, that underwent a change in the therapeutic regimen due to intolerance or poor adherence, that were classified into clinical and/or immunological AIDS (had an AIDS defining disease or CD4+ counts equal or less than 200 cells/mm³, according to the U.S. Centers for Diseases Control and Prevention criteria, 2014) [52] or that were under follow-up at the reference center for less than 12 months were excluded from the analysis. By using those strict criteria, we have been able to enroll a total of 32 patients among the whole patients followed-up at that site. This research was approved on March 29th, 2016 by the Ethics Committees in Research of the Brazilian National Cancer Institute (INCA) and of Hospital Federal de Ipanema (CAAE 52862016.9.0000.5274).

2.2. DNA Extraction and PCR of Proviral DNA

Plasma and buffy coats were separated by centrifugation of the whole blood, and the latter were used for genomic DNA extraction with the Genomic DNA Extraction Kit (Real Genomics, BioAmerica, Inc., Miami, FL, USA) following manufacturer’s specifications.

Nested PCR was carried out for the amplification of near full-length HIV genomes using a set of five overlapping fragments of approximately 2 kb each, or alternatively four fragments, with ~3 kb each, as previously described [53,54]. Each fragment overlapped the adjacents by an average of 400 bp (minimum 89 bp, maximum 555 bp). All reactions were performed in a Veriti® 96 Well Thermal Cycler (Life Technologies, Carlsbad, CA, USA) using Platinum™ Taq DNA Polymerase High Fidelity (Life Technologies) in a final volume of 25 µL. An overview of all PCR fragment coverage regions and list of primers used are provided in Table S1.

PCR products were visualized in 1% agarose gels. Duplicates were made for each PCR and pooled before proceeding to the purification step with the GFXTM PCR DNA and Gel Band Purification Kit (GE Healthcare, Chicago, IL, USA). The purified fragments were quantified in a NanoDrop ND
1000 apparatus (Thermo Scientific, Waltham, MA, USA) and diluted to 4 ng/µL before being pooled per sample. The final product pool per sample was diluted to 0.4 ng/µL for library construction.

2.3. Library Construction and NGS

Libraries were prepared using the Nextera XT DNA Sample Preparation kit (Illumina Inc., San Diego, CA, USA) according to the manufacturer’s protocol. Briefly, after the fragmentation step using transposon technology, libraries were subjected to a PCR where they were tagmented with Illumina sequencing adaptors and molecular tags (indexes) to identify each sample. A purification step was performed to select fragments of ~800 bp and libraries were quantified by qPCR with the KAPA library quantification kit (Kapa Biosystems, Wilmington, MA, USA). Individual libraries were then diluted to 4 nM, considering the mean size of each library and the quantification performed, and pooled. The final product was diluted to 12 pM and sequenced in a MiSeq Illumina platform (2 × 301 paired-end run) (Illumina) with 1% denatured PhiX DNA as a sequencing control.

2.4. Data Analysis

The analysis of the obtained files was performed in Geneious v.9.1.3 as described by Dudley et al. [55]. Briefly, the two fastq files generated per sample after the demultiplexing process were paired and trimmed at the ends to an error rate below 0.1%. The products were then used in the assembly of the viral genome sequence using an annotated HIV-1 HXB2 reference sequence made available by Dudley and colleagues, including information on resistance mutation-associated positions, genes, CDS and mature peptide positions (https://dholk.primate.wisc.edu/wiki/dho/public/page.view?name=default). The alignment parameters used in the assembly were as previously described [55]. For this alignment, ten iterations were used, where the first of the ten alignments was performed against the annotated reference sequence and the remaining nine used the consensus obtained in the previous step as reference, in order to reduce the influence of the reference used in the alignment product.

