Analysis of the Complete Genome Sequence Reveals a High Degree of Conservation of the HIV-1 Subtype B in Cuba

Madeline Blanco de Armas1a*, Liuber Yans Machado Zaldivar1a, Héctor M. Díaz Torres1,2, Enrique Noa1, Dania Mercedes Romay Franchi1

1 AIDS Research Laboratory, Carretera de Tapaste and Autopista Nacional
San José de las Lajas, Mayabeque, Cuba.
2 Hospital Hermanos Ameijeiras, Havana, Cuba
*a both authors equally contributed to this work

*Corresponding Author: Madeline Blanco de Armas, AIDS Research Laboratory, Carretera de Tapaste and Autopista Nacional, San José de las Lajas, Mayabeque, Cuba, E-mail: madeniba@infomed.sld.cu

Abstract: The analysis of the structural genes of HIV-1 in different Cuban patients showed a preserved behavior for subtype B, which motivated to study the complete length of the viral genome in patients with this pattern and to demonstrate the high degree of conservation of this genetic variant. More than 8,000 bp of the HIV-1 genome were sequenced from four HIV-positive individuals with no epidemiological linkage, previously identified as subtype B by pol gene analysis. Phylogenetic and recombination analysis showed the conserved pattern of circulating subtype B in the Cuban seropositive population. The results of the present study strengthened the epidemiological surveillance system of the National STI / HIV / AIDS Program in Cuba.

Keywords: HIV HIV-1, full genomes, molecular epidemiology, subtype B, Cuba

Abbreviations: Circulating recombinant forms: CRF; Unique recombinant forms: URF ; Men having sex with men: MSM

1. INTRODUCTION

The complete HIV-1 genome sequence was first published in 1985 by Ratner et al and subsequently numerous reports documented the extraordinary genetic diversity of HIV-1, derived from the high rate of mutations and recombination of this virus. Frequency of circulating recombinant forms (CRFs) and unique recombinant forms (URFs) has increased in multiple geographic areas, where they co-circulate different genetic forms. Population movements have also led to the spread and epidemic expansion by the world of genetic variants and recombinant forms. On the other hand, the event of transmission of the same genetic form and its propagation within a group with specific risk behaviors (founding effect), has resulted in the establishment of the epidemic in that area, where certain genetic variants predominate, over others. However, due to the rapid evolution of HIV and its global spread, other variants emerge and recombine.

In Cuba, in the 1980s, subtype B had its founding effect on the population of men who have sex with men (MSM), which is also the majority risk group. Other non-B subtypes were introduced by people heterosexual men, who acquired the infection in different African countries and, on their return, transmitted it to their (female) sexual partners and have spread less in the HIV-positive population. After more than two decades, the progressive appearance of different CRFs and multiple URFs has been observed, due, among other causes, to the recombination of the subtypes circulating in Cuba. Different studies carried out in Cuba for the molecular characterization of the structural genes of HIV-1 showed in patients characterized as subtype B, a conserved pattern, for each of the genes studied. Subtype B is the genetic variant of HIV-1 most disseminated in the Cuban epidemic from its beginnings to the present day, so it is of interest to study the complete genome length in patients with this pattern and to demonstrate the high degree of conservation observed in the predominant subtype in Cuba.
2. MATERIALS AND METHODS

2.1. Samples Preparations

Ten mL of whole blood were collected from four HIV-1 seropositive patients with no documented epidemiological link and previously classified as subtype B according to the pol gene. The proviral DNA was obtained using the High Pure Viral Nucleic Acid Kit (Roche Diagnostics GmbH, Germany) according to the manufacturer’s instructions.

The viral load, at the time of sampling, was determined using the Cobas AmpliPrep - Cobas TaqMan HIV-1 test version 2.0 and the CAP-CTM 48 (Roche) kit. The CD4 values in percent and absolute value at the time of sampling were obtained from the SIDATRAT database of the Institute of Tropical Medicine Pedro Kouri.

Sequencing of the complete HIV-1 genome was performed according to the protocol described by Nadai et al in 2008 and using the CEQ Dye-labeled dideoxy terminator cycle sequencing kit (Beckman Coulter, Company, Beverly, United States United).

The sequences were assembled and edited using the Sequencer v 5.0 program (GenCode Corporation), for further realization of a BLAST (http://www.ncbi.nlm.nih.gov/BLAST). The subtype was determined by the REGA HIV (http://www.ncbi.nlm.nih.gov/BLAST). The trees were visualized and edited by the Muscle program present in the SITARATRAT database of the Institute of Tropical Medicine Pedro Kouri.

The sequences were assembled and edited using the Sequencer v 5.0 program (GenCode Corporation), for further realization of a BLAST (http://www.ncbi.nlm.nih.gov/BLAST). The subtype was determined by the REGA HIV (http://www.ncbi.nlm.nih.gov/BLAST). The sexual orientation of the patients studied was MSM. Table 1 illustrates the epidemiological and clinical characteristics of the patients analyzed.

2.2. Phylogenetic Analysis

Multiple sequence alignment was performed through the Muscle program present in the MEGA v 6.0. Samples were aligned with subtype B sequences from the Americas, Europe, Asia and Africa obtained from the HIV database -1 of Los Alamos (Supplementary material 1). The analysis for the determination of possible recombinants was performed using the RDP3 program.

