Sequence and Phylogenetic Analysis of Human T Cell Lymphotropic Virus Type 1 from Tumaco, Colombia

Norman Balcázar/[+], Gloria I Sánchez*, Felipe Garcia-Vallejo**

Cra 51 D # 62-29, Grupo de Inmunovirología *Grupo Infección y Cáncer, Facultad de Medicina, Universidad de Antioquia, A.A. 1226, Medellín, Colombia **Departamento de Ciencias Fisiológicas, Facultad de Salud, Universidad del Valle, AA 25360, Cali, Colombia

Human T cell lymphotropic virus type 1 (HTLV-1) is a retrovirus that causes leukemia and the neurological disorder HTLV-1 associated myelopathy or tropical spastic paraparesis (HAM/TSP). Infection with this virus – although it is distributed worldwide – is limited to certain endemic areas of the world. Despite its specific distribution and slow mutation rate, molecular epidemiology on this virus has been useful to follow the movements of human populations and routes of virus spread to different continents. In the present study, we analyzed the genetic variability of a region of the env *gene of isolates obtained from individuals of African origin that live on the Pacific coast of Colombia. Sequencing and comparison of the fragment with the same fragment from different HTLV-1 isolates showed a variability ranging from 0.8% to 1.2%. Phylogenetic studies permit us to include these isolates in the transcontinental subgroup A in which samples isolated from Brazil and Chile are also found. Further analyses will be necessary to determine if these isolates were recently introduced into the American continent or if they rather correspond to isolates introduced during the Paleolithic period.*

Key words*:* human T cell lymphotropic virus type 1 - HTLV-1 associated myelopathy/tropical spastic paraparesis - HTLV-1 phylogeny - Colombia

Human T cell leukemia virus type 1 (HTLV-1) of the Retroviridae family (Poiesz et al. 1980) is the etiological agent of adult T-cell leukemia (ATL; Hinuma et al. 1981) and of the HTLV-1 associated myelopathy (HAM), known also as tropical spastic paraparesis (HAM/TSP; Gessain et al. 1985, Osame et al. 1986). HAM/TSP is a chronic and progressive neuromyelopathy characterized by paraparesis and spasticity of the inferior members, and occasionally changes are observed in both the sensorial-superficial parts and the inferior part of the spine (Osame et al. 1986). Specific IgG antibodies against HTLV-1 have been detected in individuals from several Caribbean islands, South America, the Seychelles Islands, parts of Africa, the Southern part of Japan, Melanesia and Australia. It is considered that of the 15 to 25 million HTLV-1-infected individuals worldwide, 1 to 5% will develop ATL or HAM/TSP (Green & Chen. 2001). The first cases of TSP in Colombia were described in 1981 (Zaninovic et al. 1981) and their association with HTLV-1 in 1988 (Arango et al. 1988) with a prevalence of 0.098% in Tumaco, a town where most people have African ancestors (Zaninovic et al. 1988). Although genome variation among strains of HTLV-1 is very low, a few nucleotide changes seem to be specific of the geographical origin of the strains. However, the genomic sequence of HTLV-1 isolated from patients with ATL and/or HAM/TSP show neither meaningful biological nor biochemical differences (Daenke et al. 1990, Paine et al. 1991, Ratner et al. 1991).

To determine the characteristics of the HTLV-1 strains that infect individuals of the Pacific coast of Colombia, we report the sequences of the pro-virus HTLV-1 strains identified in individuals from Tumaco. Polymerase chain reaction (PCR) was used to amplify pro-viral segments that included parts of the *pol* and *env* genes, and these sequences were analyzed by restriction fragment length polymorphism (RFLP) using 10 different restriction enzymes. Furthermore, a segment of the *env* gene amplified from several of these isolates was sequenced and phylogenetic analyses of the sequences were conducted to determine their possible correlations to strains from different geographic regions of the world.

MATERIALS AND METHODS

HTLV-1 samples and PCR - The collection of clinical specimens and clinical diagnoses of HAM/TSP of volunteers from whom HTLV-1 positive peripheral mononuclear blood cells (PMBC) were collected, have been described (Arango et al. 1988). Briefly, 10 ml of heparinized blood was obtained from 10 individuals living in Tumaco who had been diagnosed with HTLV-1 or who clinically had been confirmed with TSP (Table). PMBC were Ficoll hypaque separated and the high molecular weight DNA was extracted by classical proteinase K digestion and phenol-chloroform extraction (Sambrook et al. 1989).

To identify the HTLV-1 positive samples, PCR was performed using the specific primers SK110(5'-CCCTACAATCCAACCAGCTCA – 3' 4733-4755) and SK111(5'-GTGGTGAAGCTACCATCGGGTTTA-3'4902-4922) (Fig. 1) that amplify a fragment of 189 bp of the *pol* gene of HTLV-1. The PCR mixture (50 µl) contained 200 ng of extracted DNA, 50 mM KCl, 10 mM Tris-Cl (pH 9.0), 0.01% Triton X-100, 2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 10 pmol of each primer and 2.5 units of Taq polymerase.