2.4.1. Analysis of Resistance Mutations

The annotated drug resistance mutations were based on the consensus of the International Antiviral Society [56] and the Stanford HIV Drug Resistance Database (available at http://hivdb.stanford.edu/hiv/). Transmitted drug resistance mutations (TDRM) were also included and defined according to the classification of the World Health Organization established by Bennett et al. [11] and in the TDRM database of the Stanford HIV Database. The HIV-1 HXB2 reference annotation was manually updated to include new positions at the lists mentioned above and resistance mutations described in the literature at the C-terminal region of RT covering the connection (CN) and RNase H (RH) subdomains [57–64]. The mutations evaluated in the RT C-terminal region are listed in detail in Table 1.

The variant finder of Geneious v.9.1.3 was used to call nucleotide variants from the reference sequence at frequency higher than 1%. This tool evaluates nucleotide substitutions found in the alignment along with their frequency and whether they are synonymous or non-synonymous. The frequency and the number of reads representing each ARV-resistant variant was determined in relation to the annotated reference and exported to an Excel file. Variants between 20% and 1% of frequency were considered as minor variants, since 20% is the minimum frequency associated with detection of mutations by commercial genotype resistance assays that use Sanger sequencing [65].
<table>
<thead>
<tr>
<th>Subdomain</th>
<th>Mutation</th>
<th>ARV Class Associated with Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN</td>
<td>E312Q</td>
<td>NRTI</td>
</tr>
<tr>
<td></td>
<td>Y318F/W</td>
<td>NNRTI</td>
</tr>
<tr>
<td></td>
<td>G335D/C (polymorphism)</td>
<td>NRTI</td>
</tr>
<tr>
<td></td>
<td>N348I</td>
<td>NRTI and NNRTI</td>
</tr>
<tr>
<td></td>
<td>A360I/V</td>
<td>NRTI and NNRTI</td>
</tr>
<tr>
<td></td>
<td>V365I</td>
<td>NRTI</td>
</tr>
<tr>
<td></td>
<td>T369I/V</td>
<td>NRTI and NNRTI</td>
</tr>
<tr>
<td></td>
<td>A371V</td>
<td>NRTI and NNRTI</td>
</tr>
<tr>
<td></td>
<td>A376S</td>
<td>NRTI and NNRTI</td>
</tr>
<tr>
<td></td>
<td>E399D/G</td>
<td>NRTI and NNRTI</td>
</tr>
<tr>
<td></td>
<td>A400T (polymorphism)</td>
<td>NRTI</td>
</tr>
<tr>
<td>RH</td>
<td>D488E</td>
<td>NRTI</td>
</tr>
<tr>
<td></td>
<td>Q509L</td>
<td>NRTI and NNRTI</td>
</tr>
<tr>
<td></td>
<td>Q547K</td>
<td>NRTI</td>
</tr>
</tbody>
</table>

NRTI, nucleotide/nucleoside reverse transcriptase inhibitors; NNRTI, non-nucleoside reverse transcriptase inhibitors; * References [57–64].

2.4.2. Phylogenetic Analysis

A consensus sequence was derived for each sample from the reference assembly with Geneious using the 50% stringency setting. HIV-1 subtype classification of each query sequence was inferred through phylogenetic analysis performed with the maximum likelihood (ML) method using PhyML v.3.0 [66] and the best model of nucleotide substitution was inferred with Model Generator [67]. Sequences suggestive of intersubtype recombination were further analyzed with the bootscanning tool of Simplot v.3.5.1 [68] for determining patterns of recombination and the HIV-1 subtypes involved in the recombination event. The following parameters were used: window = 400 pb; steps = 40 pb; T/t = 2.0; gapstrip = on; replicas = 100; nucleotide substitution model = F84; method = Maximum Likelihood. Recombinant strains were further confirmed by phylogenetic analysis of individual HIV-1 subtype genomic fragments as suggested by the bootscanning breakpoint analysis.

