Estimating of Origin and Evolutionary History of Human Immunodeficiency Virus Type 2 in Cuba

Liuber Y Machado^{1,*}, Orlando Martínez², Héctor M Díaz^{1,3}, Enrique Noa¹, Madeline Blanco¹, Dayamí Martín¹, Dervel F Díaz¹, María L Sánchez⁴, Marta Dubed¹

¹AIDS Research Laboratory, Cuba

²University of Information Science, Cuba

³Hermanos Ameijeiras Hospital, Cuba

⁴Provincial Center of Hygiene, Epidemiology and Microbiology, Cuba

Copyright ©2016 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License.

Abstract Background: Infection with human HIV-2 is endemic in West Africa. The virus originated from West African sooty mangabeys during the first half of the 20th century and an epidemic initiation in Guinea Bissau that coincides with the independence war (1963-1974). The HIV-2 group A is categorized as epidemic group. The presence of HIV-2 group A in Cuba has been previously documented. However, the evolutionary history of HIV-2 group A in the Cuban epidemic is unknown. The aim of this work is to estimate the origin and evolutionary history of the HIV-2 group A in Cuba. **Methods:** We used a Bayesian coalescent method to analyze the *env* gene of Cuban HIV-2 group A. The rate of nucleotide substitution was determined and was used to date the phylogenies and reveal the evolutionary history of HIV-2 group A in Cuba. **Results:** Multiple introductions of HIV-2 group A, mainly from Guinea Bissau and Portugal were detected. The most recent common ancestor of Cuban HIV-2 groups A was dated back to about 1972 (95 % HPD: 1966-1978). The rate of nucleotide substitutions was 5.02×10^{-3} substitutions per site per years (95 % HDP: $4.51-5.52 \times 10^{-3}$). **Conclusions:** The results of this study allowed for the first time to estimate the evolutionary history of HIV-2 in Cuba and establish the basis for phylogeographic and phylodynamics studies.

Keywords HIV-2, Origin, Evolution, Phylodynamic

and at lower rates than HIV-1, several countries have reported the presence of infection (3-6).

The origin of HIV-2 is due to an event of zoonotic transmission from simian immunodeficiency virus (SIV) from sooty mangabey (7-9) occurred during the first half of the twentieth century, followed by an initiation possible epidemic in Guinea Bissau, this coincides with the development of its war of independence (1963-1974) (10) and epidemiological factors that favored its expansion (11).

Phylogenetic analysis of HIV-2 sequences has resulted in nine monophyletic groups, with a predominance of A and B. HIV-2 group A is the predominant group in West Africa (Senegal and Guinea Bissau) and HIV-2 group B predominates in the Ivory Coast. The other groups have been documented in one or two individuals. Except for the G and H groups, groups C, D, E, F and I were isolated rural areas where people are in frequent contact with SIV infected mangabey (12, 13). In HIV-2 infected individuals the presences of a circulating recombinant form (CRF) 01_AB, has been detected (14, 15).

In Cuba, from 1986 to December 2015, 22 individuals were serologically confirmed as positive for antibodies to HIV-2 in the AIDS Research Laboratory (National Reference Laboratory for Human Retrovirus) (16; 17). Subsequently, a study of genetic characterization described the HIV-2 group A in Cuban patients, suggesting the occurrence of multiple introductions of this group in Cuban patients (18); which led to delve into the origin and evolutionary history of this retrovirus in Cuba.2. Materials and Methods

1. Introduction

The occurrence of infections caused by the human immunodeficiency virus type 2 (HIV-2) was initially restricted to West Africa, where the first isolates were obtained in patients with AIDS originating in Cape Verde and Guinea-Bissau (1, 2). Although HIV-2 is less expanded

2.1. Selection of the Sequences and Multiple Alignments.

Thirteen *env* gene sequences in Group A of HIV-2 from Cuban patients involved in the study by Machado *et al* in 2014 (18) were selected. GenBank accession numbers for the sequences reported here are from KJ677041 to

KJ677053. Epidemiological and demographic characteristics of individuals were collected at the time of the study.

Cuban sequences were aligned with 244 sequences of the *env* gene of HIV-2 group A, available in the database of Los Alamos, who met the requirement of equal or greater in length to the Cuban sequences. The alignment was performed using the Clustal X program and manually edited by the Bioedit program (19).

2.2. Phylogenetic Analysis

Maximum Likelihood (ML) phylogenetic trees were inferred under the TIM2+I+G (alpha parameters=0.547) nucleotide substitution model selected using the jModeltest program (20). The ML trees were reconstructed with PhyML program. Hueristic tree search was performed using the SPR branch swapping algorithm and the reliability of the obtained topology was estimated with the approximate likelihood-radio test (α LTR) based on the Shimodaira-Hasegawa-like procedure. The ML trees were visualized using the FigTree v 1.1.2 (<http://treebioedacuk/software/figtree>).

