Identification and Characterization of a New and Distinct Molecular Subtype of Human T-Cell Lymphotropic Virus Type 2

NOBUTAKA EIRAKU,¹ PATRICIA NOVOA,² MARIZETE DA COSTA FERREIRA,² CLAUDE MONKEN,³
RICARDO ISHAK,⁴ ORLANDO DA COSTA FERREIRA,² SHI WEI ZHU,¹ ROSEMARIE LORENCO,²
MARLUISA ISHAK,⁴ VANIA AZVEDO,⁴ JOAO GUERREIRO,⁴ MARIA POMBO DE OLIVEIRA,⁵
PAULA LOUREIRO,⁶ NELSON HAMMERSCHLAK,⁷ SHINJI IJICHI,⁸ AND WILLIAM W. HALL^{1*}

Laboratory of Medical Virology, The Rockefeller University, New York, New York 10021¹; Emilo Ribas Institute of Infectious Diseases² and Hospital Israelita Albert Einstein,⁷ Sao Paulo, Federal University of Para, Belem,⁴ National Cancer Institute, Rio de Janeiro,⁵ and Hemope, Recife,⁶ Brazil; Department of Medicine, Thomas Jefferson University, Philadelphia, Pennsylvania 19107³; and Department of Internal Medicine, Kagoshima University, Kagoshima, Japan⁸

Received 16 August 1995/Accepted 28 November 1995

Molecular studies have demonstrated the existence of at least two major subtypes of human T-cell lymphotropic virus type 2 (HTLV-2), designated HTLV-2a and HTLV-2b. To further investigate the heterogeneity of this family of viruses, we have characterized the HTLV-2 subtypes present in several urban areas in Brazil. DNAs from peripheral blood mononuclear cells of a large number of infected individuals, the majority of whom were intravenous drug abusers, were analyzed by using PCR with restriction fragment length polymorphism and nucleotide sequencing analysis. Restriction fragment length polymorphism analysis of the *env* region suggested that all individuals were infected with the HTLV-2a subtype, and this was confirmed by nucleotide sequence analysis. In contrast, nucleotide sequence analysis of the long terminal repeat demonstrated that although the viruses were more related to the HTLV-2a than to the HTLV-2b subtype, they clustered in a distinct phylogenetic group, suggesting that they may represent a new and distinct molecular subtype of HTLV-2. This conclusion was supported by nucleotide sequence analysis of the pX region, which demonstrated that the Tax proteins of the Brazilian viruses differed from that of prototype HTLV-2a isolates but were more similar to that of HTLV-2b in that they would be expected to have an additional 25 amino acids at the carboxy terminus. In transient expression assays, the extended Tax proteins were found to be much more potent transactivators of the virus long terminal repeat than the Tax protein of the prototype HTLV-2a subtype. The studies suggest that the Brazilian viruses analyzed in this study, while being phylogenetically related to the prototypic HTLV-2a seen in North America, are phenotypically more related to HTLV-2b and can be justifiably classified as a new molecular subtype, which has been tentatively designated HTLV-2c.

Human T-cell lymphotropic virus type 1 (HTLV-1) and type 2 (HTLV-2) are closely related mammalian retroviruses which share a number of biological properties and have a tropism for T lymphocytes (22, 27, 44, 54). HTLV-2 infection is endemic in a number of American Paleo-Indian populations (1, 2, 4, 6, 7, 10, 14, 25, 26, 29, 33, 35, 37, 41, 43, 51, 56), and high rates of infection occur in intravenous drug abusers (IVDAs) in parts of North America, Europe, and Southeast Asia (13, 22, 44, 57). A number of studies of the viruses present in these populations have demonstrated that there are two major molecular subtypes of HTLV-2, which have been designated HTLV-2a and HTLV-2b (5, 8, 22, 26, 44, 53). Whereas HTLV-2a has been clearly shown to be the predominant infection in IVDAs in urban areas of North America (8, 44), HTLV-2b predominates in the Paleo-Indian groups (5, 22, 41, 44, 52). To date, HTLV-2b has been shown to be the exclusive infection in the Guaymi of Panama (43), the Wayu of Colombia (29, 51), and the Toba and Mataco groups of Argentina (10). In contrast, studies of Indian populations in North America have shown that whereas HTLV-2b infection appears to predominate, a number have infection with both subtypes (26). It seems possible that these populations have endemic infection with one subtype and that the second has only recently been introduced

by interactions with individuals who had other risk factors for infection.

In contrast to the studies on the aforementioned Indian groups, we have recently demonstrated that the Kayapo Indians, an indigenous population of Amazon region of Brazil, are infected with a distinct variant of the HTLV-2a subtype (30). Unfortunately, because of the small sample number and the limited molecular analysis, it was impossible to determine if this virus should be considered a distinct and unique molecular subtype. In the present study, we have attempted to characterize HTLV-2 infections in urban areas of Brazil and to compare these viruses with those present in the Kayapo Indian group. This effort has involved detailed nucleotide sequence analysis of three regions of the provirus, the *env*, long terminal repeat (LTR), and pX regions. The studies have demonstrated that the viruses in the urban areas are closely related to that present in the Kayapo Indians but that they are phylogenetically distinct from the HTLV-2a and HTLV-2b subtypes present in North America and Europe. This finding together with preliminary analysis of their phenotypic properties supports the view that these isolates represent and can be justifiably classified as a new and distinct molecular subtype of HTLV-2.

MATERIALS AND METHODS

* Corresponding author. Mailing address: Laboratory of Medical Virology, The Rockefeller University, 1230 York Ave., New York, NY 10021. Phone: (212) 327-8581. Fax: (212) 327-8582. Electronic mail address: Hall@rockvax.rockefeller.edu.

Study population. Twenty-eight individuals from Sao Paulo ($n = 24$), Rio de Janeiro ($n = 2$), and Recife ($n = 2$), 27 of whom were intravenous drug abusers (IVDAs), were identified as being seropositive for HTLV-2 by enzyme-linked immunosorbent assay (ELISA) and confirmatory Western blot (immunoblot)

analysis; 26 of the latter had concomitant human immunodeficiency virus infection. Seropositive individuals were identified by initial screening for antibodies to HTLV-1 and/or HTLV-2 by ELISA (Coulter Immunology, Hialeah, Fla.). Reactive samples were analyzed by Western blotting (HTLV Blot 2.3; Diagnostic Biotechnology, Singapore) to confirm seropositivity and to differentiate HTLV-1 and HTLV-2 infections. The origins of the samples obtained from the Kayapo Indian group have been previously described (30).

Peripheral blood mononuclear cells from infected individuals were separated by Ficoll-Hypaque centrifugation, and DNA was extracted by standard phenol-chloroform extraction methods. Amplification of the *env* region encoding the transmembrane glycoprotein gp21 for restriction fragment length polymorphism (RFLP) and nucleotide sequence analysis was carried out by using nested PCR with methods previously described (23, 30). For amplification and nucleotide sequencing analysis of the LTR, nested PCR was carried out as follows. Reactions were performed in 100- μ l volumes containing 1 μ g of extracted DNA, 225 μ M each deoxynucleoside triphosphate, 100 pmol of each primer, 50 mM KCl, 2.0 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), and 0.5 U of *Taq* polymerase (Perkin-Elmer Cetus Corp., Norwalk, Conn.). Primer pairs used in the first round were 5'-CCTTACCCACTTCCCTAGCGCTGA-3' and 5'-GGGAAAGCCCGTGATTTGCCCAT-3', corresponding to nucleotides (nt) 194 to 218 and 807 to 831 of the HTLV-2-MoT isolate. The second round was carried out with the nested primers 5'-AAAAGCGCAAGGACAGTTCAGGAGG-3' and 5'-ATCCGGACGAGAGAGCCCCACTTGT-3', corresponding to nt 290 to 314 and 759 to 783. Five microliters of the initial amplified product was used in the nested reaction. For RFLP analysis of the LTR, PCR was used to generate a 553-bp amplified product. After the first-round PCR as just described, a second primer pair, 5'-TGACGATTACCCCTGCCATAAA-3' and 5'-ATCCCGGACGAGCCCCACTTGT-3', corresponding to nt 231 to 254 and 759 to 783 of the HTLV-2-MoT isolate, was used in the reaction to produce the nested product. The synthesis of this LTR product, which is larger than that used in the nucleotide sequence analysis, was necessary to achieve satisfactory resolution of all DNA fragments in the RFLP analysis of the HTLV-2a subtype (8). In each experiment, after an initial denaturation for 5 min at 94°C, 35 cycles were carried out in a DNA thermocycler (Perkin-Elmer Cetus). Each cycle consisted of denaturation for 40 s at 94°C, annealing for 30 s at 57°C, and extension for 1 min at 72°C. The 35 cycles were followed by a 10-min extension at 72°C. For amplification and nucleotide sequencing of the entire pX region, nested PCR was carried out to prepare two overlapping products (nt 7192 to 7791 and 7713 to 8285). Primers in the first round for the first product were pX101 (nt 7153 to 7173; 5'-GGCAATCTCCTAAAATAGTCT-3') and pX102 (nt 7815 to 7795; 5'-TGTGTGTAGGAACATTTTGA-3'); primers in the nested reaction were pX103 (nt 7192 to 7212; 5'-TTACAATCCTGTCTCTCTCA-3') and pX104 (nt 7791 to 7771; 5'-AAGTCTCTTAATCGTTTATAG-3'). For the second overlapping product, the first-round primers were pX105 (nt 7683 to 7703; 5'-GCTTCCCCACCATGACATG-3') and pX106 (nt 8326 to 8306; 5'-GGGCGCGTGTTTCAGTTCCTA-3'), and primers in the nested PCR were pX107 (nt 7713 to 7733; 5'-ACCCATGTATATTCTGCCA-3') and pX108 (nt 8285 to 8265; 5'-AGCCTTTACTTGGGATTGTTT-3'). Reaction conditions were identical and, following initial denaturation at 94°C for 5 min, involved 35 cycles of denaturation at 94°C for 40 s, annealing at 52.5°C for 30 s, and chain elongation at 72°C for 40 s. Upon completion of each PCR, an additional extension at 72°C for 10 min was carried out.