2.4.3. Analysis of Viral Tropism

For the phenotypic prediction of HIV co-receptor usage, the reads spanning the complete region of V3 loop of the env gene were selected and unique haplotypes were identified and their relative frequency considered. Amino acid sequences of the haplotypes were analyzed only if started and ended with cysteine residues and contained 32–38 amino acid residues, a pattern consistent with functional sequences [69]. Valid haplotypes were analyzed with the Geno2Pheno algorithm (available at http://coreceptor.geno2pheno.org) [70] using a false positive rate cut-off of 10% for classification as R5- or non-R5-using viruses, as recommended by the European Consensus Group on Clinical Management of HIV-1 Tropism Testing and used in several studies [54,71–74].

3. Results

The median age of the patients included in the study was 38 years and 75% were male (Table 2). The median time of antiretroviral treatment at enrollment was approximately three years and the median baseline CD4⁺ T-cell count was 712.5 cells/mm³. All patients have reported being infected by sexual transmission, 56% of which were men who have sex with men. Most patients (19%; 59%) were under HAART composed of tenofovir (TDF), lamivudine (3TC) and efavirenz (EFV) at the time of sample collection, but less frequent ARV schemes were also used (Table 3).
Table 2. Demographic and clinical characteristics of the 32 human immunodeficiency virus (HIV)-positive participants of the study.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males (%)</td>
<td>24 (75%)</td>
</tr>
<tr>
<td>Age (years) (mean ± SD)</td>
<td>40 ± 12.3</td>
</tr>
<tr>
<td>Median baseline CD4+ T-cell counts (cells/mm³; IQR₅₀)</td>
<td>712.5 (606.5–856)</td>
</tr>
<tr>
<td>Median baseline CD8+ T-cell counts (cells/mm³; IQR₅₀)</td>
<td>657.5 (529–1047.25)</td>
</tr>
<tr>
<td>Median time since HIV diagnosis (years; IQR₅₀)</td>
<td>4.7 (3.9–6.5)</td>
</tr>
<tr>
<td>Median time from HIV diagnosis to antiretroviral therapy initiation (years; IQR₅₀) *</td>
<td>1.2 (0.6–2.8)</td>
</tr>
<tr>
<td>Median time of treatment (years; IQR₅₀)</td>
<td>3.1 (2.4–3.9)</td>
</tr>
</tbody>
</table>