2.3. Reconstruction of Evolutionary History

The age of most recent common ancestor (TMRCA, years) and the evolutionary rate (μ , nucleotide substitutions per site per year) were estimated using a Bayesian Markov Chain Monte Carlo (MCMC) approach implemented in BEAST v 1.7.5 (21). Analyses were performed using the TIM2+I+G nucleotide substitution model and a relaxed uncorrelated lognormal molecular clock model. A Bayesian Skyline coalescent tree prior was first used to estimated μ and the TMRCA. MCMC chains were run for 10×10^6 generations. Effective Sample Size (ESS) and 95 % Highest Probability Density (HDP) values were inspected using Tracer v 1.6 (<http://evolve.zoo.ox.ac.uk/software/tracer>) to asses convergence and uncertainty of parameter estimates. (21).

3. Results

3.1. Phylogenetic Relationship of Cuban Sequences

Of patients, 30.8% acquired HIV-2 in countries where the infection is endemic and 69.2% were infected in Cuba. The route of transmission was sexual, predominate heterosexual behavior (Table 1)

The ML analysis of HIV-2 group A from Cuba and other countries revealed that Cuban sequences branched with multiple sequences for endemic and non-endemic regions of HIV-2 (Figure 1). Cuban sequences grouped in the cluster I and corresponding to residents in central and eastern regions of Cuba are more closely related to sequences from Guinea Bissau (α LRT = 0.82). The 12CU14 and 12CU15 sequences corresponded to a pair of seropositive of east region, where one acquired the infection in Guinea Bissau. The cluster II is

composed of two subcluster (II-I and II-II), comprising HIV-2 sequences belonging to individual residents in Havana. The subcluster II-I grouped the Cuban HIV-2 sequences of individuals who acquired their infection in Africa. The subcluster II-II included sequences corresponding to infected individuals in Cuba, which are related to sequences from different geographical origins, where in addition to Guinea Bissau, countries such as Portugal, Japan and India were included.

3.2. Age and Evolutionary Rate of the Population in Group A of HIV-2 in Cuba

The median TMRCA of the HIV-2 sequences involved in phylogenetic analysis was 1958 (95%, HPD=1929-1977) and the TMRCA Cuban sequences is close to 1972 (95%, HPD= 1966-1978) (Figure 2), time interval during which 30.8% of studied patients were stationed in endemic areas or with presence of HIV-2.

The μ *env* gene of HIV-2 sequences studied was 4.83×10^{-3} substitutions per site per year (95% HPD= 4.58 - 5.00×10^{-3} subs/ site / year).

4. Discussion

The accurate and timely laboratory diagnosis and epidemiological surveillance of circulating viral variants in HIV-positive population are essential components of the National HIV/AIDS Program of Cuba (16). The first HIV-2-infected individual was diagnosed in Cuba in 1987 and to date has detected the presence of this retrovirus in 22 people. A previous study by our working group, described the presence of group A in 13 HIV-2 infected individuals, representing so far 65% of those diagnosed by this retrovirus and indicates several independent introductions of HIV-2 into Cuba (18). These facts led to delve into the origin of HIV-2 in Cuba and analyze the behavior of evolutionary history, by employing a larger number of viral sequences from different geographical regions.

Although HIV-2 is less transmissible than HIV-1, the heterosexual transmission it is predominant in countries where the virus is endemic and non-endemic (22). In Cuba, the predominant mode of transmission in HIV is sexual, predominantly in the group of men who have sex with men (MSM) (23); however the results of this study are consistent with those described in other countries that have reported infected with HIV-2 individuals (4-6).

Grouping of Cuban sequences with viral sequences from regions where HIV-2 is endemic and non-endemic reflected in the formation of two large clusters on the phylogenetic tree. An infected patient in Guinea Bissau, which then transmitted the virus to their pair, is grouped in cluster I. The close phylogenetic relationship between the two Cuban sequences and from Guinea Bissau, reinforcing the information obtained through the epidemiological survey. Regarding the origin of this retrovirus epidemic in Guinea

Bissau, it has been hypothesized that the war of independence made in that country, a former Portuguese colony, in the period 1963-1974, was the epicenter of the beginning of the epidemic through the spread of this virus by the Portuguese army (11). Other factors played a role in the initiation epidemic during this period, as were the increase of access to blood transfusions not researched, the increased prostitution in the region, and the ritual of female circumcision and mass vaccination campaigns (10, 24, 25).