TABLE

Epidemiological and clinical status of human T cell lymphotropic virus type 1 (HTLV-1) infected individuals
of the Tumaco population

| Sample code | Age (yr)/Sex | Clinical status | Serological status | | | HTLV-1(+) by SK110/SK111 |
			RIA P24	ELISA	WB p21	PCR
TUM0072	46/M	TSP/HAM	+	+	+	+
TUM3687	48/M	HC	ND	ND	ND	+
TUM3033	42/M	TSP/HAM	-	-	-	-
TUM0310	48/F	HC	+	+	+	+
TUM0654	45/F	HC	-	+	+	+
TUM3071	45/M	HC	+	+	+	-
TUM3049	47/F	HC	+	+	+	+
TUM0112	46/F	HC	+	+	+	-
TUM0361	41/M	HC	+	+	+	+
TUM0385	48/M	HC	-	+	+	+

F: female; M: male; RIA: radioimmunoassay; ELISA: enzyme-linked immunosorbent assay; WB: western blot; ND: not determined;
HC: healthy carrier; serological assay used in Arango et al. 1988.

The PCR conditions were 94°C for 7 min, 40 cycles at 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min (Heneine et al. 1992). To confirm that the amplified fragments indeed corresponded to HTLV-1, they were transferred and hybridized with the $[\gamma\text{-}^{32}P]$ ATP labeled probe SK112 (5'-GTACTTTACTGACAAACCCGACCTAC – 3') (Fig. 1) that hybridizes to a specific region of the *pol* gene.

After confirming HTLV-1 infection, semi-nested PCR was carried out to amplify a fragment of 1,093 bp that spans the 3' end of the *pol* gene and the 5' end of the *env* gene. In the first round, oligonucleotides SK110 and *Env* B (-) (5'- TGGGCACTTTAAGGAACAAG- 3' 5846-5866) were used as external primers and in the second round oligonucleotides SK110 and EU3(-) (5'- GTTAACTGG ACAAGGGTCAGG- 3' 5806-5826) (Fig. 1) as internal primers. The PCR conditions were as described above, except that the annealing temperature was at 50°C instead of 55°C and the final step at 72°C was for 3 instead of 2 min. In the initial round, 100 ng of target DNA was used, and in the second round, 2 µl of the initial PCR reaction. For each reaction, DNA from an HTLV-1 seronegative individual as well as a sample with the complete PCR mixture but without DNA, and another with DNA obtained from the HTLV-1 chronically infected cellular line MT-2, were included as controls. Samples that where not amplified with any of the HTLV-1 primers were amplified with β-globin primers HG063 (5'- ACTCCCAGGAGCAGGGAGGG-3') and HG064 (5' - TGGTCTCCTTAAACCTGTCT-3') (Saiki et al. 1985) to determine the quality and quantity of the DNA. These oligonucleotides amplify a 249 bp fragment from nt -69 to 180 of the human β-globin gene. PCR conditions were similar to those used for the *pol* gene segment except that the annealing temperature was 60°C. In all cases, 10 µl of the amplified products were analyzed by electrophoresis in 1-2% agarose gels.
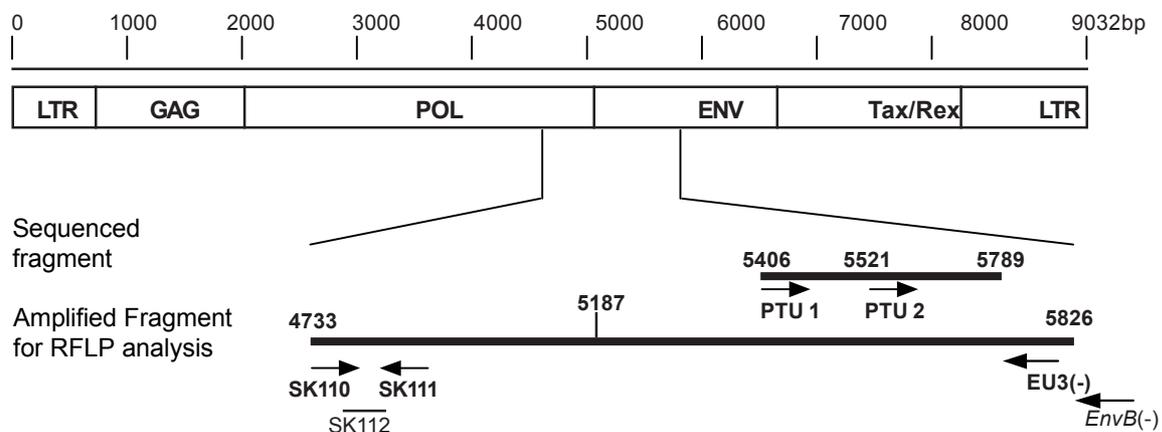

Fig. 1: schematic representation of the complete human T cell lymphotropic virus type 1 genome and the amplified semi-nested PCR fragment of 1093 bp. The 383 bp segment sequenced using primers PTU I and PTU 2 as well as the position of the SK112 probe are shown. Nucleotide positions correspond to the nucleotide sequence of the ATK-1 isolate (Accession nr J02029).