* Missing in one patient; SD, standard deviation; IQR₅₀, interquartile range.
Table 3. Distribution of the antiretroviral resistance mutations and major envelope tropism found across the 23 HIV-1 genome sequences analyzed.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Protease Mutations (Coverage; Frequency)</th>
<th>Reverse Transcriptase Mutations (Coverage; Frequency)</th>
<th>RT Connection Mutations (Coverage; Frequency)</th>
<th>RT RNase H Mutation</th>
<th>Integrase Mutations (Coverage; Frequency)</th>
<th>Envelope Mutations (Coverage; Frequency)</th>
<th>HAART Regimen</th>
<th>Tropism ¹</th>
<th>Subtype/URF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>AZT + 3TC + NVP</td>
<td>100.0%</td>
<td>X4/R5X4</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>AZT + 3TC + EFV</td>
<td>63.0%</td>
<td>X4/R5X4</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>97.2%</td>
<td>R5</td>
</tr>
<tr>
<td>4</td>
<td>147V (3589; 10.3%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>AZT + 3TC + ATV</td>
<td>100.0%</td>
<td>B</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>E399D (2701; 99.2%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>97.1%</td>
<td>R5</td>
</tr>
<tr>
<td>6</td>
<td>D30N (8012; 2.8%) M46I (8639; 1.9%)</td>
<td>M41L (5444; 99.8%)</td>
<td>T369V (9726; 39.6%) E399G (8334; 4.9%)</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>96.1%</td>
<td>R5</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>84.8%</td>
<td>R5</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>97.2%</td>
<td>R5</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>97.0%</td>
<td>R5</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>91.6%</td>
<td>R5</td>
</tr>
<tr>
<td>11</td>
<td>D30N (4574; 49.3%) M46I (4378; 45.6%)</td>
<td>M41L (5791; 99.5%)</td>
<td>D67N (6465; 72.9%) K70R (6475; 77.2%) T215Y (9219; 58.2%)</td>
<td>E399D (2051; 99.9%)</td>
<td>NA</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>99.4%</td>
<td>R5</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>99.7%</td>
<td>R5</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>96.0%</td>
<td>R5</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>X4/R5X4</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>E399D (14,872; 1.5%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + FPV</td>
<td>NA</td>
<td>B</td>
</tr>
<tr>
<td>16</td>
<td>D30N (4574; 49.3%) M46I (4378; 45.6%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>99.3%</td>
<td>R5</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>R5</td>
</tr>
<tr>
<td>18</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>R5</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>R5</td>
</tr>
<tr>
<td>20</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>R5</td>
</tr>
<tr>
<td>21</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + EFV + FTC</td>
<td>99.9%</td>
<td>R5</td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>R5</td>
</tr>
<tr>
<td>23</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>96.0%</td>
<td>R5</td>
</tr>
<tr>
<td>24</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>X4/R5X4</td>
</tr>
<tr>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>99.9%</td>
<td>R5</td>
</tr>
<tr>
<td>26</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>99.3%</td>
<td>R5</td>
</tr>
<tr>
<td>27</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>99.1%</td>
<td>R5</td>
</tr>
<tr>
<td>28</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>99.7%</td>
<td>X4/R5X4</td>
</tr>
<tr>
<td>29</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>R5</td>
</tr>
<tr>
<td>30</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>R5</td>
</tr>
<tr>
<td>31</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>R5</td>
</tr>
<tr>
<td>32</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>TDF + 3TC + EFV</td>
<td>100.0%</td>
<td>R5</td>
</tr>
</tbody>
</table>

* The near full-length genomes are in bold; NA, not available; -, no mutations found; * only partial sequence available; ¹ X4/R5X4: CXC chemokine receptor 4 (CXCR4) and/or CCR5/CC chemokine receptor 5 (CCR5) tropism profile; R5: CCR5 tropism profile.
Nine samples (28%) failed to have more than one viral DNA fragment PCR-amplified and were excluded from further analyses, while 23 samples amplified at least two fragments and were sequenced using NGS. Of the latter, eleven samples had the near full-length genome (NFLG) sequenced, and the remaining samples were missing one or two DNA fragments over the genome. The average number of reads obtained per sample was 774,536 (374,780–1,647,090). After assembling, the average coverage per nucleotide position was 7193 and it was homogeneous over the regions sequenced. The Gag CDS was complete for 22 samples (96%), the Pol CDS for 15 (65%) and the Env CDS for 17 (75%).

Most samples (18; 78%) carried archived viruses harboring at least one mutation associated with antiretroviral resistance (Table 3). Among them, 13 samples had mutations in the RT region, including the connection domain. Four carried resistance only to NRTI, six only to NNRTI, while three presented resistance mutations to both NRTI and NNRTI. The ART regimen of these 13 patients included two NRTI and one NNRTI. One sample (patient #28) presented four thymidine analogue-associated mutations (TAMs), with frequency varying from 38.7 to 99.5%, that are associated with intermediate resistance to TDF, part of the patient’s current therapy regimen. Three samples presented viruses harboring the M184V mutation (frequency varying from 1.3 to 8%) that confers high-level resistance to 3TC, used by those patients. However, this mutation is known to increase the susceptibility to zidovudine (AZT) and TDF, that were also part of their ART regimen. Two samples presented the E399G mutation with frequency of 1.5% and 4.9%. This mutation is associated with resistance to EFV which was included in these patients’ therapy. Tree samples presented major resistance mutations in the protease sequence, and while two of them used protease inhibitors (PI), the mutations observed were not related to the PI used. Resistance mutations in the integrase region were found in five samples. Most of them were minority variants, present in frequency below 11%, while one (T66I) was found with frequency of 31% in patient #29. Only one mutation associated with resistance to entry inhibitors was found in the env gene, with a frequency of 3% in patient #13.

