Sequence Note:

Title: Phylogenetic analysis of Human T-Cell Lymphotropic Virus Type 1 isolated from Cuban individuals.

Running head: Subtypes HTLV-1 from Cuban individuals

Authors: Liuber Y Machado, Leonor M Navea, Héctor M Díaz, Madeline Blanco, Marta Dubed, Dania M Romay, Eladio Silva, Lester J Pérez.

AIDS Research Laboratory, Carretera de Tapaste and Autopista Nacional, San José de las Lajas, Mayabeque, Cuba, Postal Code 32700.


both authors equally contributed to this work.

*correspondence author e mail: liuberyans@infomed.sld.cu
Abstract

The presence of infection by Human T cell lymphotropic virus type 1 (HTLV-1) in Cuba has been previously documented. However, the genetic information on the strains that circulate in the Cuban people remains unknown. The present work constitutes the first study about phylogenetic relationship of HTLV-1 Cuban isolates. 12 Cuban patients who were diagnosed with HTLV-1 infection and different clinical manifestations were performed. The sequences 3’LTR were analyzed for the construction of a phylogenetic tree with reference sequences of HTLV-1 of different geographical origins. Phylogenetic analysis of 3’ LTR gene showed that all the Cuban samples clustered in Transcontinental subgroup of the Cosmopolitan subtype. The phylogenetic analysis suggests multiple introductions of HTLV-1 in Cuba and the African possible origin of the samples. The results of the study will reinforce the program of epidemic surveillance of the infection in Cuba.
The human T-cell lymphotropic virus type 1 (HTLV-1) was the first human onco-retrovirus isolated from a patient with a cutaneous T-cell lymphoma in 1980 (1). Actually it is considered the causal agent of adults T-cell leukemia/lymphoma (ATLL) and its implication in the etiology and pathogenesis of tropical spastic paraparesis/HTLV-1 associated myelopathy (TSP/HAM), uveitis and chronic infective dermatitis is well recognized (2). It has been calculated that 15 to 20 million people are infected around the world. Higher prevalence zones, considered as HTLV-1 infection endemic areas are the Southwest of Japan, the center and South of Africa, the Central America, the Southeast of the USA, and the Caribbean basin. Cases are also reported in other geographic areas from immigrants of endemic areas (3).

Four main HTLV-1 subtypes have been identified based in phylogenetic analysis of the 3´LTR gen: the Cosmopolitan (a), the Central African (b and d), and the Melanesian (c). The Cosmopolitan group is disseminated around the world and has been divided in five subgroups: Transcontinental (A), Japanese (B), West African/Caribbean (C), North African (D), and Black Peruvian (E) (4). In the Americas, it has been suggested that the Cosmopolitan group was introduced during the post- Columbus migrations due to the African slave trade and some Japanese migrants (5). In Cuba, the first evidence of HTLV-1 infection was found by serological surveillance carried out in different risk groups where an antibody seroprevalence of 0.037% was reported (6, 7) and later in a patient with clinical manifestations of ATLL diagnosed in 1990 (8).

Since 1990, a total of 40 individuals have been serologically confirmed as positive for antibodies against HTLV-1 at the AIDS Research Laboratory - the Cuban National Reference Laboratory for Human Retrovirus. Nevertheless, there are not available reports about the genetic information of circulating HTLV-1 strains in the Cuban population. The actual research was done in order to know the genetic diversity of HTLV-1 strains isolated from asymptomatic Cuban individuals and/or from Cuban patients with HTLV-1 associated diseases.
Between January 2010 and December 2011 a total of 12 HTLV-1-infected individuals were investigated and 10 mL EDTA-blood sample was taken, after the completion of a written informed consent and the recording of clinical and epidemiological data, including patient’s age, sex, origin and contagious way (Table 1). The infection source or acquisition mode was determined by the analysis of epidemiological evidences and the serologic investigation to family (parents, children, brothers and sisters) and sexual contacts. All sera samples were screened with an HTLV-1 antibodies immunoassay test (DAVIH HTLV-1, DAVIH Laboratories, Cuba) according of manufacturer’s directions. Positive samples were later confirmed using a Western blot assay (DAVIH-blot HTLV-1, DAVIH Laboratories, Cuba) according of manufacturer’s directions. Positive individuals were classified as asymptomatic carriers, TSP/HAM, and ATLL patients, according to the World Health Organization (WHO) directives (9).