In the cluster II-I, Cuban sequences are grouped with sequences from Portugal and Guinea Bissau. Patients to whom belong these viral sequences reported to have been infected in Cape Verde and Angola, countries that maintain a constant political, economic, social and cultural exchange with Portugal, a country with a high incidence of HIV-2 (25-27).

The stay of a group of Cuban individuals in Africa in the late 70s and early 80s, reaffirm the estimated TMRCA Cuban sequences in this study. The spread of group A HIV-2 occurred in our country after the return of this group of patients grouped in cluster I and II with Guinea Bissau viral sequences obtained between 1987 and 2012.

One aspect to consider when analyzing the demographic

and epidemiological data from patients living in Havana, is the absence of epidemiological link between them, something that contradicts the results shown in the phylogenetic tree, all grouped in the subcluster II-II sequences of isolates of HIV-2 in non-endemic countries. Clustering sequences from Japan and France confirm the several independent introductions of this viral variant in Cuba, as suggested by Machado et al (18). These elements reaffirm the importance of finding and active investigation of contacts of individuals studied, as epidemiological tools that will clarify the above stated.

It has been suggested that high values of μ *env* gene of human retroviruses is that the proteins encoded by this gene are under the selective pressure of the immune system, which favors the conformational change of envelope proteins and thus evading the immune system virus in each cycle of viral replication (28). The result obtained in this study is high and close to that obtained by Lemey et al (29)

This study allowed us to deepen the dynamics of the HIV-2 in Cuba and estimate for the first evolutionary history of the retrovirus in our country, allowing the National HIV/AIDS Program of Cuba develop strategies aimed at prevention and treatment.

Table 1. Baseline Characteristics of Patients Included in the Study

Individuals	Diagnostic years	Isolation years	Sex	Sexual conduct	Location	Country of infection
12CU01	2011	2012	M	MSM	Havana	Angola
12CU02	2011	2012	F	HT	Havana	Cuba
12CU03	1997	2012	M	MSM	Havana	Cuba
12CU04	2004	2012	M	MSM	Havana	Cuba
12CU05	1997	2012	M	HT	Havana	Cabo Verde
12CU06	2003	2012	M	MSM	Havana	Cuba
12CU07	2005	2012	F	HT	Central region	Cuba
12CU09	2008	2012	F	HT	Havana	Cuba
12CU10	2001	2012	F	HT	Central region	Cuba
12CU11	1997	2012	M	HT	Havana	unknown
12CU14	2004	2012	F	HT	Eastern region	Cuba
12CU15	2004	2012	M	HT	Eastern region	Guinea Bissau
12CU16	1996	2012	F	HT	Eastern region	Cuba

Figure 1. Maximum likelihood (ML) tree of group A HIV-2 sequences. ML analysis was performed under TIM2+I+G nucleotide substitution model. All 13 Cuban sequences are shown in red. Numbers on branches indicate the degree of support for each node.

Figure 2. Bayesian phylogenetic tree of group A env gen HIV-2. Years are reported on the scale axis below. The colors shown the country or region of the sequences (red: Cuba, blue-Guinea Bissau, green-Ivory Coast, orange: Portugal, purple: India, olive green: Japan, black-others countries).

REFERENCES

- [1] Barin F, M'Boup S, Denis F, Kanki P, Allan JS, and Lee TH et al.: Serological evidence for virus related to simian T-lymphotropic retrovirus III in residents of West Africa. *Lancet* 1985; 2: 1387-1389.
- [2] Clavel F, Mansinho K, Chamaret S, Guetard D, Favier V, Nina J, San-tos-Ferreira MO, Champalimaud JL, Montagnier L:
Human immunodeficiency virus type2 infection associated with AIDS in West Africa. *N Engl J Med* 1987, 316:1180-1185.
- [3] Matheron S, Mendoza-Sassi G, Simon F, Olivares R, Coulaud JP, and Brun-Vezinet F: HIV-1 and HIV-2 AIDS in African patients living in Paris. *AIDS* 1997; 11: 934-936.
- [4] Dougan S, Patel B, Tosswill JH, and Sinka K. Diagnoses of HIV-1 and HIV-2 in England, Wales, and Northern Ireland associated with West Africa. *Sex Transm Infect* 2005; 81: 338-341.
- [5] Valadas E, Franca L, Sousa S, and Antunes F: 20 years of HIV-2 infection in Portugal: Trends and changes in epidemiology. *Clin Infect Dis* 2009; 48: 1166-1167.
- [6] D'Ettorre G, Lo Presti A, Gori C, Cella E, Bertoli A, Vullo V, et al. An HIV type 2 case series in Italy: A phylogenetic analysis. *AIDS Res Hum Retroviruses* 2013; 29 (9): 1254-9.
- [7] Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR. An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature*. 1989; 339:389-392.
- [8] Gao F, Yue L, White AT, Pappas PG, Barchue J, Hanson AP,