For the selection of the best nucleotide substitution model, the jModel Test v 2.1.4 program was used. Phylogenetic trees were constructed using the maximum likelihood method, using the GTR + G + I model. To verify the support of the branches and calculate the numbers near the nodes, the SH-like and bootstrap tests were used (1000 Replicates) and considered reliable, those with values greater than 0.90 and above 70%.

The PhyML v 3.0 was used. The trees were visualized and edited by the Fig Tree program. Mutations associated with antiretroviral resistance were determined according to the algorithm available in the Stanford Database (http://hivdb.stanford.edu). The prediction of viral tropism was determined using the geno2pheno coreceptor tools (http://coreceptor.geno2pheno.org) and WebPSSM (https://indra.mullins.mcriobiol.org). Was hington.edu/webpssm) .

3. RESULTS AND DISCUSSION

The four patients included in the study were male and resided in Havana and had no epidemiological link. The sexual orientation of the patients studied was MSM. Table 1 illustrates the epidemiological and clinical characteristics of the patients analyzed.
Analysis of the Complete Genome Sequence Reveals a High Degree of Conservation of the Hiv-1 Subtype B in Cuba

Splicing of approximately 8000 bp was achieved, with the exception of the LTR sequences. The analyzed samples were classified as subtype B by the REGA HIV subtyping tool program and no presence of recombination was detected by analyzing the genome of the patients studied, as illustrated in the analysis of sample 12CU087 (Figure 1).

Figure 1. Recombinants Analysis by Boots can of Sequence 12CU087. the Boots can Analysis Performed with Windows Size 400 and Step Size 20

Phylogenetic analysis grouped the Cuban sequences with sequences belonging to the pandemic subtype B (Figure 2). Samples 12CU087 and 14CU05 were grouped with sequences from the United States and France, while samples 14CU06 and 14CU07 were related to sequences from Colombia, South America.

Figure 2. Phylogenetic Tree of Four Cuban Sequences of HIV B Subtype Full Genome. Tree Constructed using Maximum Likelihood (ML); Genetic Distance Estimated by Calculating Matrices According to GTR+G+I Model. Cuban Sequences are shown with Symbol ● and from other Geographic Regions were Grouped and Are Represented with Different Tones

* - Started antiretroviral therapy two days after sampling
** - x4 (FPR <5%), dual r5x4 (FPR ≥5% <20%), r5 (FPR ≥20%)
The HIV/AIDS epidemic in Cuba has been characterized since its beginnings to present a high genetic diversity, translated in the wide variety of introduced subtypes and the circulation of recombinant forms. The first molecular epidemiology studies of HIV-1 in Cuba were directed to the most hyper variable regions of the HIV-1 env gene and showed a predominance of subtype B and association of this variant in the population of MSM. The use of highly active antiretroviral therapy (HAART) in Cuba since 2001 and the subsequent introduction of genotypic assays for the determination of HIV-1 resistance to antiretroviral (ARV) allowed the analysis of a larger number of samples from treated and untreated patients. Although the results of the studies, mainly focused on the regions coding for the protease and reverse transcriptase enzymes of the pol gene, indicated an increase in recombinant forms, subtype B continued as a majority genetic variant in the Cuban seropositive population. The high genetic diversity of HIV-1 described in Cuba favored the origin of new viral variants, to the detriment of pure subtypes. However, the founding effect of subtype B in the epidemiological context of our country and the origin of many of the sequences of this subtype as of 1977, can explain the conservation of the genome of this genetic variant in Cuba. Several studies have linked genetic diversity with the presence of mutations associated with HIV-1 resistance to ARV, with no relationship between these variables. Non-detection of viral variants with ARV resistance profile favors the use of an effective HAART, which entailed the therapeutic success achieved by the patient identified with the sample number 12CU087. However, the presence of X4 and dual R5X4 viruses in two recently diagnosed patients is interesting. HIV-1 variants with these characteristics are most frequently detected in the final stages of infection, relating them to the rapid progression to AIDS and the early approach to HIV diagnosis in high-risk and early onset populations of HAART, should be one of the main strategies of health systems in the world, if it is to stop the epidemic by 2020.

Subtype B had multiple introductions in Cuba and is phylogenetically related to sequences belonging to the pandemic variant that circulates mainly in North America and Europe, which supports the epidemiological history reported by various authors since the beginning of the HIV/AIDS epidemic in our country. A recent study by Machado et al presents the relationship of Cuban sequences with those from the United States, Canada and Europe. However, Junqueria et al., 2011, supports the origin of Cuban subtype B with South America. Population movement and continuous exchange explain the phylogenetic relationship of the sequences studied with countries of North America, Europe and South America, respectively.

4. CONCLUSIONS

The above mentioned elements show a conserved pattern of subtype B in the samples studied, a result that feedback to the national STI / HIV / AIDS program in Cuba on the evolution of the majority genetic variant and its possible implications in the diagnosis, vaccine development and new therapeutic strategies.

Accession Number of Nucleotide Sequences

The accession numbers in the GenBank sequences studied in the present study are: KR914675-KR914678.

REFERENCES


Analysis of the Complete Genome Sequence Reveals a High Degree of Conservation of the Hiv-1 Subtype B in Cuba


Copyright: © 2017 Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.