Follow-up data of the patients under study (i.e., HIV viral load and CD4+ T-cell counts) were obtained during the preparation of this report to evaluate the clinical progress of the viral infection. The time information varied between 9 and 17 months after the initial collection time. Median baseline CD4+ T-cell counts were 758 cells/mm$^3$ ($n = 31$; IQR$_{50}$ 715–963) and median baseline CD8+ T-cell counts were 8675 cells/mm$^3$ ($n = 25$; IQR$_{50}$ 657.75–1035), higher compared to the values observed at the time of patient enrollment. With respect to HIV viral load ($n = 31$), only one patient presented a detectable load (patient #8; 222 copies/mL) approximately one year after inclusion in the study. There was no change in the antiretroviral regimen of any patient, except for patients using LPV/r, who were switched to DRV/r due to changes in first-line regimens recommended by the Brazilian Ministry of Health.

The HIV-1 consensus sequence was generated for each sample for phylogenetic analysis. Most NFLG obtained were from HIV-1 subtype B ($n = 9$; 82%; Figure 1). The two remaining were classified as distinct unique recombinant forms (URF) comprising subtypes B and F1 based on Simplot analysis (Figure 2). The sequences from the twelve non-complete genomes were predominantly of subtype B ($n = 10$), and two URF were found, one URF-BF1 and one URF-BC (Figure 2).

Two of the 11 NFLG described harbored stop codons in one or more open reading frames (ORF). One (patient #5) had multiple stop codons in the gag and one in the pol gene, consistent with APOBEC-mediated hypermutation (G→A changes), yet the number of substitutions did not allow the classification of this sequence as hypermutated. The other (patient #15) had a frameshift deletion (7 bp) within the pol gene, generating multiple stop codons in that ORF.

Virus envelope tropism analysis was performed with the reads spanning the V3 env region of the sequences and the results are depicted in Table 3. All but one sample contained the V3 loop region sequenced and were analyzed. We found that 17 (77%) patients showed CC chemokine receptor 5 (CCR5) usage as the major tropism profile (R5), while 5 patients (23%) presented predominantly the CXC chemokine receptor 4 (CXCR4) tropism profile (X4 and/or R5X4).
Viruses 2017, 9, 392 8 of 18

Figure 1. Phylogenetic maximum likelihood analysis of HIV near full-length genomes obtained in this study. The analysis was conducted with 1000 bootstrap iterations and included eleven HIV-1 proviral sequences from Hospital Federal de Ipanema, Rio de Janeiro (represented in bold) and references of HIV-1 subtypes (represented by the subtype, country, year and GenBank accession number). Only bootstrap values greater than 0.7 are shown. The gray box highlights the sequences classified as subtype B, among which nine of the eleven sequences were placed. The two remaining sequences represented unique recombinant forms comprising subtype B and other subtypes.

Figure 2. Phylogenetic classification of recombinant viruses considering the phylogeny and similarity analyses. The gray shading patterns represent the different subtypes: black for subtype B, gray for subtype F1 and white for subtype C. Sample IDs are represented at the left of each virus structure, which is in-scale relative to the genomic coordinates of the reference HXB2 genome at the top.

4. Discussion

Several viral factors are known to affect and modulate progression of HIV-1-related AIDS from the early asymptomatic phase of the disease, like the virus genetic diversity, viral fitness and co-receptor tropism [75–79]. In this context, we analyzed the archived HIV-1 proviral sequences from HIV patients
in the early chronic phase of infection under HAART and with undetectable HIV viral load through NGS. This type of population is poorly explored in high-throughput studies, despite being of great relevance due the increased access to HAART worldwide. In the present report, we describe the genetic variability, prevalence of drug resistance-associated mutations and viral tropism of those patients at an ultra-deep level. It is noteworthy that this is the first study investigating the presence of minority drug resistance mutation among Brazilian patients under virologic control.