DNA was extracted from whole blood by the midi spin columns from the High Pure Viral Nucleic Acid kit (Roche Diagnostic, IN, USA) following the manufacturer’s directions. Samples were subjected to an “in house” nested polymerase chain reaction (PCR) to amplify tax and pol genes. Amplification of the pol region was performed with outer primers SK-110/SK-111-I and inner primers pol 1.1/pol 3.1 and pol 1.2/pol 3.2 (10). Amplification of the tax region was carried out with outer primers SK-43-I/SK-44-I and inner primers SK-43/SK-44 (11). The PCR reactions were performed by use of Fast Start High Fidelity PCR System (Roche Diagnostics GmbH, Mannheim, Germany). The PCR products were 135 bp for pol and 128 bp for tax amplification. In order to perform the phylogenetic analysis, the 3’LTR region was amplified by heminested PCR by using 8200LA/3Vext as outer primers and 8200LA/3Vint as inner primers (528 bp, ATK-1 genome positions 8196-8699) (12). Sequencing in both directions was performed using the Genome Lab™ Dye Terminator Cycle Sequence with Quick Start kit following the manufacturer’s directions (Beckman Coulter, Inc, CA) and with the inner primers of the heminested PCR employed (12). Sequencing products were read on a CEQ™ 8800 genetic analyzer (Beckman Coulter, Inc, CA). Neighbor-joining (NJ) and maximum likelihood (ML) trees were generated by PAUP 4.0 (13). The TIM+G model
was selected as the best model for the phylogenetic analysis (alpha parameter=0.7966). The nucleotide model was inferred using Modeltest (14). The reliability of the NJ trees was assessed by analyzing 1000 bootstrap replicates. For ML trees, a heuristic search was performed with a subtree pruning regrafting branch swapping algorithm using the NJ tree as the starting material, including its optimized parameters. The likelihood ratio test (RT) method was used to calculate statistical support for the branches. Bootstrap values for the ML analysis were obtained with the PHYML program (15).

A number of 7 (58.3%) of the studied samples were from patients with HTLV-1 associated diseases. From them, five (41.6%) have developed HAM/TSP, one patient developed a cutaneous T-cell lymphoma and the other one uveitis (Table 1).

The 3’LTR region phylogenetic analysis allowed grouping the 12 patient’s nucleotide sequences in the Cosmopolitan subtype (Transcontinental (A) subgroup), with a bootstrap value of 73% (p≤0.001 for the ML analysis). The majority of our studied sequences (n=10) were grouped by their homology with sequences coming from French Guyana. The 11CUHT01 sample was grouped inside the Latin American cluster beta. The 11CUHT12 formed group with sequences coming from French Guyana and Chile (Figure 1).

Although the presence of HTLV-1 in Cuba was already documented (6-8), this report is about the first study of circulating HTLV-1-genetic variants in the Cuban seropositive population.

The presence of the **HTLV-1 subtype A subgroup C** in the Caribbean is supported by the introduction of this lineage from Western Africa during the slave trade (4). Nevertheless, circulation of the Transcontinental subgroup has been detected in the area, so it is not surprising that all Cuban strain sequences were found inside this subgroup. Their grouping in different individual clusters suggests that there were multiple introductions of the HTLV-1 in Cuba. Nevertheless, it is important to remark that the sample 11CUHT01, grouped together with sequences from Brazil in the Latin American cluster beta, was
isolated from a female patient who has had a sexual intercourse with an individual infected in Africa, according to the epidemiological surveillance data. Several phylogenetic studies carried out in some Brazilian populations have also shown association between the Latin American cluster beta with sequences from African origin (16, 17).

Inside the group of samples with sequences similar to the ones from French Guyana origin, the 11CUTH02 sample is from a patient who infected both 11CUTH03 and 11CUHT11 samples by sexual and vertical transmission, respectively (18, 19). The high degree of homology of the studied sequences in these three samples is an evidence of the same genetic variant transmission, in spite of the fact that there is a little distance between then, inside the same phylogenetic cluster.

The 11CUHT06 sample was obtained from a direct descendent of Haitians and he is the father of patients with the 11CUHT05 and CUHT09 samples. This result suggest that one of the possible origin of the HTLV-1 infection in Cuba is the one coming from the migration of people from Haiti, a country where TSP is frequent, and TSP has been used as a sentinel disease indicating a high prevalence of HTLV-1 infection. The facts of the Afro-Caribbean origin of part of the Cuban population (20) together with the molecular evidences of the arrival of HTLV-1 to the Americas coming from the African population introduced by the European colonizers (21, 22) are factors that can explain the origin of HTLV-1 in Cuba.

The 11CUHT12 sample is from a female patient born in Santiago de Cuba province who develops HAM/TSP and, it was grouped by homology with nucleotide sequences coming from the Latin American area (French Guyana, Chile and Brazil). The Santiago de Cuba province, with its Caribbean Sea coast is a know area used for the entrance of Haitians immigrants to Cuba, may reinforce the relation between this studied sample and the already mention Caribbean origin sequences samples.