RFLP and southern blot analysis - For the RFLP analysis, the fragment amplified from the different isolates after processing of the seminested PCR was purified from agarose gels using the USB CLEAN Kit following the manufacturer's instructions (Amersham Pharmacia Biotech, UK). Digestions were carried out using 5 µl of the semi-nested PCR product in a final volume of 20 µl and 1 X of the appropriate enzyme reaction buffer with 10 units each of the following enzymes, Hae III, Taq I, Hind III, Pvu II, Bam HI, Dde I, Nco I, Bst NI, Sal I or Hae II (New England Biolabs, Beverly, MA). The reaction mixtures were incubated for 4-6 h at 37°C. The digested DNA fragments were electrophoretically separated in 2-3% agarose gels, transferred to nylon membranes (Hybond-N+, Amersham Pharmacia Biotech, Buckinghamshire, UK) and the membranes hybridized with the random primer-labeled plasmid pMT-2 which contains the complete sequence of the HTLV-1 provirus MT-2 (Miyoshi et al. 1981).

The membranes were pre-hybridized in 0.25 M Na-phosphate buffer pH 7.2, 1 mM EDTA, 1% BSA and 7% SDS at 65°C for 16 h. The filters were washed twice for 30 min at 65°C in 2 X SSC and 0.1% SDS followed by two additional washes in 0.2 X SSC and 0.1% SDS for the same time and temperature. Finally the filters were exposed for 4 h to X-omat X-ray films.

DNA sequencing and analysis - The semi-nested PCR products obtained from samples TUM0072, TUM0310, TUM0385, and TUM3049 that correspond to the segment of 1093 bp that spans the 3' end of the *pol* gene and the 5' end of the *env* gene (Fig. 1), were cloned into the pCR II™ cloning vector (Invitrogen, San Diego, CA). Positive clones were selected on Luria Broth (LB) agar plates colored with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and isopropyl-β-D-thiogalactopyranoside (IPTG). Three recombinant clones of each sample were isolated and a 383 bp segment from 5406 to 5789 of the proviral genomic sequence that corresponds to amino acid positions 77 to 203 of the *env* gene were sequenced by the dideoxy sequencing method using the Sequenase 2.0 Kit (Amersham Pharmacia Biotech Buckinghamshire, UK) and primers PTU 1 (5'- GTTCGTGGACCCTCGACC-3' 5406-5424) and PTU 2 (5'- CCTTGTTCCTTAA AGTGCC- 3' 5521-5540) as described (Soto-Ramirez et al. 1995; Fig. 1). These were aligned and compared with the sequence of the ATK-1 isolate.

Phylogenetic analysis - A phylogenetic analysis of the 383 bp sequenced segments was conducted. To this end the neighbour joining (NJ; Saitou & Nei 1987), the UPGMA (Nei 1991) and the maximum parsimony (MP; Fitch 1977) methods that accommodate variable rates of genetic changes, were used to construct unrooted phylogenetic trees. The methods used to determine the evolutionary distances between the isolated samples were the model of one parameter of Jukes-Cantor (Jukes & Cantor 1969) and the model of two parameters of Kimura (Kimura 1980). To verify the reliability of the final topologies of the trees, the bootstrap model of probabilities was used. All the above-mentioned procedures were carried out using the software package MEGA (molecular evolutionary genetics analysis, PA) version 1.01 (Kumar et al. 1993).

RESULTS

Identification of HTLV-1 - From seven (samples TUM0072, TUM0310, TUM0385, TUM3049, TUM3687, TUM0361 and TUM0654) out of the 10 samples taken from individuals clinically diagnosed with TSP and/or with positive serology, it was possible to amplify a fragment of the expected size of 189 bp with primers SK110 and SK111. The fragments amplified from these 7 samples hybridized with probe SK112 that specifically recognizes a fragment of the *pol* gene of HTLV-1. Taken together these results permit us to conclude that these 7 samples contain HTLV-1.

The DNA of the remaining three samples that showed no amplification with these primers was used as target in another PCR reaction using primers HG063 an HG064 that amplify a 249 bp fragment of the human β-globin gene. Although it was possible to amplify a fragment of the expected size from these samples (not shown), it was not possible to amplify any HTLV-1 fragment even when the primers that amplify fragments from the *gag* and *tax-rex* genes were included in the PCR reaction (not shown).

RFLP analysis - We first conducted an RFLP analysis on the 1093 bp fragment amplified from the semi-nested PCR from all seven samples that had hybridized to probe SK112.

Analysis of the digestion pattern of this fragment with the enzymes Hae III, Taq I, Hind III, Pvu II, Bam HI, Nco I, Bst NI, Sal I or Hae II showed a very similar digestion pattern of the fragment amplified from the 7 samples of this study as well as with the DNA amplified from the MT-2 cell line using the same set of primers (not shown).