The majority of our patients self-reported as men who have sex with men (56%), had a median time of HIV-1 diagnosis to initiation of ART of about 1.2 years and they have been successfully treated for approximately three years. With the new recommendations by the Brazilian Ministry of Health implemented at the end of 2013, it is expected that the time lapse between HIV diagnosis and HAART initiation approaches zero, which contributes to reducing virus transmission. On the other hand, the test-and-treat approach highlights the importance of monitoring the prevalence of antiretroviral resistance and TDRM at population level, to avoid resistance dissemination and therapeutic failures due to transmitted drug resistance.

As seen in Latin America and the Caribbean countries, our study found a prevalence of subtype B infections (19; 83%) [80–82]. Although subtype B is the major HIV-1 genetic clade circulating in the Brazilian epidemic, the overall prevalence of non-B strains, like URF_BF1, URF_BC, and particularly subtype C and CRF31_BC in the South of Brazil, is increasing [83–92]. None of the BF1 (n = 3, 13%) or the BC recombinant structure (n = 1, 4%) identified in this study showed similarity with the recombination patterns of known CRFs or other recombinants. CRFs were also not detected among HIV-1 NFLG obtained from children and adolescents from São Paulo, Brazil, which corroborate the hypothesis that novel recombinants are continually arising in the Brazilian epidemic [89]. The analysis of HIV-1 NFLG, as opposed to most previous Brazilian HIV subtype studies that analyzed HIV sequence information only from shorter fragments, may unveil an underestimation of the occurrence of recombinant viruses in the country.

The integrated state of HIV allows its persistence as archived proviruses in PBMCs. Consequently, this may potentially compromise the efficacy of targeted antiretroviral drugs exerting a long-term impact on responses to HAART [31,93]. Also, it should be noted that routine genotyping tests focus on plasma circulating variants and on a limited number of short genomic fragments, despite the increasing need to simultaneously probing multiple genomic regions that target different steps of the viral life cycle. These facts, coupled with previous observations that standard bulk sequencing cannot fully access the spectrum of viral variants archived in the proviral DNA, justify the use of NGS technologies to study proviral DNA from PMBC as a valuable source for resistance analysis [17,94–96].

We found 18 patients (78%) who were successful responders to first-line ART (undetectable HIV viral load) that carried archived viruses harboring at least one drug resistance-associated mutation. Only a single study has been previously published in this regard on HIV patients experiencing suppressive therapy [51]. The authors evidenced the presence of antiretroviral resistance mutations above 1% of frequency in five of the eleven patients analyzed (45%). The difference in prevalence found in both studies can be attributed to differences in the ultradepth sequencing method used (MiSeq - Illumina versus 454 Life Science - Roche GS Junior), the regions covered by the study (NFLG versus Gag, Pol and Nef regions) and the associated coverage obtained, as two patients of their cohort only presented mutations under 1%.

Eight of our proviral sequences (8/23, 22%) harbored only minority variants with drug resistance-associated mutations. This finding underscores the use of more sensitive HIV genotyping techniques, since the standard genotypic resistance testing using bulk sequencing of the viral population can only detect viral variants that constitute over 15–20% of the total viral population [65], and likely underestimates the overall prevalence of resistant variants and may impact on the surveillance of HIV resistance and on the clinical management of treated patients.

The clinical significance of drug-resistant minority variants is still uncertain and has been addressed by several studies, due to its potential impact on the response to future treatment schemes.
Some studies point out that these resistance variants would be favored by the ARV selective pressure, leading to a substitution of the wild-type virus and consequently to therapeutic failure [21,33–38]. However, other studies found no influence of these minority drug-resistance variants on treatment response, highlighting the role of the majority, not the minority variants, in this outcome [39–45].