For amplification of the pX region which contains the expected stop codon of *tax*, the following reaction conditions were used. The initial PCR was carried out with primers pX105 and pX106 as described above and was followed by nested PCR using primers pX106 (see above) and LSI (nt 8140 to 8160; 5'-GAATACCAACATCCCTGTC-3'). Reaction conditions were as described above, and amplified products were analyzed by agarose gel electrophoresis with ethidium bromide staining.

In view of the large number of PCRs and particularly those involving the LTR, maximum precautions were taken to avoid contamination (32). Physically separate rooms were used for DNA isolation, preparation of reaction mixes, the PCR, and analysis of amplified products. All reactions were carried out with appropriate positive and negative controls. It should also be noted that contamination between the Kayapo Indian and IVDA samples would be highly unlikely, as studies of the former were carried out in different laboratories and at an earlier time.

RFLP analysis. The amplified *env* gene product was directly digested with *XhoI* as described previously (23). Digestion of the amplified product of HTLV-2a by this enzyme would be expected to generate two fragments (430 and 180 bp), while the amplified product of HTLV-2b lacks the restriction site and is not cleaved. Digestion products were visualized in agarose gels after staining with ethidium bromide. RFLP analysis of the LTR was carried out as described previously (8). PCR products were initially purified by using Wizard PCR Preps (Promega), and 10 μ l of the purified DNA was directly digested with 5 U of *AvaII*, *BanII*, *BglI*, *Bsu36I*, *MseI*, and *SmaI* (Boehringer Mannheim, Indianapolis, Ind.), either singly or in combination, using conditions recommended by the manufacturer. Products were electrophoresed on 3.0% agarose gels and visualized by ethidium bromide staining.

Cloning and nucleotide sequence analysis of PCR-amplified products. Following PCR and electrophoresis on low-melting-temperature agarose gels, the amplified products from the *env*, LTR, and pX regions were excised and cloned

into the pCR-Script SK(+) vector (Stratagene, La Jolla, Calif.), using conditions recommended by the manufacturer and as previously described (8). Plasmid DNA from recombinant clones was extracted by using the Promega miniprep system (Promega). Nucleotide sequence analysis was carried out on two to three clones of each sample. DNAs were sequenced in both directions, using an automatic sequencing apparatus (Applied Biosystems, Norwalk, Conn.).

Phylogenetic analysis. Nucleotide sequences were aligned by using CLUSTAL W (version 1.5), obtained from the EMBL file server (12, 42). Phylogenetic trees were constructed by using programs from both the PHYLIP (version 3.5c) package (9) and the CLUSTAL W package (55). The nucleotide distances among the envelope clones were determined on nonidentical sequences, using the DNADIST program and correcting for evolutionary distance by using the Kimura two-parameter model (31). A phylogenetic tree was constructed by least-squares analysis using the FITCH program. To assess the reliability of the groupings obtained with FITCH, 1,000 bootstrap replicates were generated by using the SEQBOOT program, and the trees were analyzed by the CONSENSE program. Equivalent phylogenetic relationships were found among the envelope clones by using the maximum-likelihood analysis program DNAML and by the neighbor-joining method (45), using the NEIGHBOR program. Phylogenetic trees for the LTR clones were constructed by using the neighbor-joining method (45) as implemented in CLUSTAL W. The reliability of these trees was statistically evaluated by using 10,000 bootstrap replicates. Equivalent phylogenetic relationships among the LTR clones were found by using a least-squares analysis (FITCH) and a maximum-likelihood analysis (DNAML). The analysis included the 8 Brazilian IVDA samples and 25 other HTLV-2 LTR sequences, 18 sequenced in our laboratory (8) and 7 from the published literature. Details of the former have been previously published; these samples included 8 which had been identified as HTLV-2a and 10 identified as HTLV-2b (8). The latter included the two prototype subtypes of HTLV-2a and -2b, MoT (48) and NRA (34); G12, an isolate from a Guaymí Indian (43); two isolates from the Wayu Indians of Colombia (51); PygCam1, an isolate from a prostitute from Cameroon (38); and KT, an isolate from a prostitute from Ghana (28). Also included were two LTR sequences of the virus endemic in the Kayapo Indians of Brazil, one of which has been recently published (30).

Tax constructs. The expression vector pCG, used for the expression of Tax proteins, was kindly provided by M. Yoshida, Tokyo University, Tokyo, Japan. The pCG vector contains the cytomegalovirus promoter, the herpes simplex virus *tk* gene, 5' untranslated leader, and initiation codon, rabbit β -globin splicing and polyadenylation signals, and the replication origin of simian virus 40. To prepare the HTLV-2a Tax protein (Tax a), the extended Tax protein representative of the HTLV-2b isolates, and selected deletion mutants of these proteins, PCR was used to amplify the corresponding gene regions with primers containing *XbaI* and *BamHI* restriction endonuclease sites in the 5' and 3' ends, respectively. Primers used in the PCR were as follows. An upstream primer 5'-CGGGCTTCTAGACATGGCCCATTTCCAGGATT-3' (*XbaI* site underlined) was used in the preparation of both Tax a and the extended Tax proteins. The downstream primers for Tax a and the extended Tax protein were 5'-GTAAGTGGATCCCTAGTCGCCATTGTCATCCGA-3' and 5'-CGCGCGGATCCTTACTTGGGATTGTTGTGTG-3', respectively (*BamHI* sites underlined). To prepare Tax a with a 55-amino-acid deletion in the carboxy terminus, the downstream primer 5'-CGCGCGGATCCCTAGGAGAGATTGTTATAG-3' (*BamHI* site underlined) was used. Thirty-five cycles of denaturation at 94°C for 30 s, annealing for 1 min at 63°C, and chain elongation at 72°C for 2 min were used in all reactions. Amplified products were initially cloned in the pCR-Script SK(+) vector as described above, and the purified plasmids were subsequently digested with *XbaI* and *BamHI*. Inserts were separated by agarose gel electrophoresis and directly ligated to the pCG vector which had been previously digested with *XbaI* and *BamHI*.

Transfections. Jurkat cells were transfected with the pCG Tax constructs and pHTLV-II-LTR-CAT (originally derived from pVS2, CAT and kindly provided by K. Shimotohno, National Cancer Center Research Institute, Tokyo, Japan), using Lipofectin (Promega) as described by the manufacturer. One microgram of each plasmid was used to transfect 5×10^6 cells which had been grown in antibiotic-free RPMI medium containing 10% fetal calf serum and subsequently washed twice with Opti-MEM (Gibco) without serum. Following transfection, cells were incubated for 3 h at 37°C, at which time the Lipofectin-DNA complexes were removed. Cells were then incubated in RPMI containing 10% fetal calf serum without antibiotics for an additional 48 h.

CAT assays. Cells were collected, washed twice in ice-cold phosphate-buffered saline, resuspended in 300 μ l of Reporter lysis buffer (Promega), and incubated at 60°C for 15 min to inactivate endogenous deacetylase activity. After centrifugation, supernatants (100 μ l) were incubated with 3 μ l of [14 C]chloramphenicol (0.05 mCi/ μ l) and 5 μ l of *n*-butyryl coenzyme A (5 μ g/ μ l), and reactions were allowed to proceed at 37°C for 2 h. Five hundred microliters of ethyl acetate was added to each sample, and after vortexing and centrifugation at 14,000 rpm for 5 min, the upper organic phase was removed and evaporated to dryness on a Savant DNA SpeedVac apparatus. The residues were suspended in 30 μ l of ethyl acetate, and 10- μ l samples were spotted onto thin-layer chromatography silica gel plates (J. Baker SG/IBZ plates) which had been preequilibrated for 1 h with chloroform-methanol (97:3). Chromatography was continued for 60 min, and the plates were then dried at room temperature and exposed to X-Omat X-ray film for 24 h with intensifying screens. After chromatography, the butyrylated prod-

ucts were scraped from the plate and counted by liquid scintillation. Chloramphenicol acetyltransferase (CAT) activity was averaged from a minimum of three separate assays.

