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,1 ORLANDO DA COSTA FERREIRA,2 SHI WEI ZHU,1 ROSEMARIE LORENCO,2 MARLUIA AZVEDO,1 JOAO GUERREIRO,4 MARIA POMBO DE OLIVEIRA,5 PAULA LOUREIRO,6 NELSON HAMMERSCHLAK,7 SHINJI IICHI,8 AND WILLIAM W. HALL1*

Laboratory of Medical Virology, The Rockefeller University, New York, New York 10021; Emilo Ribas Institute of Infectious Diseases8 and Hospital Israelita Albert Einstein, São Paulo, Federal University of Para, Belem, National Cancer Institute, Rio de Janeiro, and Hemope, Recife;4 Brazil; Department of Medicine, Thomas Jefferson University, Philadelphia, Pennsylvania 19107; 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.

DNA 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-2 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

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.). Active samples were analyzed by Western blotting (HTLV Blot 2.3, Diagnostics Biotechnology, Singapore) to confirm seropositivity and to differentiate HTLV-1 and HTLV-2 infections. The origin of the sequences obtained from the Kayapo Indian group have been previously described (30).

Peripheral blood mononuclear cells from infected individuals were separated by Ficol-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 the LTR, nested PCR reactions were performed in 100-μl volumes containing 1 μg of extracted DNA, 225 μM each deoxynucleoside triphosphate, 100 pmol of each primer, 50 mM KCl, 2.0 mM MgCl2, 10 mM Tris-HCl (pH 8.3), and 0.5 μU of Taq polymerase (Perkin-Elmer Cetus Corp., Norwalk, Conn.). Primer pairs used in the first round were 5′-CTTACCACATCCTCCACTGCGGTA-3′ and 5′-GGGAAAGCCCTCAGAATTTTGA-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′-AAAGCCGAAAGAGGGACTGAA-3′ and 5′-ATCCGAGAACCCACCTGTTT-3′, corresponding to nt 290 and 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 carried out as described in the first-round reaction, and nested PCR reactions were performed in 100-μl volumes containing 1 μg of DNA, 225 μM each deoxynucleoside triphosphate, 100 pmol of each nested primer, and 50 mM KCl. The 35 cycles were followed by a 10-min extension at 72°C. Amplification and nucleotide sequencing of the entire pX region, nested PCR was carried out as described for the first-round reaction, and two overlapping primer pairs 5′-TGAGTCTTACCCCCCTGCCCATAA-3′ and 5′-ATCCGGACGAGCCCCCACTTGTTT-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 nested PCR product used in 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 thermal cycler (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 env region, nested PCR was carried out as described in the first-round reaction, and two overlapping primer pairs 5′-AAGTTCTTCTAATCGTTTTAG-3′ and 5′-GTTTCAGTTCCTA-3′ (FITCH) and a maximum-likelihood analysis (DNAML). The analysis included the 8 Brazilian IVDA sequences and 25 other HTLV-2 LTR sequences, 18 sequences from the published database, and 7 sequences from the HTLV-1A gag database of this laboratory. Sequences were identified as HTLV-2a and 10 identified as HTLV-2b (8). The latter included the two prototype subtypes of HTLV-2a and -b, MoT (48) and NPA (34); 48, is an isolate from a Guaymi Indian from Panama; NPA, from Colombia (51); PyCam1, 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 (25).

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 g-globin gene, a human initiation codon, rabbit β-globin polyadenylation signals, and the replication origin of simian virus 40. To prepare the HTLV-2 Tax protein (Tax), the extended Tax protein representative of the HTLV-2b isoforms, and isolated deletion mutants of these proteins, PCR was used to amplify the corresponding regions with primers containing XhoI and BamHI restriction endonuclease sites in the 5′ and 3′ ends, respectively. Primers used in the PCR were as follows. An upstream primer 5′-CCCGGCTTCTGAGACATTCCACCTCAAGGCAGATGAG-3′ (XhoI 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′-GTAACCTGCACTAGTCTTTCACTCACAAGGATCCTGGA-3′ and 5′-CGCGGCGATCTTACCTGGAGTTGTTG-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′-TCTGACTAGTCTTTCACTCACAAGGATCCTGGA-3′ and 5′-CGCGGCGATCTTACCTGGAGTTGTTG-3′, corresponding to nucleotides nt 7683 to 7703 and 7791 to 7771. PCR products were sequenced in both directions, using an automated sequencing apparatus (Promega). Nucleotide sequences were initially purified by using Wizard PCR Prep (Promega), and 10 μl of the purified DNA was directly digested with 5 U of AvaII, BanII, BglII, Bsu36I, MseI, and Smal (Boehringer Mannheim, Indianapolis, Ind.) in 20 μl of reaction mix, containing reactions 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 amplification of the env, LTR, and pX regions were excised and ligated 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 automated sequencer apparatus (Promega). 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 (55) 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 using the PROTDIST program. The other groups obtained with FITCH, 1000 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 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 sequences from the published database, and 7 sequences from the HTLV-1A gag database of this laboratory. The origin of these sequences was previously described (30).

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 (55) 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 using the PROTDIST program. The other groups obtained with FITCH, 1000 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 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 sequences from the published database, and 7 sequences from the HTLV-1A gag database of this laboratory. The origin of these sequences was previously described (30).

Transfections. Jurkat cells were transfected with the pCG Tax constructs and pHTLV-II-LTR-CAT (originally derived from pVS, CAT and kindly provided by K. Shimotomoto, National Cancer Center Research Institute, Tokyo, Japan), using Lipofectin (Promega) as described by the manufacturer. One microgram of each plasmid was used to transfet 5 × 106 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. Cells were washed twice in ice-cold phosphate-buffered saline, resuspended in 300 μl of Reporter lysis buffer (Promega), and incubated at 60°C for 15 min to inactivate endogenous deacetylase activity. After centrifugation, supernatants (100 μl) were incubated with 5 μl of [3H]chloramphenicol and 5 μl of p-nitrophenylacetate, which were carried out as previously (8). PCR products were initially purified by using Wizard PCR Prep (Promega), and 10 μl of the purified DNA was directly digested with 5 U of AvaII, BanII, BglII, Bsu36I, MseI, and Smal (Boehringer Mannheim, Indianapolis, Ind.) in 20 μl of reaction mix, containing reactions recommended by the manufacturer. Products were electrophoresed on 3.0% agarose gels and visualized by ethidium bromide staining.
HTLV-2 separate assays. Phenylacetyltransferase (CAT) activity was averaged from a minimum of three experiments. Scrapeouts were scraped from the plate and counted by liquid scintillation. Chloramphenicol acetyltransferase (CAT) activity was averaged from a minimum of three separate assays.

**RESULTS**

RFLP and nucleotide sequence analysis of the *env* region. RFLP analysis using *Xhol* 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 prototype 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 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, 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). Additional 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 Guaymi Indian, and PH230CAM 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.
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.
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 a1 to a4 (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 AvaI 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 AvaII 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 endonucleases AvaII, BanII, MseI, Bsu36I, SmaI, and BglI. 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 BanII 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 BanII 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
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.

<table>
<thead>
<tr>
<th>HTLV-2b (Mo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
</tr>
</tbody>
</table>
FIG. 4—Continued.
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-
Extended 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.

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 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
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 prototype 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 viral 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 support 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 IVDAAs 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 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 among pygmies in Cameroon (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