Despite the high prevalence of drug resistance mutations (18/23) and of mutations associated with the current ARV therapy of the patients (6/23) in our study, only one patient had detectable viral load 15 months after inclusion in the study. Unfortunately, we failed to PCR-amplify the HIV RT polymerase domain of this patient, and only the E399D mutation was found in RT connection domain of that virus. Most of the drug resistance mutations associated with the current therapy regimens of the patients analyzed were found at low prevalence at the viral population (1.3%–8%), except for the patient harboring TAMs (patient #28; Table 3). Surprisingly, this patient maintained an undetectable viral load 14 months after inclusion in the study, paralleled by an increase in the CD4+ T-cell counts (681 cells/mm³) compared to those at the time of enrolment (535 cells/mm³). The low prevalence observed in most of the mutations found and the high adherence reported by the clinicians of the program may have contributed to the high therapeutic success rate observed in these patients one year after inclusion in the study.

In Brazil, two entry inhibitors/antagonists, enfuvirtide (ENF) and maraviroc (MVC), have been used since 2005 and 2007, respectively, in therapeutic rescue strategies for patients failing previous ARV regimens [97,98]. The resistance mutation to ENF observed in one of our patients (#13; Table 3) emphasizes the importance of including this region at the genotypic resistance analyses, since resistance to ENF is characterized by a low genetic barrier [99–101] and can lead to therapeutic failure if used in patients carrying resistance to that drug.

One patient (#11) was found to harbor proviruses with two transmitted drug resistance mutations to protease inhibitors (PI), all of them as minority variants (frequency between 1.9% and 2.8% of the virus population). The treatment history of each patient was considered to define the transmitted resistance. All patients were under NRTI use, so it was not possible to characterize NRTI-associated TDRM. The prevalence of TDR found in this study (1/20; 5%) was in agreement with the moderate TDR prevalence reported by other Brazilian studies, usually ranging from 5 to 10% [102–107].

Only three patients studied herein were under PI use in their treatment schemes, and two of those harbored resistance mutations in the protease. While five patients carried integrase inhibitor-related mutations, none used integrase inhibitors, which were only approved as rescue therapy in Brazil at the time of study enrollment. In those cases, mutations were carried from the virus donor to the recipients (transmitted resistance) or arouse in the patient’s virus quasispecies.

With respect to HIV coreceptor tropism, R5 viruses usually predominate during primary HIV infection, whereas the transition to X4 viruses occurs in later stages of HIV disease, being associated with more rapid CD4+ T-cell depletion and consequently to AIDS progression [108–112]. Genotypic predictors prove to be highly concordant with phenotypic data and can be reliably used to determine viral tropism with better results in PBMC than in plasma samples [113]. In this study, we used the Geno2pheno algorith with a false positive rate (FPR) cutoff of 10% [54,71–74]. Prediction of coreceptor usage showed that most individuals (17/22; 77%) presented only R5-tropic viruses, while five individuals presented X4/R5X4-tropic viruses. The prevalence of R5-tropic viruses was similar to the 78% prevalence found by de Azevedo et al. [114] in another Brazilian cohort.

Hypermutated proviral sequences were detected in only one individual (patient #26) through the identification of an excessive G→A change pattern, consistent with APOBEC3G/F signature. Several early stop codons resulted from these nucleotide substitutions were observed along that HIV genome. The evidence of hypermutation as an apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC) action was also highlighted by the Stanford HIV drug resistance database during the analysis of resistance mutations for that virus.

An important limitation of the present study is that the proviral sequences have not been fully evaluated for replication ability, and may be therefore defective or inactive. We only retrieved NFLG
sequences from 11 of the 23 patients studied, and no ORF intactness was evaluated for the incomplete genomes. Two of the 11 NFLG sequences had stop codons due to hypermutation or to a frameshift deletion, and are thought to be defective for replication. However, missense mutations that could render the virus defective were not evaluated, and the number of defective proviruses is likely underestimated here.