Accession numbers. The GenBank accession numbers (first term) of the HTLV-2 *env* fragments included in the phylogenetic study (second term) are as follows: U32905, 408N; U32901, MSA1bp; U32903, DOG (WH2); U32906, 130P; U32904, 72969N; U32902, GAR (WH6); U32900, PAR (WH7); U19110, Kayapo 1 (KAY1); U19109, KAY2; U32899, Sao Paulo 1 (SP1); U32898, SP2; U32897, SP3; U32896, SP4; U32895, SP5; Z46837, PH230PCAM; PygCam1 (17). A number of additional sequences of the *env* region which were available for analysis were not included, as these were found to have sequences identical to those of the representative HTLV-2a (DOG) and HTLV-2b (130P) samples included in the analysis (23, 26). The GenBank accession numbers of the HTLV-2 LTR fragments sequenced in our laboratory and included in the phylogenetic study are as follows: L37129, WY; L37130, CAM; L37131, DOG; L37132, DSA; L37133, FH; L37134, FLN; L37135, FUC; L37136, GAR; L37137, JD; L37138, MIN; L37139, SAC; L37140, SC; L37141, VIN; L37142, AG; L37143, ASB; L37144, MER; L37145, PAR; L37146, WEN; U25135, KAY1; U32887, KAY2; U32888, SP1; U32889, SP2; U32890, SP3; U32891, SP4; U32892, SP5; U32893, SP6; U32894, SP7; U32886, Rio de Janeiro 1 (RJ1). Seven published LTR sequences were also included in the analysis. In addition to sequences of the prototype (MoT, NRA, and G12) viruses (accession numbers are listed below), two sequences of HTLV-2b from two Wayu Indians (51), WYU-1 and WYU-2 (U12792 and U12794), and two of HTLV-2a from African prostitutes (28, 38) (PH230PCAM [GenBank accession number Z246838] and KT [28]) were included. The GenBank accession numbers for pX fragments are as follows: U32885, FLN; U32871, DOG; U32883, SAC; U32884, MIN; U32882, FUC; U32880, PAR; U32881, GAR; U32879, VIN; U32875, KAY1; U32874, KAY2; U32873, SP1; U32872, SP2; U32878, SP3; U32877, SP4; U32876, SP5. The GenBank accession numbers for the complete nucleotide sequences of the prototype HTLV-2a (MoT) (40) and for HTLV-2b (NRA) (30) and G12 (35) are MID060, L20734, and L11456, respectively.

RESULTS

RFLP and nucleotide sequence analysis of the *env* region.

RFLP analysis using *Xho*I digestion of the PCR product of the *env* region encoding the transmembrane glycoprotein, gp21, was carried out to differentiate infections by the two known HTLV-2 subtypes. Previous studies have shown that whereas all HTLV-2a isolates contain this enzyme site, it is absent in HTLV-2b isolates (23). It could be shown that the amplified products of all 28 seropositive samples were effectively cleaved, yielding two fragments with expected sizes of 450 and 180 nt and demonstrating that all were infected with a virus related to the HTLV-2a subtype. To confirm infection by the HTLV-2a subtype and to determine the relationship of the viruses both to prototype isolates and to the virus endemic in the Kayapo Indians, the nucleotide sequence of the amplified *env* product was determined in five randomly chosen samples (SP1 to SP5). Phylogenetic analysis of a 589-nt region clearly demonstrated that all of the Brazilian sequences segregated with prototypic HTLV-2a isolates from North America (Fig. 1). While the isolates did not cluster on a single branch, distinct phylogenetic groups within the HTLV-2a subtype could not be clearly identified. This reflects the marked conservation of nucleotide sequence in this region of the proviruses (23, 26, 29). Examination of the nucleotide sequences of the *env* clones revealed that HTLV-2a (Mo) (40) and HTLV-2b (NRA) (30) differ by 24 residues in the region examined. The Kayapo clones matched HTLV-2a at 20 positions and HTLV-2b at 4 positions and had one unique residue. The isolates from the urban areas, SP1 to SP5, matched HTLV-2a at 21 positions and HTLV-2b at 3 positions and had two unique residues.

Nucleotide sequence analysis of the LTR region. To further clarify the relationship of the Brazilian viruses to prototype HTLV-2 isolates, the phylogenetic relationship of a 449-nt region of the LTR was analyzed. Previous studies have shown that analysis of the LTR region permitted the identification of distinct phylogenetic clusters within both virus subtypes (8, 43). Figures 2A and B show unrooted and rooted phylogenetic trees, respectively; in these analyses, 8 sequences (SP1 to SP7

FIG. 1. Phylogenetic analysis of 589 nucleotides the *env* gene region encoding the transmembrane glycoprotein gp21 of HTLV-2 isolates from five IVDAs from Sao Paulo (SP1 to SP5) and two Kayapo Indians (KAY1 and KAY2) with representative HTLV-2a and HTLV-2b isolates, including the prototype HTLV-2a isolate (MoT) and the prototype HTLV-2b isolate (NRA), representative isolates of HTLV-2a from a New York City IVDA (DOG) and North American Indians (130P and MSA1bp), and corresponding HTLV-2b isolates from New York City IVDAs (GAR and PAR) and North American Indians (408N and 60405N). Other available sequences of both subtypes in these two populations (26) were intentionally not included in the analysis, as they were found to have sequences identical to those of the representative isolates indicated above. Also included in the analysis were isolate G12, from a Guaymí Indian, and PH230PCAM and PygCam1 from African prostitutes. The origins of these isolates are detailed in Materials and Methods. The bootstrap statistical analysis was applied, using 1,000 bootstrap replicates with HTLV-1 (ATK) as the outgroup.

and RJ1) analyzed in the present study were compared with the sequence of the virus present in two Kayapo Indians and an additional 25 sequences representative of HTLV-2a and -2b isolates. It could be shown that all of the Brazilian viruses, while being more closely related to the HTLV-2a than to the HTLV-2b subtype, clearly clustered in a distinct phylogenetic group. The reliability of the phylogenetic trees was statistically evaluated by using 10,000 bootstrap replicates and provided a confidence level of 97 to 98% in the branching of the Brazilian sequences. Equivalent phylogenetic relationships were obtained with each method used. Examination of the nucleotide sequences of the LTR clones revealed that HTLV-2a and HTLV-2b differ at 20 positions in this region. The Brazilian sequences matched HTLV-2a at 17 positions and HTLV-2b at 3 positions and had unique residues at 4 positions. The results also established the close relationship of the viruses present in the urban areas of Brazil to that present in the Kayapo Indians.

FIG. 2. (A) Unrooted phylogenetic analysis of a 449-nt region of the LTR. Included in the analysis were isolates from eight IVDAs from Brazil (SP1 to SP7 and RJ1) and two Kayapo Indians and 25 other viruses representative of HTLV-2a and HTLV-2b subtypes. Details of the latter viruses are outlined in Materials and Methods. The bootstrap statistical analysis was applied, using 10,000 bootstrap replicates. (B) Rooted phylogenetic analysis of the same 449 nt of the LTR with HTLV-1 (ATK) as the outgroup. Sequences analyzed were as in panel A.

FIG. 3. Comparative RFLP analysis of prototype HTLV-2a isolates from urban IVDAs and Kayapo Indians in Brazil. (A) Predicted restriction endonuclease maps of a 553-nt region of the LTR, based on computer analysis of the nucleotide sequences. RFLP groups a_1 to a_4 were as previously described (8) and are representative of the isolates shown in Fig. 2. The predicted restriction map of the Brazilian isolates is shown at the bottom, and the positions of the cleavage sites of the enzymes (*Ava*II, *Ban*II, *Bgl*I, *Bsu*36I, *Mse*I and *Sma*I) used in the RFLP analysis are indicated. In all, 24 of 28 samples (22 from IVDAs and 2 from Kayapo Indians) had identical patterns. The four remaining Brazilian IVDAs samples had an additional *Ban*II [*Ban*II] site in the 5' of the amplicon. All Brazilian samples could be differentiated from the a_1 to a_4 groups by the presence of an *Ava*II site in 3' end of the amplicon. (B) Differentiation of the Brazilian isolates from the a_1 to a_4 groups by *Ava*II digestion. Agarose gel electrophoresis of digested products show the presence of the additional enzyme site in the Brazilian isolates (Bzl). Sizes are indicated in nucleotides.

On the basis of their distinct phylogenetic clustering, the Brazilian viruses were tentatively considered to represent a distinct molecular subtype. An interesting finding but unrelated to the main focus of this study was the observation that the three HTLV-2b isolates from the Wayu Indians of Colombia formed three distinct phylogenetic clusters (WY, WYU-1, and WYU-2). The reasons for this are unclear but might suggest that in addition to endemic infection, there may have been a recent introduction(s) of one or more of these viruses into this population (51). As far as we are aware, there has only been one other phylogenetic analysis of the LTR of HTLV-2 isolates (52). However, the results of that study and ours are difficult to compare, as the former involved a different region of the LTR; in addition, it included only one sample from Brazil. However, it is important to note that this sample also appeared to cluster distinctly from the North American HTLV-2a subtypes.