Analysis of the digestion with Dde I showed two different patterns (Fig. 2). Fragments amplified from samples TUM0310, TUM0385, TUM3687, TUM0361 and TUM0654
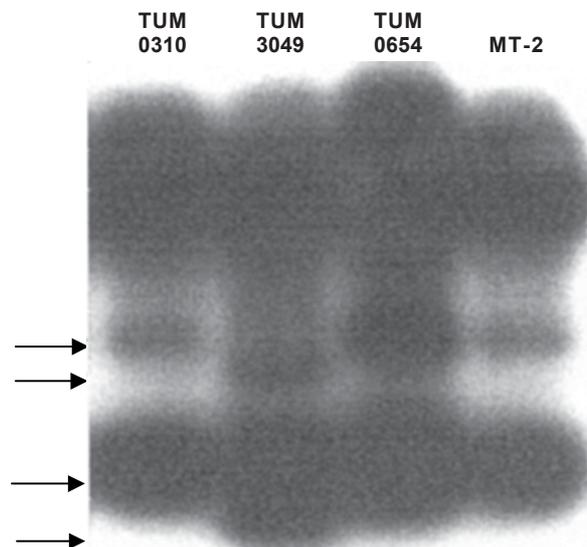

Fig. 2: restriction fragment length polymorphism analysis and southern blot detection showing the two different patterns obtained by digestion with Dde I. Analysis of fragments amplified from samples TUM0310, TUM3049, TUM0654 and MT-2 are shown. Arrows show bands of different sizes after digestion with Dde I and after hybridization with probe pMT-2.

and from MT-2 DNA, produced the same set of bands, whereas digestion of the fragments amplified from samples TUM0072 and TUM3049 similar pattern but different from the former five samples as well as from the fragment amplified from MT-2 DNA.

Sequence analysis - We therefore sequenced a 383 bp segment (Fig. 1) from samples TUM0072 and TUM3049 and from two samples TUM0310 and TUM0385 chosen at random which represent the different RFLP pattern. To this effect, the amplified fragment from these samples was cloned into the pCR II cloning vector and three clones of each fragment from the same sample were sequenced. Sequences of the clones obtained from the same sample showed complete agreement showing that the amplified fragments were composed of a homogeneous population of PCR products, and the changes observed were consequently not artifacts, but reflect changes in the viral genomic sequences. The nucleotide sequences were aligned and compared with the sequences of the HTLV-1 strains from various geographic regions including Japan (ATK-1, H5, TSP-1, MT-2; Seiki et al. 1983, Gray et al. 1990), the Caribbean area (HS-35, CH; Malik et al. 1988, Ratner et al. 1991), Brazil (pt-8; Schultz et al. 1991), Romania (H990; Schultz et al. 1991), Melanesia (MEL-1; Gessain et al. 1993), the United States (SP; Paine et al. 1991), a variant from Gabon (GeneBank Accession number L33266; Moynet et al. 1995), Chile (ST; Dekaban et al. 1992) and Zaire (EL; Paine et al. 1991, Ratner et al. 1991), as well as the sequences of Simian T-cell lymphotropic virus type I (STLV-I) strain ptM3 from a pig-tailed macaque (*Macaca nemestrina*) from Indonesia (Watanabe et al. 1985) and sequences of HTLV-1 strain Mo from the United States (Shimotohono et al. 1985)**.**

Aligment of the sequences of the 383 bp segment of the Japanese HTLV-1 strain ATK-1 (Fig. 3), with sequences of the strains from Tumaco revealed a 2.3 to 2.6% variability. The variability between the sequences obtained from the Tumaco strains was 0.5%. The amino acid sequences deduced for this fragment showed that there are changes in amino acid 89 (Thr→Ile, all Tumaco strains), 137 (His→Gln, all Tumaco strains), 166 (Asp →Gly, strain TUM 3049), 168 (Pro→Gln, strain TUM0385), 169 (Gly→Glu, strain TUM0310), 194 (Ser→Tyr, strain TUM0072), when compared to the ATK-1 strain.

Despite this overall low degree of amino acid sequence dissimilarity, the amino-terminal neutralizing domain, comprising amino acids 88 to 98 on the external envelope glycoprotein gp46, was totally conserved between the HTLV-1 Tumaco strain sequences examined here.

Phylogenetic analysis - The phylogenetic trees constructed using the sequences of the 383 bp fragment of the *env* gene, were very similar whichever of the three methods was used (UPGMA, NJ, or MP) with very similar evolutionary distances estimated either by the Jukes-Cantor or the Kimura method. Altough similar, we chose the NJ method because it gave the most representative phylogenetic tree. As shown in Fig. 4, three main groups were clearly identified. In the first group appeared strain ptM3 of HTLV-1 that represents an early ancestor from which the Melanesian and cosmopolitan strains diverged,

whereas in the second group appeared the Melanesian MEL-1 strain. The third group or cosmopolitan subgroup appeared clearly separated from the simian and Melanesian groups and contained the Tumaco strains described here.

Isolates TUM0310, TUM3049 and TUM0072 from Tumaco and isolate pt-8 from Brazil form a subgroup with a bootstrap value of 52%. Although this significance is low and does not permit us to conclude that these isolates form a completely separate group, they nevertheless form a subgroup. Our results suggest that these Colombian isolates can be included in the same group as other South American isolates such as CH from the Caribbean area and pt-8 from Brazil. Although the analysis shows that the strains from Japan and Africa do not appear in the same branch as the strains described here from Tumaco, they do not form a completely separate subgroup.