- et al. Human infection by genetically diverse SIVSM-related HIV-2 in west Africa. *Nature*. 1992; 358:495–499.
- [9] Peeters M, D'Arc M, Delaporte E. The origin and diversity of human retroviruses. *AIDS Rev* 2014; 16 (1): 23-34
- [10] Poulsen AG, Aaby P, Jensen H, Dias F. Risk factors for HIV-2 seropositivity among older people in Guinea-Bissau. A search for the early history of HIV-2 infection. *Scand J Infect Dis* 2000; 32:169–175.
- [11] de Silva T; van Tienne C, Onyango C; Jabang A, Vincent T, Schim van der Loeff M et al. Population dynamics of HIV-2 in rural West Africa: comparison with HIV-1 and ongoing transmission at the heart of the epidemic. *AIDS* 2013, 27: 125-134.
- [12] Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med*. 2011; 1:a006841. Review.
- [13] Ayoub A, Akoua-Koffi C, Calvignac-Spencer S, et al. Evidence for continuing cross-species transmission of SIVsmm to humans: characterization of a new HIV-2 lineage in rural Côte d'Ivoire. *AIDS*. 2013; 27:2488–2491.
- [14] Yamaguchi J, Vallari A, Ndambi R, Coffey R, NGansap C, Mbonya D, et al. HIV type 2 intergroup recombinant identified in Cameroon. *AIDS Res Hum Retroviruses* 2008; 24 (1): 86-91.
- [15] Ibe S, Yokomaku Y, Shiino T, Tanaka R, Hattori J, Fujisaki S et al. HIV-2 CRF01_AB: First circulating recombinant form of HIV-2. *J. Acquir Immune Defic Syndr* 2010; 54: 241-247.
- [16] Díaz HM, Pérez MT, Lubián AL, Nibot C, Cruz O, Silva E, et al. HIV Detection en Cuba: Role and Results of the National Laboratory Network. *MEDICC Review* 2011; 13 (2): 9-13.
- [17] Díaz DF, Ortiz E, Martín D, Nibot C, Rizo A, Silva E. HIV-2 antibody detection after indeterminate or negative HIV-1 western blot in Cuba, 2005-2008. *MEDICC Review* 2012; 14 (1): 25-29.
- [18] Machado LY, Díaz HM, Noa E, Martín D, Blanco M, Díaz DF, et al. Phylogenetic analysis of Human Immunodeficiency Virus type 2 isolated from Cuban individuals. *AIDS Res Hum Retroviruses* 2014; 30 (8): 823-826.
- [19] Hall TA. Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp Ser* 1999; 4: 95-98.
- [20] Posada D, Crandall K. ModelTest: testing the model of DNA substitution. *Bioinformatics* 1998; 14: 817–818.
- [21] Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 2007; 7: 214.
- [22] Schimvan der Laeff MF, Aaby P. Towards a better understanding of the epidemiology of HIV-2. *AIDS* 1999; 13: 869-84.
- [23] Gorrry C. Cuba's National HIV/AIDS Program. *MEDICC Review* 2011. 13(4): 5-8.
- [24] Jensen ML, Dave S, Schim van der Loeff M, da Costa C, Vincent T, Leligdowicz A, et al. Vaccinia scars associated with improved survival among adults in rural Guinea-Bissau. *PLoS One* 2006; 1: e101.
- [25] Pepin J, Plamondon M, Alves AC, Beaudet M, Labbe AC. Parenteral transmission during excision and treatment of tuberculosis and trypanosomiasis may be responsible for the HIV-2 epidemic in Guinea-Bissau. *AIDS* 2006; 20: 1303–1311.
- [26] Carvalho AC, Valadas E, Franca L, Carvalho C, Aleixo MJ, Méndez J et al. Population mobility and the changing epidemics of HIV-2 in Portugal. *HIV Med* 2012; 13 (4): 219-25.
- [27] Faria N, Hodges-Maameletzis Y, Silva JC, Rodés B, Erasmus S, Paolucci S et al. Phylo geographical footprint a colonial history in the global dispersal of HIV-2 group A. *J Gen Virol* 2012; 93: 889-99.
- [28] Freed EO, Mouland AJ. The cell biology of HIV-1 and other retroviruses. *Retrovirology* 2006; 3: 77
- [29] Lemey P, Pybus OG, Wang B, Saksena NK, Salemi M, Vandamme AM. Tracing the origin and history of HIV-2 epidemic. *PNAS* 2003; 100 (11): 6588-92.