RFLP analysis of the LTR. To determine if the HTLV-2 in our urban populations represented infection by the same or a closely related virus, RFLP analysis of the LTR was used to compare 28 of the viruses (26 IVDA and 2 Kayapo isolates) included in this study. Previous studies from our laboratory have demonstrated that analysis of the LTR of HTLV-2a isolates from North America allowed their differentiation into four distinct RFLP groups, which were designated a_1 to a_4 (Fig. 3A) (8). Computer-generated restriction endonuclease maps of the LTR sequences obtained for the two Kayapo and eight IVDA isolates shown in Fig. 2 suggested that the Brazilian samples would have a distinct RFLP pattern, which was tentatively designated Bzl (Fig. 3A). All of the samples were found to have a characteristic *Ava*II site in the 3' end of the amplicon which would allow their differentiation from the four RFLP groups present in the North American samples (Fig. 3A). A typical RFLP analysis demonstrating the ability of *Ava*II digestion to differentiate the Brazilian from North American viruses is shown in Fig. 3B. The RFLP patterns of all 28 samples were analyzed by using the six restriction endo-

nucleases *Ava*II, *Ban*II, *Mse*I, *Bsu*36I, *Sma*I, and *Bgl*II. It could be shown that 24 (two Kayapo and 22 IVDA) samples had a pattern identical to that designated Bzl (Fig. 3A). The remaining four samples (all from IVDAs) had an additional *Ban*II site in the 5' end of the amplicon ([*Ban*II] in Fig. 3A). Nucleotide sequence analysis revealed that in the four LTRs, this was generated by a single nucleotide substitution (T→G) at nt 315. The nucleotide sequence of the remaining region of the amplicon was unchanged. The results demonstrate that the majority (24 of 28) of the Brazilian HTLV-2 samples in this study involved a closely related if not identical virus. The remaining 4, while having a different RFLP pattern characterized by an additional *Ban*II site, were considered on the basis of nucleotide sequence analysis to be extremely closely related to the former 24 samples. The observation that all 28 Brazilian samples could be readily differentiated from the HTLV-2a present in North America suggests that at least in our study population, infection by viruses from the latter geographic region had not occurred.

Nucleotide sequence analysis of the pX region. To support the view that the Brazilian viruses represent a unique molecular subtype and to determine if there may be phenotypic differences between the subtypes, we focused on the nucleotide sequence and on preliminary functional analysis of products of the pX region. The pX region is known to encode at least two regulatory proteins, Rex and Tax. Previous studies have shown that the Tax protein of the HTLV-2b subtype (Tax b) contains an additional 25 amino acids at the carboxy terminus compared with Tax a as a result of a nucleotide substitution(s) in the stop codon present in HTLV-2a isolates (22, 40, 42, 43). However, it is unclear if these additional amino acids may result in differences in the biological properties or phenotype of this subtype. Nucleotide sequence analysis demonstrated that the Brazilian viruses differed from HTLV-2a but were similar to HTLV-2b in that the expected stop codon was also abrogated by a single nucleotide substitution (Fig. 4). As a result, these

	7192 Start TAX/REX →
HTLV-IIa (Mo)	<u>TTACAATCCTGTCTCCTCTCAGCCCATTTCCAGGATTGGACAGAGCCTCCTATATGGATACCCCGTCTACGTGTTT</u>
KAYAPO-1	
KAYAPO-2	
SAO PAULO-1	
SAO PAULO-2	
HTLV-IIb (FUC)	C.....
HTLV-IIb (PAR)	C.....
HTLV-IIb (NRA)	C.....
HTLV-IIb (G12)	C.....
	7270
HTLV-IIa (Mo)	GGCGATTGTGTACAGGCCGATTGGTGTCCCGTCTCAGGTGGTCTATGTTCACCCCGCTACATCGACATGCCCTCCTG
KAYAPO-1	
KAYAPO-2	
SAO PAULO-1	A.....
SAO PAULO-2	
HTLV-IIb (FUC)	.A.....
HTLV-IIb (PAR)	
HTLV-IIb (NRA)	
HTLV-IIb (G12)	
	7348
HTLV-IIa (Mo)	GCCACCTGTCCAGAGCACCAACTCACCTGGGACCCCATCGATGGACGCGTTGTACAGCTCTCCTCTCCAATACCTTATC
KAYAPO-1	
KAYAPO-2	
SAO PAULO-1	
SAO PAULO-2	A.....
HTLV-IIb (FUC)	G.....
HTLV-IIb (PAR)	G.....
HTLV-IIb (NRA)	G.....
HTLV-IIb (G12)	G.....
	7426
HTLV-IIa (Mo)	CCTCGCCTCCCTCCTTCCCCACCCAGAGAACCTCAAGGACCTCAAGGTCTTACCCCTCCCACTCCTGTCTCC
KAYAPO-1	A.....
KAYAPO-2	
SAO PAULO-1	G.....
SAO PAULO-2	
HTLV-IIb (FUC)	G.C.A.....
HTLV-IIb (PAR)	G.C.A.....
HTLV-IIb (NRA)	G.C.A.....
HTLV-IIb (G12)	C.A.....
	7504
HTLV-IIa (Mo)	CCCAAGGTTCCACCTGCCTTCTTTCAATCAATGCGAAAGCACACCCCTACCGAAATGGATGCCTGGAACCAACCTC
KAYAPO-1	
KAYAPO-2	C.....
SAO PAULO-1	
SAO PAULO-2	
HTLV-IIb (FUC)	C.....C.....T.C.....
HTLV-IIb (PAR)	C.....C.....T.C.....
HTLV-IIb (NRA)	C.....C.....T.C.....
HTLV-IIb (G12)	C.....C.....T.C.....
	7582
HTLV-IIa (Mo)	GGGGATCAGCTCCCTCCCTCGCCTTCCCCGAACCTGGCCTCCGTCCCCAAAACATCTACACCACCTGGGGAAAAACC
KAYAPO-1	T.....
KAYAPO-2	T.....
SAO PAULO-1	T.....
SAO PAULO-2	T.....
HTLV-IIb (FUC)	T.....
HTLV-IIb (PAR)	T.....
HTLV-IIb (NRA)	T.....
HTLV-IIb (G12)	T.....
	7660 → End REX
HTLV-IIa (Mo)	GTA GTATGCCTATACCTATACAGCTTTCCCCACCCATGACATGGCCACTTATACCCCATGTCATATCTGCCACCCC
KAYAPO-1	
KAYAPO-2	
SAO PAULO-1	C.....
SAO PAULO-2	C.....
HTLV-IIb (FUC)	G.....G.....T.....C.....C.....A
HTLV-IIb (PAR)	G.....G.....T.....C.....C.....A
HTLV-IIb (NRA)	G.....G.....T.....C.....C.....A
HTLV-IIb (G12)	G.....G.....T.....C.....C.....A

FIG. 4. Comparison of the nucleotide sequences of the pX region encoding Tax and Rex of isolates from two IVDAs from Sao Paulo and two Kayapo Indians with the published sequences of HTLV-2-Mo, the prototype HTLV-2a (40), and HTLV-2b from two IVDAs from New York City (FUC and GAR), G12, a Guaymi Indian isolate, and HTLV-2-NRA, the prototype HTLV-2b (30). With the exception of G12, a single nucleotide substitution, T→C at position 8203, was present in the Sao Paulo, Kayapo, and HTLV-2b samples. In G12, two substitutions were found in the stop codon. These substitutions abrogated the stop codon present in the HTLV-2-Mo provirus.