DISCUSSION

A strong correlation exists between the phylogenetic HTLV-1 genotypes and the geographic origin of virus carriers, although it has so far not been possible to determine if a link exists between strain variability and disease outcome. Additionally, molecular epidemiology studies of HTLV-1 strains from different ethnic groups have permitted the identification of the time of introduction of strains into different continents along with the migration of human populations. These facts validate the analysis of the genetic variability of HTLV-1 strains. Here we report the partial sequences of a fragment encompassing the genes *env* and *pol* of HTLV-1 isolates of individuals of Tumaco, a town on the Pacific Coast of Colombia, where most inhabitants are of African origin. The level of antibodies against HTLV-1 in these individuals is high compared to that in populations from other regions of the world and the prevalence of TSP is one of the highest in the world (Arango et al. 1990).

HTLV-1 infection in the 10 samples collected from individuals who had been clinically diagnosed with HAM/TSP, or healthy carriers that were identified as positive for HTLV-1 by three different serology methods were studied for proviral HTLV-1 presence. Seven samples were identified as carrying the HTLV-1 provirus by PCR and hybridization. Samples TUM3071 and TUM0112 were positive with all three serology methods but were negative by PCR and hybridization. Since the sensitivity of the test was not determined to detect HTLV-1, we conclude that they are indeed from HTLV-1 infected individuals, but they were not detected in our PCR experimental conditions. The sample from one individual (TUM3033) clinically diagnosed with TSP was negative by either the search for anti-HTLV-1 antibodies (Arango et al. 1988) or by the PCR and hybridization methods. Therefore his/her clinical state should be associated with an etiology different from infection with HTLV-1.

The RFLP analysis was conducted with several restriction enzymes indicated that all but Dde I produced a similar pattern of digestion on samples from Tumaco, the isolate ATK-1 (GeneBank accession numbers J02029) and the HTLV-1 infected cell line MT-2. The digestion pattern

Fig. 3: aligment and comparison of nucleotide sequences and the deduced amino acid sequences of the 383 bp fragment of the *env* gene (nucleotides 5428 to 5811; amino acids 77 to 203 of gp46) from Tumaco strains. Dots indicate sequence identity with HTLV-1 ATK-1. GeneBank accession nr: U66294 for TUM0072; U66296 for TUM0385; U66295 for TUM0310; U66297 for TUM3049.

observed with Dde I led us to initially infer the presence of variation in the sequence which was specific of certain isolated samples from Tumaco.

Of the HTLV-1 structural genes, the envelope gene seems to be subject to the greatest variablity (Malik et al. 1988, Gray et al. 1990, Shirabe et al. 1990). For this reason the present study focused on a segment of 383 bp of the *env* gene, to study the genetic variability in four of the isolated samples (TUM0072, TUM0310, TUM0385 and TUM3049) since such a segment contains the sequences that code for immunodominant epitopes on gp46 (Lipka et al. 1990). We, and others (Gray et al. 1990, Schultz et al. 1991, Paine et al. 1991) have observed C → T and G → A changes at nucleotide position 5468 and 5470 respectively, when compared to the ATK-1 isolate; this could be due to a mistake in the sequencing process of the ATK-1 proto-type strain.

When the sequences of the reference strains and sequences from isolates of this study were aligned with other HTLV-1 sequences of geographically distant strains, they showed a low degree of variability when compared to the ATK-1 prototype strain. Homology variability of the isolates was: ST (Chile) 0.8%, pt-8 (Brazil) 0.8%, ATK-1, MT-2, TSP-1, H-5 (Japan) 1.2– 2.3% and CH (Caribbean) 1.4%.

To verify the reliability of the topology of the phylogenetic trees constructed, three different methods, UPGMA, NJ and MP were used. From this analysis it is possible to observe that isolates TUM0310, TUM3049, TUM0072, TUM0385 and pt-8 from Brazil are closely related and that the strains examined here belong to the Cosmopolitan group. In this analysis is also possible to observe that the Japanese strains MT-2, TSP-1 as well as strains EL from Zaire, a variant from Gabon, and HS-35 from the Caribbean area locate to a different group. However, due to the low Bootstrap percentages, they cannot be considered as a monophyletic clade.