The clinical presentation of HTLV-1 infection in the Caribbean differs from the observed in Japan and Sub-Saharan Africa. In the Caribbean HAM/TSP
predominates while in Japan is more frequent ATLL, with 86 new cases of the leukemia and only 3 patients with the myelopathy per 100,000 inhabitants per year (23). Our phylogenetic results have epidemiological importance and can have prognostic value, because it has been described that individuals infected with HTLV-1 Transcontinental subgroup A may be associated with a higher risk for HAM/TSP development compared with HTLV-1 subgroup B (24).

In conclusion, this paper is showing the presence of the subtype A Transcontinental subgroup A of HTLV-1 in the Island. Finally, the obtained results will be a contribution for the epidemiological surveillance program; they will reinforce the studies about the origin of HTLV-1 in Cuba and the detection of evolution patterns for the HTLV-1-circulating genetic variants in the country.

**Nucleotide sequence accession number:**

GenBank accession numbers for the sequences reported here are from JX194169-JK194179, and JX871882.

**Competing interests:**

No competing financial interests exist
References:


Figure 1. Maximum likelihood (ML) tree of 3’LTR HTLV-1 sequences. ML analysis was performed under TIM+G nucleotide substitution model. All 12 Cuban sequences are shown with symbol ♦. Numbers on branches indicate the
degree of support for each node. MEL5 was used as the out-group. Geographic origin is given in parentheses.
Table 1: Epidemiological and clinical data of HTLV-1 infected Cuban individuals

<table>
<thead>
<tr>
<th>Individuals</th>
<th>Age (years)</th>
<th>Gender</th>
<th>Province/Country</th>
<th>Geographic origin</th>
<th>Clinical status</th>
<th>Acquisition mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>11CUHT01</td>
<td>51</td>
<td>F</td>
<td>Havana/Cuba</td>
<td>Cuba</td>
<td>Asymptomatic</td>
<td>Sexual</td>
</tr>
<tr>
<td>11CUHT02</td>
<td>49</td>
<td>F</td>
<td>Havana/Cuba</td>
<td>Cuba</td>
<td>Asymptomatic</td>
<td>Sexual</td>
</tr>
<tr>
<td>11CUHT03</td>
<td>29</td>
<td>F</td>
<td>Havana/Cuba</td>
<td>Cuba</td>
<td>uveitis</td>
<td>Vertical</td>
</tr>
<tr>
<td>11CUHT04</td>
<td>39</td>
<td>M</td>
<td>Havana/Cuba</td>
<td>Cuba</td>
<td>Asymptomatic</td>
<td>Vertical</td>
</tr>
<tr>
<td>11CUHT05</td>
<td>48</td>
<td>M</td>
<td>Camagüey/Cuba</td>
<td>Cuba</td>
<td>TSP/HAM</td>
<td>Vertical</td>
</tr>
<tr>
<td>11CUHT06</td>
<td>83</td>
<td>M</td>
<td>Camagüey/Cuba</td>
<td>Haiti</td>
<td>TSP/HAM</td>
<td>Vertical</td>
</tr>
<tr>
<td>11CUHT07</td>
<td>56</td>
<td>F</td>
<td>Matanzas/Cuba</td>
<td>Cuba</td>
<td>Asymptomatic</td>
<td>Vertical</td>
</tr>
<tr>
<td>11CUHT08</td>
<td>43</td>
<td>F</td>
<td>Matanzas/Cuba</td>
<td>Cuba</td>
<td>TSP/HAM</td>
<td>Vertical</td>
</tr>
<tr>
<td>11CUHT09</td>
<td>40</td>
<td>F</td>
<td>Camagüey/Cuba</td>
<td>Cuba</td>
<td>TSP/HAM</td>
<td>Vertical</td>
</tr>
<tr>
<td>11CUHT10</td>
<td>51</td>
<td>F</td>
<td>Havana/Cuba</td>
<td>Cuba</td>
<td>ATLL</td>
<td>unknown</td>
</tr>
<tr>
<td>11CUHT11</td>
<td>63</td>
<td>M</td>
<td>Havana/Cuba</td>
<td>Cuba</td>
<td>Asymptomatic</td>
<td>Sexual</td>
</tr>
<tr>
<td>11CUHT12</td>
<td>41</td>
<td>F</td>
<td>Santiago de Cuba/Cuba</td>
<td>Cuba</td>
<td>TSP/HAM</td>
<td>unknown</td>
</tr>
</tbody>
</table>

F: female; M: male; TSP: tropical spastic paraparesis; ATLL: adult T cell leukemia/lymphoma