	7738	
HTLV-IIa (Mo)	AGACAATTAGGAGCCTTCCTCACCAAGGTGCCTCTAAAACGATTAGAAGAACCTTCTATACAAAATGTTCTTACACACA	
KAYAPO-1		
KAYAPO-2		
SAO PAULO-1		
SAO PAULO-2		
HTLV-IIb (FUC)	G.....C.....	
HTLV-IIb (PAR)	G.....C.....	
HTLV-IIb (NRA)	G.....C.....	
HTLV-IIb (G12)	G.....C.....	
	7816	
HTLV-IIa (Mo)	GGGACAGTCATAGTCCTCCCGGAGGACGACCTACCCACCACAATGTTCCAACCCGTGAGGGCTCCCTGTATCCAGACT	
KAYAPO-1	..G.....	
KAYAPO-2	..G.....	
SAO PAULO-1	..G.....G.....	
SAO PAULO-2	..G.....G.....	
HTLV-IIb (FUC)	..AG.G.....G.....A.....	
HTLV-IIb (PAR)	..AG.G.....G.....A.....	
HTLV-IIb (NRA)	..AG.G.....G.....A.....	
HTLV-IIb (G12)	..AG.G.....G.....A.....	
	7894	
HTLV-IIa (Mo)	GCCTGGGTGTACAGGACTTCTCCCTTACTCCTCATCTTAACAACCCAGGTCTAATATGGACCTTCAATGACGGCTCA	
KAYAPO-1		
KAYAPO-2		
SAO PAULO-1		
SAO PAULO-2		
HTLV-IIb (FUC)	C.....C.....C..T.....	
HTLV-IIb (PAR)	C.....C.....C..T.....	
HTLV-IIb (NRA)	C.....C.....C..T.....	
HTLV-IIb (G12)	C.....C.....C..T.....	
	7972	
HTLV-IIa (Mo)	CCAATGATTTCGGCCCTTACCCCAAAGCAGGGCAGCCATCTTTAGTAGTTTCAGTCCCTCCTATTAACTCTTCGAAAA	
KAYAPO-1	G.....G.....	
KAYAPO-2	G.....G.....	
SAO PAULO-1	G.....G.....	
SAO PAULO-2	G.....G.....	
HTLV-IIb (FUC)	G..T..G.....A.....T.....	
HTLV-IIb (PAR)	G..T..G.....A.....T.....A.....	
HTLV-IIb (NRA)	G..T..G.....A.....T.....	
HTLV-IIb (G12)	G..T..G.....A.....T.....	
	8050	
HTLV-IIa (Mo)	TTTCGAAACCAAAGCCTTCCATCCCTCCTATCTACTCTCTCATCAGCTTATACAATACTCCTCCTTCCATAACCTTCAC	
KAYAPO-1	..C.....	
KAYAPO-2	..C.....	
SAO PAULO-1	..C.....	
SAO PAULO-2	..C.....	
HTLV-IIb (FUC)	..C.....T.....A.....	
HTLV-IIb (PAR)	..C.....T.....A.....	
HTLV-IIb (NRA)	..C.....T.....A.....	
HTLV-IIb (G12)	..C.....T.....A.....	
	8128	End TAX →
HTLV-IIa (Mo)	CTTCTATTTCGATGAATACACCAACATCCCTGTCTCTATTTTATTTAATAAAGAAGAGGCGGATGACAATGGCGACTAG	
KAYAPO-1	C.....C..G.....	
KAYAPO-2	C.....C..G.....	
SAO PAULO-1	C.....C..G.....	
SAO PAULO-2	C.....C..G.....	
HTLV-IIb (FUC)	..C.....C.....G.....C.....	
HTLV-IIb (PAR)	..C.....C.....G.....C.....	
HTLV-IIb (NRA)	..C.....C..G.....C.....	
HTLV-IIb (G12)	..C.....C.....C..G.....	
	8206	End EXT-TAX → 8285
HTLV-IIa (Mo)	CCTCCCGAGCCAGCCACCCAGGGCGAGTCATCGACCCAAAAGGTCAGACCGTCTCACACAAACCAATCCCAAGTAAAGGCT	
KAYAPO-1		
KAYAPO-2		
SAO PAULO-1		
SAO PAULO-2		
HTLV-IIb (FUC)	T.....G.....	
HTLV-IIb (PAR)	T.....G.....	
HTLV-IIb (NRA)	T.....G.....	
HTLV-IIb (G12)	T.....G.....	

FIG. 4—Continued.

TAX

HTLV-IIa(Mo) MAHFPGFGQSLLYGYFVYVFGDCVQADWCPVSGGLCSTRLHRHALLATCPEHQLTWDPIDGRVVSPLQYLIPRLPSFPTQRTSRTLKVL
 KAYAPO-1
 KAYAPO-2
 SAO PAULO-1Q.....
 SAO PAULO-2S.....
 HTLV-IIb(FUC)D.....AK.....
 HTLV-IIb(PAR)AK.....
 HTLV-IIb(NRA)AK.....
 HTLV-IIb(G12)K.....

HTLV-IIa(Mo) TPPTTFVSPKVPFAFFQSMRKHTPYRNGCLEPTLGDQLPSLAFPEPGLRPQNIYTTWGKTVVCLYLYQLSPPMTWPLIPHVIFCHPRQLG
 KAYAPO-1I.....
 KAYAPO-2
 SAO PAULO-1
 SAO PAULO-2
 HTLV-IIb(FUC)F.....S.....
 HTLV-IIb(PAR)F.....S.....
 HTLV-IIb(NRA)F.....
 HTLV-IIb(G12)F.....

HTLV-IIa(Mo) AFLTKVPLKRLEELLYKMFLLHTGTIVLPEDDLPTMTFQPVRAFCIQTAWCTGLLPYHSILTTPLGIWTFNDGSPMISGPYPKAGQPSLV
 KAYAPO-1A.....C.....
 KAYAPO-2A.....C.....
 SAO PAULO-1A.V.....C.....
 SAO PAULO-2A.V.....C.....
 HTLV-IIb(FUC)A.....C.....
 HTLV-IIb(PAR)A.....C.....
 HTLV-IIb(NRA)A.....C.....
 HTLV-IIb(G12)A.....C.....

HTLV-IIa(Mo) VQSSLLIFEKFFETKAFHPSYLLSHQLIQYSSFHNLHLFDEYTNIPVSILFNKEEADDNGD*
 KAYAPO-1Q.....T.....QPPEPATQGESSTQKVRFSHTNNPK*
 KAYAPO-2Q.....Q.....
 SAO PAULO-1Q.....Q.....
 SAO PAULO-2Q.....Q.....
 HTLV-IIb(FUC)Q.....G.....Q.....A.....
 HTLV-IIb(PAR)K.....Q.....G.....Q.....A.....
 HTLV-IIb(NRA)Q.....Q.....A.....
 HTLV-IIb(G12)Q.....R.....A.....

REX

HTLV-IIa(Mo) MPKTRRQRTRRRARNRPPTWPISQDLDRASYMDPTSTCLAIVYRPIGVPSQVYVPPAYIDMPSWPPVQSTNSPGTSPMDALSALLSNT
 KAYAPO-1
 KAYAPO-2
 SAO PAULO-1N.....
 SAO PAULO-2
 HTLV-IIb(FUC)S.....T.....S.....
 HTLV-IIb(PAR)S.....S.....
 HTLV-IIb(NRA)S.....S.....
 HTLV-IIb(G12)S.....S.....

HTLV-IIa(Mo) LSLASPPSPPREPQGQPSRSLPLPPLLSPFRHLPSFNQCESTPPTMDAWNQPSGISPPSPSPNLASVPKSTSTPPGKEP*
 KAYAPO-1L.....
 KAYAPO-2P.....L.....
 SAO PAULO-1G.....L.....
 SAO PAULO-2L.....
 HTLV-IIb(FUC)PR.....P.....S.....IA.....L.....
 HTLV-IIb(PAR)PR.....P.....S.....IA.....L.....
 HTLV-IIb(NRA)PR.....P.....S.....IA.....L.....
 HTLV-IIb(G12)PR.....P.....S.....IA.....L.....

FIG. 5. Predicted amino acid sequences of the Tax and Rex proteins of the HTLV-2 isolates analyzed in Fig. 4. It would be expected that the Brazilian and prototype HTLV-2b isolates would have an additional 25 amino acids at the carboxy terminus. In addition, the Brazilian isolates had four amino acid differences from Mo, resulting in identical amino acids present in the HTLV-2b isolates.

proteins would also be expected to have an extended Tax protein with an additional 25 amino acids at the carboxy terminus (Fig. 5). In addition to the additional 25 amino acids at the carboxy terminus, the Brazilian viruses had consistent amino acid differences from HTLV-2-Mo, the prototype HTLV-2a subtype, resulting in amino acids which are present in all of the HTLV-2b isolates examined. To ensure that the loss of the stop codon is a common feature, PCR was used to amplify a 187-nt region flanking the stop codon from a total of 12 (10 IVDA and 2 Kayapo) Brazilian isolates and a larger

number of representative HTLV-2a and HTLV-2b subtypes. It could be shown that identical substitutions were present in the stop codons of all 12 Brazilian isolates. Representative nucleotide sequence analysis of samples from five IVDA (SP1 to SP5) and two Kayapo Indians is shown in Fig. 6. The results also confirmed that similar substitutions were present in all representative HTLV-2b isolates but that the stop codons were intact in all HTLV-2a isolates analyzed.

Transactivation of the HTLV-2 LTR by HTLV-2 Tax proteins. To compare the relative abilities of Tax a and the ex-

FIG. 6. Comparison of the nucleotide sequences of the 3' ends of the pX region of five HTLV-2a (Mo, FLN, DOE, SAC, and MIN), five HTLV-2b (FUC, PAR, GAR, VIN, and NRA), and seven representative Brazilian isolates (from two Kayapo Indians and five from IVDAs from Sao Paulo), confirming the nucleotide substitution in the stop codon at position 8203 in the HTLV-2b and Brazilian viruses.

tended Tax proteins representative of HTLV-2b and the Brazilian viruses to transactivate the HTLV-2 LTR, the proteins were cloned into the eukaryotic expression vector pCG. Constructs were prepared to exclude expression of Rex and were used to transfect Jurkat cells together with a plasmid expressing the HTLV-2 LTR with CAT as a reporter gene. Cells were harvested at 48 h after transfection, and cell extracts were analyzed by thin-layer chromatography for CAT activity. Quantification of CAT activity was assayed by measuring incorporation of radioactivity into butyrylated chloramphenicol.

FIG. 7. Representative CAT assay comparing the transactivation properties of Tax a (lane 1) with those of the extended Tax protein (lane 2) representative of HTLV-2b and -2c. Jurkat cells were transfected as described in Materials and Methods, and cell extracts were assayed for CAT activity by measuring the conversion of chloramphenicol (CM) to the butyrylated form (b-CM), using thin-layer chromatography.