Several lines of evidence have suggested that HTLV-1 was introduced into the American continent together with the first human population which came from the East through the Bering Strait, and migrated to the rest of the continent. Phylogenetic analyses of mitochondrial DNA (Horai et al. 1993) and HLA haplotype sequence comparisons (Sonoda et al. 1991) also suggest a relationship between Japanese and paleo-indian people from South America (Miura et al. 1994). In contrast, other researchers have suggested that the American isolates originated from African isolates based on the absence of HTLV-1 among Amerindians and the high prevalence of infection among individuals of African descendents in the Caribbean area and in South America (Song et al. 1995). It is possible that the introduction of HTLV-1 could have occurred by African immigrants as suggested by Saksena et al. (1992) and

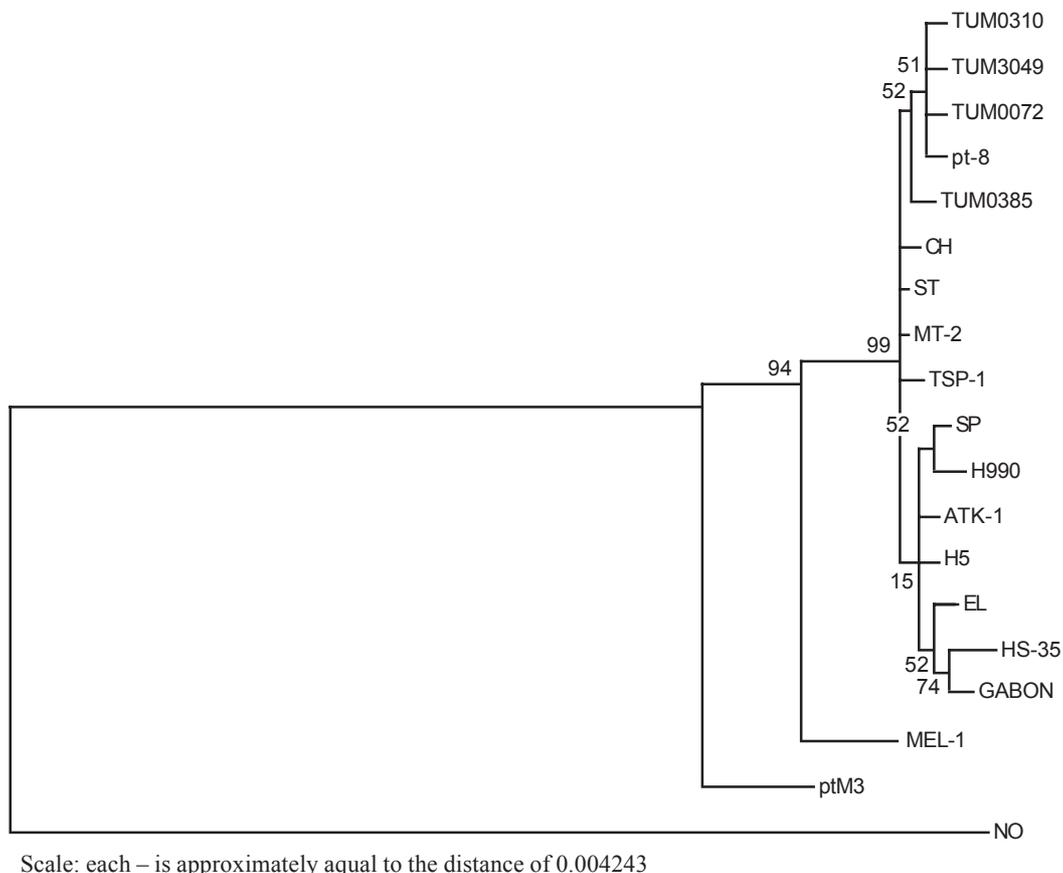Scale: each – is approximately aqual to the distance of 0.004243

Fig. 4: phylogenetic tree generated by the NJ method with a 383 bp fragment encompassing the amino terminus of gp46 for the four new sequences of human T cell lymphotropic virus type 1 (HTVL-1) strains from Tumaco and for other HTLV-1 sequences from isolates worldwide. The HTLV-2 Mo sequence was used as an out-group to root the tree. The numbers at some nodes or branches (bootstrap values) indicate frequencies of occurrence for 500 bootstrap samples.

Gessain et al. (1992). These authors suggest that HTLV-1 originated in an Indian Malayan region, from there presumably spread to Africa by marines thousand of years ago, and that later due to the slave trade in the XVI century, it spread to the New World and Japan. An analysis of the LTR region of HTLV-1 isolates from different ethnic groups of Peru have suggested that the virus may have been introduced into the American continent on a number of occasions from several African lineages, probably during and after the slave trade or through recent Japanese immigrations (Van Dooren et al. 1998). Phylogenetic analyses of the LTR region in the virus isolates examined here as well as in strains from different ethnic groups in South America will make it possible to confirm the origin of the HTLV-1 strains in these populations and to determine whether or not crossing between races has led to the introduction of viral strains among Indians and African descendents of the American continent.

It is predictable, based on the low divergence of the *env* region found in the Tumaco patients analyzed in this study as well as the low divergence of this region found by other authors in other parts of Latin America, that HTLV-1 isolates have evolved in the American population in different ethnic groups separately. However, analysis of the LTR region of HTLV-1 isolates from people of diverse ethnical origin will clarify this hypothesis.

REFERENCES

Arango C, Concha M, Trujillo JM, Biojo R 1990. Tropical spastic paraparesis associated risk factors in Tumaco, Colombia. In WA Blattner, *Human Retrovirology: HTLV,* Raven Press Ltd., New York, p. 337-383.