We are aware that additional studies should be carried out to validate our findings in a larger population. Our strict inclusion criteria did not allow the enrollment of a large cohort for this study. Another limitation was the difficulty of PCR-amplifying the archived proviral genomes from some patients in a setting of undetectable HIV viral load and early chronic infection, where PBMC archived HIV reservoirs are thought to be small. Even with the utilization of diverse strategies, some genomic regions could not be amplified for some patients. Despite we have carried out the PCR amplifications of the virus samples in duplicates to avoid false-positive mutations introduced by PCR errors, we did not quantify the input of proviral copies subjected to the analyses, what could impact the estimates of the variation found in the quasispecies. The subjects under study were in the chronic phase of infection, yet not in advanced stage of the disease (as judged by the absence of AIDS and the high number of CD4+ T-cell counts), and their proviral quasispecies population was expected to be fully established. In addition, the fact that we achieved rates of 1–2% of frequency for particular DRMs in several patients suggests that we were able to reach such sensitivity rates in our assays.

The analysis of resistance-associated mutations in HIV-positive patients is an important issue when considering the context of broad access to antiretroviral treatment and high rates of therapeutic success, one of the main goals to be achieved worldwide against HIV infection in the coming years. The high rate of resistance-associated mutations found in our cohort, composed of patients with undetectable viral load undergoing HAART as first-line therapy, directs attention to the selection of antiretroviral resistant variants in a context of therapeutic success and early in chronic infection. These observations underscore the importance of further studies in order to better correlate the presence of these drug resistance mutations and response to ART to investigate their potential association with therapeutic failure and to establish effective public policies to decrease their prevalence and transmission.

Supplementary Materials: The HIV-1 consensus sequences were deposited at the GenBank nucleotide database under the accession numbers MG571979-MG572011. The raw NGS reads were submitted to the Sequence Read Archive (SRA) under the numbers SRR6324863-SRR6324885. The following are available online at www.mdpi.com/1999-4915/9/12/392/s1.

Acknowledgments: This work was supported by the Rio de Janeiro State Science Foundation—FAPERJ (grant # E-26/201.331/2014 to Marcelo A. Soares) and by the Brazilian National Research Council—CNPq (grants # 454623/2014-3 to Marcelo A. Soares and # 460346/2014-8 to Esmeralda A. Soares). Bruna M. Alves was recipient of a scholarship from the Brazilian Ministry of Education (CAPES) to spend six months at the Lady Davis Institute for Medical Research, McGill University, Montreal, Canada. This work is part of the requirements for Bruna M. Alves to obtain her title at the Graduate Program in Oncology of the Brazilian National Cancer Institute (INCA).

Author Contributions: Esmeralda A. Soares and Marcelo A. Soares conceived and financed the study. Bruna M. Alves, Juliana D. Siqueira, Ornella M. Botelho and Isabel M. Prellwitz performed all the molecular biology and the computer analyses of the work. Marianne M. Garrido and Sayonara R. Ribeiro were responsible for patient recruitment and follow-up, and collected all clinical and laboratory data of the patients. All authors wrote and revised the manuscript, and agreed with the submission of its final version.

Conflicts of Interest: The authors declare no conflict of interest.

References


1063–1071. [CrossRef] [PubMed]

PLoS ONE

with historical antiretroviral use.

HIV Med.

with single nucleoside reverse transcriptase inhibitor resistance mutations.

Yazdanpanah, Y.; Harrigan, P. R.; et al. Highly frequent HIV-1 minority resistant variants at baseline of The Provir/Latitude 45 Project. Resistance mutations and CTL epitopes in archived HIV-1 DNA of patients

J. Virol.

variation and spurious detection rather than transmission and evolution of multiple viral variants.

J. Virol.

Prevalence and impact of minority variant drug resistance mutations in primary HIV-1 infection. PLoS ONE 2011, 6, e28952. [CrossRef] [PubMed]