The results of a typical experiment are shown in Fig. 7, and a summary of the data from several experiments is provided in Table 1. We found that the extended Tax proteins had consistently much higher transactivation activity than Tax a. Deletion (by 55 amino acids) of the carboxy terminus of Tax a resulted in a complete loss of transactivation activity, confirming previous studies on HTLV-2-Mo (an HTLV-2a subtype), which demonstrated that this region is essential in transactivation (3).

DISCUSSION

Previous studies have demonstrated the existence of two molecular subtypes of HTLV-2, HTLV-2a and HTLV-2b (5, 8, 22, 23, 26, 44, 53). While the majority of IVDAs in North

TABLE 1. Transactivation of the HTLV-2 LTR by Tax proteins^a

Construct	% Butyrylation of [¹⁴ C]chloramphenicol in transfected cells ^b
Tax a (amino acids 1–331)	100
Tax b (amino acids 1–356)	480
Deleted Tax b (amino acids 1–331)	95
Deleted Tax a (amino acids 1–276)	<1

^a Jurkat cells were cotransfected with the pHTLV-II-Mo-LTR-CAT and pCG Tax constructs as described in Materials and Methods.

^b Normalized to that of the complete Tax a (100%). The values are representative of an averages of at least three assays performed with different plasmid preparations. All results were obtained within the linear range of CAT activity.

America are infected with the HTLV-2a subtype (8, 23, 44), several investigations have shown that HTLV-2b is clearly the predominant cause of infection in the majority of American Paleo-Indian groups so far analyzed (5, 22, 26, 41, 44, 52). The one notable exception is the Kayapo Indian group, whose members reside in the Amazon region of Brazil and have been shown to have endemic infection with a distinct variant of the HTLV-2a subtype (30).

In this study, we have attempted to analyze the molecular characteristics of HTLV-2 isolates present in urban areas of Brazil and to compare these isolates both with the viruses endemic in the Kayapo group and with prototype members of the HTLV-2a and HTLV-2b subtypes. RFLP analysis of a PCR-amplified *env* product from 28 infected individuals living in these urban areas suggested that all were infected with the HTLV-2a subtype. This finding was directly confirmed by nucleotide sequence analysis of the same region, which demonstrated that the viruses clustered with prototypic HTLV-2a isolates. As this region is highly conserved (23, 26, 29), with only a small number of nucleotide changes between isolates, it was not possible to confidently identify distinct phylogenetic subclusters in this region within either of the subtypes. In contrast, phylogenetic analysis of the nucleotide sequences of the LTR demonstrated that while the Brazilian viruses were much more closely related to HTLV-2a than to HTLV-2b, they clearly clustered in a distinct phylogenetic group. Moreover, the analysis of the LTR clearly demonstrated that the viruses endemic in the Kayapo Indians and those present in these urban areas of Brazil were very closely related. The distinct phylogenetic grouping of the latter viruses from individuals of such diverse backgrounds suggests that not only might they be of common origin but they should be considered a distinct and unique molecular subtype. This view was also supported by the RFLP analysis of the LTR, which demonstrated that all of the Brazilian viruses analyzed were very closely related and could be differentiated from the prototype HTLV-2a present in North America.

The classification of the Brazilian viruses as a distinct subtype was supported by sequence analysis and functional studies of the Tax protein. The pX region of the HTLV-2 provirus encodes at least two important regulatory proteins, Tax and Rex (47, 49). Tax is necessary for the transactivation of virus transcription, and this involves indirect interactions of the protein with the LTR via a number of cellular transcription factors (3, 47, 49). Studies on HTLV-2 Tax, which are now known to have been carried out on the prototype HTLV-2a subtype (HTLV-2-Mo) (48), have shown that sequences in both the amino and carboxy termini of the Tax protein are required for transactivation of the virus LTR (3). Recent nucleotide sequence analysis has shown that Tax b would be expected to differ from Tax a in that the former has an additional 25 amino acids at the carboxy terminus (22, 40, 42, 43). This results from a nucleotide substitution(s) which eliminates the stop codon present in all HTLV-2a isolates (22, 40, 42, 43). In the present study, we have also shown that the nucleotide sequences of the pX regions of all of the Brazilian viruses analyzed also have substitutions in this stop codon and would, like HTLV-2b, be expected to encode an extended Tax protein. Thus, our findings indicate that despite being phylogenetically closely related to HTLV-2a, the Brazilian viruses are phenotypically more similar to HTLV-2b. These findings support the proposal that the Brazilian viruses can be justifiably classified as a new and distinct molecular subtype of HTLV-2. The choice of nomenclature for the Brazilian viruses is made somewhat difficult by their close phylogenetic relationship to HTLV-2a. However, the bootstrap values obtained in the phylogenetic analysis sup-

port the conclusion that this is a unique subtype. In addition, the phenotypic differences between HTLV-2a and the Brazilian viruses with respect to the Tax proteins add strong support for their classification as a distinct molecular subtype. We tentatively propose that this subtype be designated HTLV-2c.

The demonstration that the proposed HTLV-2c subtype has an extended Tax protein has important implications for the use of serological methods employing Tax-derived peptides to differentiate the HTLV-2a and HTLV-2b subtypes. Previous studies have shown that the additional 25 amino acids at the carboxy terminus of the HTLV-2b tax protein contain a linear epitope (40, 42) which when used as an antigen in serological assays permits identification of infection by HTLV-2b (13, 40, 42). It has been shown that whereas the majority of individuals infected with HTLV-2b have immunoreactivity, as expected those infected with HTLV-2a do not (40). Preliminary studies from our laboratory have also demonstrated reactivity to this region in sera from both the infected Kayapo Indians and the Brazilian IVDAs analyzed in this study (21a). Thus, this assay would appear to have limited usefulness in this geographic area, as reactive Brazilian samples might be erroneously considered to represent the HTLV-2b subtype.

In this investigation, we have also carried out preliminary studies to compare the abilities of the Tax proteins to transactivate the HTLV-2 LTR, employing transient expression systems with CAT as a reporter gene. It could be shown that the transactivation abilities of the extended Tax proteins were consistently greater than that of Tax a, a finding which provides evidence for the first time that there may be differences in phenotype among the HTLV-2 subtypes. In agreement with previous studies, we found that the carboxy terminus of Tax a is required for effective transactivation of the virus LTR. It is unknown how the additional amino acids at the carboxy terminus of the HTLV-2b and HTLV-2c Tax proteins might contribute to the increased transactivation. It is known that Tax does not bind directly but rather interacts with a range of cellular factors which then interact with specific regions of the LTR (3, 47). It is possible, but remains to be established, that the additional amino acids either enhance these interactions or provide a second transactivation motif which could possibly lead to the more effective recruitment of cellular transcription factors. Further investigations using site-directed mutagenesis of the Tax proteins and identification of the cellular transcription factors involved should help to resolve these questions. In this study, we analyzed only the transactivation of the LTR of the HTLV-2a subtype. However, as the 21-bp repeats in the U3 region of the LTR which are involved in transactivation are unchanged in both HTLV-2b (53) and the Brazilian viruses (21a), it is unlikely that the results of the transactivation studies would be different if the LTRs of the latter were used in the experiments.

The close relationship of the viruses present in the urban areas of Brazil and those endemic in at least one indigenous Amazonian population suggests a common origin of the viruses, and particularly this subtype in this geographic region. It seems highly probable that as HTLV-2 infection was likely introduced to Brazil many thousands of years ago by the ancestors of present-day Indian groups, this subtype was only recently introduced into urban populations of the country. It is now probably maintained by intravenous drug abuse, contaminated blood transfusions, and perhaps sexual and vertical transmission (22). It seems less likely, but is certainly also possible, that the virus instead has only recently been introduced into the indigenous populations by interactions with infected individuals from urban areas.

Although we have analyzed only a small sample number, the

nucleotide sequence and RFLP analysis show that the Brazilian viruses can be readily distinguished from prototype North American HTLV-2a. The demonstration by RFLP analysis of the absence of the HTLV-2b and prototypic North American HTLV-2a subtypes in this IVDA population is intriguing and suggests that at least the population involved in our study has had little interaction with individuals or blood products from other geographic areas. These findings differ from those of a recent study which also used RFLP analysis of the LTR to characterize HTLV-2 infection in three Brazilian blood donors (52). While the RFLP methods used in the two studies cannot be directly compared because of differences in the restriction endonucleases employed, the conclusion of the previous study appears to differ from ours. In that study, it was reported that all three Brazilian samples had an RFLP pattern similar to that of HTLV-2-Mo, the prototype North American HTLV-2a isolate. Unfortunately, no information was available on the background of these three individuals. Thus, it is unclear if they may have been infected outside Brazil or whether they might have received North American blood products in Brazil or perhaps had other risk factors for HTLV-2 infection. Studies using larger numbers of samples from that country with detailed demographic data of the infected individuals may help to resolve these questions.