Arango C, Concha M, Zaninovic V, Corral R, Biojo R, Borrero I, Rodgers-Johnson P, Mora C, Garruto RM, Gibbs CJ 1988. Epidemiology of tropical spastic paraparesis in Colombia and associated HTLV-I infection. *Ann Neur 23*: S161-S165.

Daenke S, Nightingale S, Cruickshank JK, Banghan CR 1990. Sequence in variants of human T-cell lymphotropic virus type I from patients with tropical spastic paraparesis and adult T-cell leukemia do not distinguish neurologica1 from leukemia isolates. *J Virol 64*: 1278-1282.

Dekaban GA, King EE, Waters D, Rice GP 1992. Nucleotide sequences of an HTLV-1 isolate from a Chilean patient with HAM/TSP. *AIDS Res Hum Retroviruses 8*: 1201-1207.

Fitch WM 1977. On the problem of discovery of the most parsimonious tree. *Am Natur 111*: 223.

Gessain A, Barin F, Vernant JC, Gout O, Maurs L, Calender A, de The G. 1985. Antibodies to human T-lymphotropic virus type-1 in patients with tropical spastic paraparesis. *Lancet 2*: 407-410

Gessain A, Gallo RC, Franchini G 1992. Low degree of human T cell leukemia/lymphoma virus type I genetic drift in vivo as a means of monitoring viral transmission and movement of ancient human populations. *J Virol 66*: 2288-2295.

Gessain A, Boeri E, Yanagihara R, Gallo RC, Franchini G 1993. Complete nucleotide sequence of a highly divergent human T cell leukemia (lymphotropic) virus type I (HTLV-1) variant from Melanesia: genetic and phylogenetic relationship to HTLV-I strains from other geographical regions. *J Virol 67*: 1015-1023.

Gray GS, White M, Bartman T, Mann D 1990. Envelope gene sequence of HTLV-I isolate MT-2 and its comparison with other HTLV-I isolates. *Virology 177*: 391-395.

Green PL, Chen SY 2001. Human T-cell leukemia virus types 1 and 2. In DM Knipe, PM Howley (eds), *Fields Virology*, 4th ed., Lippincott Williams & Wilkins, Philadelphia, PA.

Heneine W, Khabbaz RF, Lal RB, Kaplan JE 1992. Sensitive and specific polymerase chain reaction assays for diagnosis of human T-cell lymphotropic virus type I (HTLV-I) and HTLV-II infections in HTLV-I/II seropositive individuals. *J Clin Microbiol 30*: 1605-1607.

Hinuma Y, Nagata K, Hanoaka M, Nakai M, Matsumoto T, Kinoshita KI, Shirakawa S, Miyoshi I 1981. Adult T-cell leukemia: antigen in an ATL cell line and detection of antibodies to the antigen in human sera. *Proc Natl Acad Sci USA 78*: 6476-6480.

Horai S, Kondo R, Nakawa-Hattori Y, Hayashi S, Sonoda S, Tajima K 1993. Peopling of the Americas, founded by four major lineages of mitochondrial DNA. *Mol Biol Evol 10*: 23-47.

Jukes TH, Cantor RC 1969. Evolution of protein molecules. In HD Munro, *Mammalian Protein Metabolism,* Academic Press, New York, p. 82-102.

Kimura M 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J Mol Evol 16*: 111-120.

Kumar S, Tamura K, Nei M 1993. Molecular evolutionary genetics analysis (MEGA) Version 1.01. Software. Institute of Molecular Evolutionary Genetics. The Pennsylvania State University, USA.

Lipka JJ, Bui K, Reyes GR, Moeckli R, Wiktor SZ, Blattner WA, Murphy EL, Shaw GM, Hanson CV, Sninsky JJ 1990. Determination of a unique and immunodominant epitope of human T cell lymphotropic virus type I. *J Infect Dis 162*: 353-357.

Malik KTA, Even J, Karpas A 1988. Molecular cloning and complete nucleotide sequence of an adult T-cell leukemia virus/human T-cell leukemia virus type (ATLV/HTLV-I) isolate of Caribbean origin: relationship to other members of the ATLV/HTLV-I subgroup. *J Gen Virol 69*: 1965-1710.

Miyoshi I, Kubonishi I, Yoshimoto S, Akagi T, Ohtsuki Y, Shiraishi Y, Nagata K, Inhuma Y 1981. Type C virus particles in a cord T-cell line derived by co-cultivating normal human cord leukocytes and T human leukaemic T cell. *Nature 294*: 770-771.

Miura T, Fukunaga T, Igarashi T, Yamashita M, Ido E, Funahashi S, Ishida T, Washio K, Ueda S, Hashimoto K, Yoshida M, Osame M, Singhal BS, Zaninovic V, Cartier L, Sonoda L, Tajima K, Ina Y, Gojobori T, Hayami M 1994. Phylogenetic subtypes of human T-lymphotropic virus type I and their relations to the anthropological background. *Proc Natl Acad Sci USA 91*: 1124-1127.

Moynet D, Cosnefroy JY, Bedjabaga I, Roelants G, Georges-Courbot MC, Guillemain B 1995. Identification of new genetic subtypes of human T cell leukemia virus type I in Gabon from encoding sequence of surface envelope glycoprotein. *AIDS Res Hum Retroviruses 11*: 1407-1411.

Nei M 1991. Molecular population genetics and evolution. North-Holland, Amsterdam. In W Li, D Graur (eds), *Fundamentals of Molecular Evolution,* Sinauer Associates Inc. Publishers, Sunderland, MA.

Osame M, Usuku M, Izumo S, Ijichi N, Amitani H, Igata A,

Matsumoto M, Tara M 1986. HTLV-l-associated myelopathy: a new clinical entity. *Lancet 1*: 1031-1032.

Paine E, Garcia J, Philpott TC, Shaw G, Ratner L 1991. Limited sequence variation in human T-lymphotropic virus type 1 isolates from North American and African patients. *Virology 182*: 111-123.

Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC 1980. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patients with cutaneous T-cell lymphoma. *Proc Natl Acad Sci USA 77*: 7415-7419.

Ratner L, Philpott T, Trowbridge DB 1991. Nucleotide sequence analysis of isolates of human T-lymphotropic virus type 1 of diverse geographical origins. *AIDS Res Hum Retroviruses 7*: 923-941.

Saiki R, Sharf SJ, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N 1985. Enzymatic amplification of β-globulin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science 230*: 1350-1354.

Saitou N, Nei M 1987. The neighbor-joining method: a new method for reconstructing phylogenetics trees. *Mol Biol Evol 4*: 406-425.

Saksena NK, Sherma MP, Yanagihara R, Dube DK, Poiesz BJ 1992. LTR sequences and phylogenetic analyses of a newly discovered variant of HTLV-I isolated from the Hagahai of Papua New Guinea. *Virology 189*: 1-9.

Sambrook J, Fritsch EF, Maniatis T 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schultz TF, Calabro ML, Hoad JG, Carrington CV, Matutes E, Catovsky D, Weiss RA 1991. HTLV-I envelope sequences from Brazil, the Caribbean, and Romania: clustering of sequences according to geographic origin and variability in an antibody epitope. *Virology 184*: 483-491.

Seiki M, Hattori S, Hirayama Y, Yoshida M 1983. Human adult T-cell leukemia virus: complete nucleotide sequences of the provirus genome integrated in leukemia cell DNA. *Proc Natl Acad Sci USA 80*: 3618-3622.

Shimotohono K, Takahashi Y, Shimizu N, Gojobori T, Golde DW, Chen IS, Miwa M, Sugimura T 1985. Complete nucleotide sequence of an infectious clone of human T-cell leukemia virus type II: an open reading frame for the protease gene. *Proc Natl Acad Sci USA 82*: 3101-3105.

Shirabe S, Nakamura T, Tsujihata M, Nagataki S, Seiki M, Yoshida M 1990. Retrovirus from human T-cell leukemia virus type I-associated myelopathy is the same strain as a prototype human T-cell leukemia virus I. *Arch Neurol 47*: 1258-1260.

Song KJ, Nerurkar VR, Pereira-Cortez AJ, Yamamoto M, Taguchi H, Miyoshi I, Yanagihara R 1995. Sequence and phylogenetic analyses of human T-cell lymphotropic virus type 1 from a Brazilian woman with adult T-cell leukemia: comparison with virus strains from South America and the Caribbean basin. *Am J Trop Med Hyg 52*: 101-108.

Sonoda S, Arce-Gomez B, Satz ML, Gorodezky G, Juarez V, Olivo O, Debaz A, Ferreira E 1992. Ethnic report on native Americans in South America and México. In K Tsuji, M Aizawa, T Sasazuki (eds) *HLA 1991*, Oxford Univ. Press, Oxford, p. 685-688.

Soto-Ramirez LE, Garcia-Vallejo F, Rengifo B, Vergara A, Borrero I, Marlink R, Essex M 1995. Human T-lymphotropic virus type I (HTLV-I) specific antibodies and cell-free RNA in crevicular fluid-rich saliva from patients with tropical spastic paraparesis/HTLV-I-associated myelopathy. *Viral Immunol 8*: 141-143.

Van Dooren S, Gotuzzo E, Salemi M, Watts D, Audenaert E, Duwe S, Ellerbrok H, Grassmann R, Hagelberg E, Desmyter J, Vandamme AM 1998. Evidence for a post-Columbia introduction of human T-cell lymphotropic virus Latin America. *J Gen Virol 79*: 2695-2708.

Watanabe T, Seiki M, Tsujimoto H, Miyoshi I, Hayami M, Yoshida M 1985. Sequence homology of the simian retrovirus genome with human T-cell leukemia virus-type I. *Virology 144*: 59-65.

Zaninovic V, Biojo R, Barreto P 1981. Paraparesia espastica del pacifico. *Colombia Med 12*: 111-117.

Zaninovic V, Arango C, Biojo R, Mora C, Rodgers-Johnson P, Concha M, Corral R, Barreto P, Borrero I, Garruto RM 1988. Tropical spastic paraparesis in Colombia. *Ann Neurol 23*: SI27-S132.