The demonstration of the existence of a third distinct molecular subtype of HTLV-2, now designated HTLV-2c, which is phylogenetically closely related to prototypic North American HTLV-2a subtype but more phenotypically similar to HTLV-2b may also have important implications for better understanding the role of these viruses in human disease. Prospective clinical and immunological studies in populations infected with the different subtypes may reveal important differences in their relative pathogenicities, particularly with respect to disorders in which the Tax protein may have an important pathogenic role. The findings of the present study also raise questions regarding the origin(s) of this virus family. It seems almost certain that HTLV-2 originated in the Old World. This view is supported not only by the finding of HTLV-2 infection in remote and reportedly isolated groups of Africa (11, 15–17, 19, 20) but also by the identification of closely related viruses in several simian species in that country (18, 21, 36). The virus was probably introduced to the New World with the migration of the ancestors of present-day Indian populations. The finding of distinct subtypes and variants thereof raises the possibility that these may have been introduced during distinct migration periods. Continued studies of the viruses present in other Indian populations in isolated geographic regions of South America and comparison of these viruses with those present in aboriginal Asian and African populations may help to resolve these questions.

ACKNOWLEDGMENTS

N.E. and P.N. contributed equally to this work.

This study was supported by grant CA64038 from the NIH/NCI, by General Clinical Research Center grant M01-RR0012 from the National Center for Research Resources at the National Institutes of Health, and by the Japanese Foundation for AIDS Prevention.

We thank Devjani Chatterjee and Karl Smalley for providing the PHYLIP and the CLUSTAL W program packages and Joe Felsenstein, Department of Genetics University of Washington, Seattle, for valuable discussions on the phylogenetic analysis and virus nomenclature.

REFERENCES

- Biglione, M., A. Gessain, S. Quiruelas, O. Fay, M. A. Taborda, E. Fernandez, S. Lupo, A. Panzita, and G. de The. 1993. Endemic HTLV-II infection among Toba and Matakó Amerindians from North Argentina. *J. Acquired Immune Defic. Syndr.* **6**:631–633.
- Black, F. L., R. J. Biggar, J. V. Neel, E. M. Maloney, and D. J. Waters. 1994. Endemic transmission of HTLV type II among Kayapo Indians of Brazil. *AIDS Res. Hum. Retroviruses* **10**:1165–1171.
- Cann, A. J., J. D. Rosenblatt, W. Wachsmann, and I. S. Y. Chen. 1989. In vitro mutagenesis of the human T-cell leukemia virus type I and II *tax* genes. *J. Virol.* **63**:1474–1479.
- Cartier, L., F. Araya, J. L. Castillo, V. Zaninovic, M. Hayami, T. Miura, J. Imai, S. Sonoda, H. Shiraki, K. Miyamoto, and K. Tajima. 1993. Southernmost carriers of HTLV-I/II in the world. *Jpn. J. Cancer Res.* **84**:1–3.
- Dube, D. K., M. P. Sherman, N. K. Saksena, V. Bryz-Gornia, J. Mendelson, J. Love, C. B. Arnold, T. Spicer, S. Dube, J. B. Glaser, A. E. Williams, M. Nishimura, S. Jacobsen, J. F. Ferrer, D. N. Pino, S. Quiruelas, and B. J. Poiesz. 1993. Genetic heterogeneity in human T-cell leukemia/lymphoma virus type II. *J. Virol.* **67**:1175–1184.
- Duenas-Barajas, E., J. E. Bernal, D. R. Vaught, V. R. Nerurkar, P. Sarmiento, R. Yanagihara, and D. C. Gajdusek. 1993. Human retroviruses in Amerindians of Colombia: high prevalence of human T cell lymphotropic virus type II infection among the Tunebo Indians. *Am. J. Trop. Med. Hyg.* **49**:657–663.
- Echeverria de Perez, G., M. Leon-Ponte, O. Noya, C. Botto, D. Gallo, and N. Bianco. 1993. First description of endemic HTLV-II infection among Venezuelan Amerindians. *J. Acquired Immune Defic. Syndr.* **6**:1368–1372.
- Eiraku, N., C. Monken, T. Kubo, S. W. Zhu, M. Rios, C. Bianco, B. Hjelte, K. Nagashima, and W. W. Hall. 1995. Nucleotide sequence and restriction fragment length polymorphism analysis of the long terminal repeat of human T cell leukemia virus type II. *AIDS Res. Hum. Retroviruses* **11**:625–636.
- Felsenstein, J. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.
- Ferrer, J. F., N. Del Pino, E. Esteban, M. P. Sherman, S. Dube, D. K. Dube, M. A. Basombrio, E. Pimentel, A. Segovia, S. Quiruelas, and B. J. Poiesz. 1993. High rate of infection with the human T-cell leukemia retrovirus type II in four Indian populations of Argentina. *Virology* **197**:576–584.
- Froment, A., E. Delaporte, M. C. Dazza, and B. Larouze. 1993. HTLV-II among pygmies from Cameroon. *AIDS Res. Hum. Retroviruses* **9**:707.
- Fuchs, R. 1990. Free molecular biological software available from the EMBL file server. *Comput. Appl. Biosci.* **6**:120–121.
- Fukushima, Y., H. Takahashi, W. W. Hall, T. Nakasone, S. Nakaka, P. Song, D. D. Duc, B. Hien, N. X. Quang, T. N. Trinh, K. Nishioka, K. Kitamura, K. Komuro, A. Vahlne, and M. Honda. 1995. Extraordinary high rate of HTLV type II seropositivity in intravenous drug abusers in South Vietnam. *AIDS Res. Hum. Retroviruses* **11**:637–644.
- Fujiyama, C., T. Fujiyoshi, T. Miura, S. Yashiki, D. Matsumoto, V. Zaninovic, O. Blanco, W. Harrington, Jr., J. J. Byrnes, M. Hayami, K. Tajima, and S. Sonoda. 1993. A new endemic focus of human T lymphotropic virus type II carriers among Orinoco natives in Colombia. *J. Infect. Dis.* **168**:1075–1077.
- Gessain, A., C. Fretz, M. Koulibaly, M. L. Boudret, A. Bah, M. Raphael, G. de The, and J. J. Fournel. 1993. Evidence of HTLV-II infection in Guinea, West Africa. *J. Acquired Immune Defic. Syndr.* **6**:324–325.
- Gessain, A., P. Tuppin, M. Kazanji, J.-V. Cosnefroy, M.-C. Georges-Courbot, A. J. Georges, and G. de The. 1994. A distinct molecular variant of HTLV-IIb in Gabon, Central Africa. *AIDS Res. Hum. Retroviruses* **10**:753–755.
- Gessain, A., P. Maulele, A. Froment, M. Biglione, J. Y. Le Hesran, F. Tekiai, J. Millan, and G. de Thè. 1995. Isolation and molecular characterization of a human T cell lymphotropic virus type II (HTLV-II), subtype B from a healthy pygmy living in a remote area of Cameroon: an ancient origin for HTLV-II in Africa. *Proc. Natl. Acad. Sci. USA* **92**:4041–4045.
- Giri, A., P. Markham, L. Digilio, G. Hurteau, R. C. Gallo, and G. Franchini. 1994. Isolation of a novel simian T-cell lymphotropic virus from *Pan paniscus* that is distantly related to human T-cell leukemia/lymphotropic types I and II. *J. Virol.* **68**:8392–8395.
- Goubau, P., J. Desmyter, J. Ghesquire, and B. Kaseraka. 1992. HTLV-II infection among the pygmies. *Nature (London)* **359**:201.
- Goubau, P., H. F. Liu, G. G. deLarge, A.-M. Vandamme, and J. Desmyter. 1993. HTLV-II seroprevalence in pygmies across Africa since 1970. *AIDS Res. Hum. Retroviruses* **9**:709–713.
- Goubau, P., M. Van Brussel, A. M. Vandamme, H. F. Liu, and J. Desmyter. 1994. A primate T lymphotropic virus, PTLV-L, different from human T lymphotropic viruses types I and II in a wild caught baboon (*Papio hamadryas*). *Proc. Natl. Acad. Sci. USA* **91**:2848–2852.
- Hall, W. W. Unpublished data.
- Hall, W. W., T. Kubo, S. Ijichi, H. Takahashi, and S. W. Zhu. 1994. Human T cell leukemia/lymphoma virus, type II (HTLV-II): emergence of an important newly recognized pathogen. *Semin. Virol.* **5**:165–178.
- Hall, W. W., H. Takahashi, C. Liu, M. H. Kaplan, O. Scheewind, S. Ijichi, K. Nagashima, and R. C. Gallo. 1992. Multiple isolates and characteristics of human T-cell leukemia virus type II. *J. Virol.* **66**:2456–2463.
- Hall, W. W., S. W. Zhu, P. Horal, Y. Furuta, G. Zagaany, and A. Vahlne. 1994. HTLV-II infection in Mongolia. Abstract 2. *AIDS Res. Hum. Retroviruses* **10**:443.

25. Hjelle, B., R. Scalf, and S. Swenson. 1990. High frequency of human T cell leukemia-lymphoma virus type II infection in New Mexico blood donors: determination by sequence specific oligonucleotide hybridisation. *Blood* **76**: 450-454.
26. Hjelle, B., S. W. Zhu, H. Takahashi, S. Ijichi, and W. W. Hall. 1993. Endemic human T cell leukemia virus type II infection in southwestern US Indians involves two prototype variants of virus. *J. Infect. Dis.* **168**:737-740.
27. Hollsberg, P., and D. A. Hafler. 1993. Seminars in medicine of the Beth Israel Hospital, Boston. Pathogenesis of diseases induced by human lymphotropic virus type I infection. *N. Engl. J. Med.* **328**:1173-1182. (Review.)
28. Igarashi, T., M. Yamashita, T. Miura, M. Osei-Kwasi, N. K. Aysi, H. Shiraki, T. Kurimura, and M. Hayami. 1993. Isolation and genomic analysis of human T lymphotropic virus type II from Ghana. *AIDS Res. Hum. Retroviruses* **9**:1039-1042.
29. Ijichi, S., K. Tajima, V. Zaninovic, F. E. Leon, Y. Katahira, S. Sonoda, T. Miura, M. Hayami, and W. W. Hall. 1993. Identification of human T cell leukemia virus type IIb infection in the Wayu, an aboriginal population of Colombia. *Jpn. J. Cancer Res.* **84**:1215-1218.
30. Ishak, R., W. J. Harrington, Jr., V. N. Azededo, N. Eiraku, M. O. G. Ishak, J. F. Guerreiro, S. B. Santos, T. Kubo, C. Monken, S. Alexander, and W. W. Hall. 1995. Identification of human T cell lymphotropic virus type IIa infection in the Kayapo, an indigenous population of Brazil. *AIDS Res. Hum. Retroviruses* **11**:813-821.
31. Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**:111-120.
32. Kwok, S. 1990. Procedures to minimize PCR product carry over, p. 142-145. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and J. J. White (ed.), *PCR protocols*. Academic Press, New York.
33. Lairmore, M. D., S. Jacobson, F. Gracia, B. K. De, L. Castillo, M. Lareategui, B. D. Roberts, P. H. Levine, W. A. Blattner, and J. E. Kaplan. 1990. Isolation of human T-cell lymphotropic virus type 2 from Guaymí Indians in Panama. *Proc. Natl. Acad. Sci. USA* **87**:8840-8844.
34. Lee, H., K. B. Idler, P. Swanson, J. J. Aparicio, K. K. Chin, J. P. Lax, M. Nguyen, T. Mann, G. Leckie, A. Zanetti, G. Marinucci, I. S. Chen, and J. D. Rosenblatt. 1993. Complete nucleotide sequence of HTLV-II isolate NRA: comparison of envelope sequence variation of HTLV-II isolates from U.S. blood donors and U.S. and Italian i.v. drug users. *Virology* **196**:57-69.
35. Levine, P. H., S. Jacobson, R. Elliott, A. Cavallero, G. Colclough, C. Dorry, C. Stephenson, R. M. Knigge, J. Drummond, M. Nishimura, M. E. Taylor, S. Wiktor, and G. M. Shaw. 1993. HTLV-II infection in Florida Indians. *AIDS Res. Hum. Retroviruses* **9**:123-127.
36. Liu, H. F., A. M. Vandamine, M. Van Brussel, J. Desmyter, and P. Goubau. 1994. New retroviruses in human and simian T lymphotropic viruses. *Lancet* **344**:265-266.
37. Maloney, E. M., R. J. Biggar, J. V. Neel, M. E. Taylor, B. H. Hahn, G. M. Shaw, and W. A. Blattner. 1992. Endemic human T cell lymphotropic virus type II infection among isolated Brazilian Amerindians. *J. Infect. Dis.* **166**: 100-107.
38. Mauclere, P., R. Mahieux, J. M. Garcia-Calieja, R. Salla, F. Tekaia, J. Millan, G. de Thè, and A. Gessain. 1995. A new HTLV-II subtype A isolate in an HIV type I infected prostitute from Cameroon, Central Africa. *AIDS Res. Hum. Retroviruses* **11**:989-993.
39. Neel, J. V., R. J. Biggar, and R. I. Sukernik. 1994. Virologic and genetic studies relate Amerind origins to the indigenous people of the Mongolia/Manchuria/Southeastern Siberia region. *Proc. Natl. Acad. Sci. USA* **91**: 10737-10741.
40. Pardi, D., J. E. Coligan, and R. B. Lal. 1995. Mapping of linear epitopes of the regulatory proteins of human T cell lymphotropic virus (HTLV) type II: identification of an HTLV-IIb restricted epitope. *J. Infect. Dis.* **172**:554-557.
41. Pardi, D., B. Hjelle, T. M. Folks, and R. B. Lal. 1995. Genotypic characteristics of HTLV-II isolates from American and non-Indian populations. *Virus Genes* **10**:27-35.
42. Pardi, D., J. E. Kaplan, J. E. Coligan, T. M. Folks, and R. B. Lal. 1993. Identification and characterization of an extended Tax protein in human T-cell lymphotropic virus type II subtype b isolates. *J. Virol.* **67**:7663-7667.
43. Pardi, D., W. M. Switzer, K. G. Hadlock, J. E. Kaplan, R. B. Lal, and T. M. Folks. 1993. Complete nucleotide sequence of an Amerindian human T-cell lymphotropic virus type II (HTLV-II) isolate: identification of a variant HTLV-II subtype b from a Guaymí Indian. *J. Virol.* **67**:4659-4664.
44. Poiesz, B., M. Sherman, N. Saksena, D. Dube, S. Dube, J. Gavalchin, N. Fan, M. Lane, and B. Paul. 1993. The biology and epidemiology of the human T-cell lymphoma/leukemia viruses, p. 189-205. *In* H. C. Neu, J. A. Levy, and R. A. Weiss (ed.), *Focus on HIV. Frontiers of infectious diseases*. Churchill-Livingston, London.
45. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406-425.
46. Seiki, M., S. Hattori, Y. Hirayama, and M. Yoshida. 1983. Human adult T cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA. *Proc. Natl. Acad. Sci. USA* **80**:3618-3622.
47. Shimotohno, K., D. W. Golde, M. Miwa, T. Sugimura, and I. S. Y. Chen. 1984. Nucleotide sequence analysis of the long terminal repeat of human T-cell leukemia virus type II. *Proc. Natl. Acad. Sci. USA* **81**:1079-1083.
48. Shimotohno, K., Y. Takahashi, N. Shimizu, T. Gojobori, D. W. Golde, I. S. Chen, M. Miwa, and T. Sugimura. 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.
49. Shimotohno, K., M. Takano, T. Teruuchi, and M. Miwa. 1986. Requirement of multiple copies of a 21 nucleotide sequence in the U3 regions of human T cell leukemia virus type I and II long terminal repeats for trans-acting activation of transcription. *Proc. Natl. Acad. Sci. USA* **83**:8112-8116.
50. Stoeck, P. J., and R. Omond. 1989. The EMBL file server. *Nucleic Acids Res.* **17**:6763-6764.
51. Switzer, W. M., S. M. Owen, P. A. Pieniazek, V. R. Nerurkar, E. Duenas-Barajas, W. Heniene, and R. B. Lal. 1995. Molecular analysis of human T cell lymphotropic virus type II from Wayu Indians of Columbia demonstrates two subtypes of HTLV-IIb. *Virus Genes* **10**:153-162.
52. Switzer, W. M., D. Pieniazek, P. Swanson, H. H. Samdal, V. Soriano, R. F. Khabbaz, J. E. Kaplan, R. B. Lal, and W. Heniene. 1995. Phylogenetic relationship and geographic distribution of multiple human T-cell lymphotropic virus type II subtypes. *J. Virol.* **69**:621-632.
53. Takahashi, H., S. W. Zhu, S. Ijichi, A. Vahlne, H. Suzuki, and W. W. Hall. 1993. Nucleotide sequence analysis of human T cell leukemia virus, type II (HTLV-II) isolates. *AIDS Res. Hum. Retroviruses* **9**:721-732.
54. Takatsuki, K., Y. Hinuma, and M. Yoshida. 1992. Advances in adult T cell leukemia and HTLV-I research. *Monogr. Cancer Res.* **39**:261.
55. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673-4680.
56. Zaninovic, V., F. Sanzon, F. Lopez, G. Velandia, A. Blank, M. Blank, C. Fujiyama, S. Yashiki, D. Matsumoto, Y. Katahira, H. Miyashita, T. Fujiyoshi, L. Chan, T. Sawada, T. Miura, M. Hayami, K. Tajima, and S. Sonoda. 1994. Geographic independence of HTLV-I and HTLV-II foci in the Andes Highland, the Atlantic Coast, and the Orinoco of Columbia. *AIDS Res. Hum. Retroviruses* **10**:97-105.
57. Zella, D., A. Cavicchini, M. Salemi, C. Casoli, F. Lori, G. Achilli, E. Can-raneo, V. Landini, and U. Bertazzoni. 1993. Molecular characterization of two isolates of human T cell leukemia virus type II from Italian drug abusers and comparison of genome structure with other isolates. *J. Gen. Virol.* **74**:437-444.