JOURNAL OF VIROLOGY, Aug. 2006, p. 7427–7438 Vol. 80, No. 15
0022-538X/06/$08.00+0 doi:10.1128/JVI.00690-06

Ancient Origin and Molecular Features of the Novel Human T-Lymphotropic Virus Type 3 Revealed by Complete Genome Analysis

William M. Switzer,[1]* Shoukat H. Qari,[1] Nathan D. Wolfe,[2,3,4] Donald S. Burke,[2,3] Thomas M. Folks,[1] and Walid Heneine[1]

Laboratory Branch, Division of HIV/AIDS Prevention, National Center for HIV, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia 30333,[1] and Departments of Epidemiology,[2] International Health,[3] and Molecular Microbiology and Immunology,[4] Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland 21205

Received 5 April 2006/Accepted 8 May 2006

Human T-lymphotropic virus type 3 (HTLV-3) is a new virus recently identified in two primate hunters in Central Africa. Limited sequence analysis shows that HTLV-3 is distinct from HTLV-1 and HTLV-2 but is genetically similar to simian T-lymphotropic virus type 3 (STLV-3). We report here the first complete HTLV-3 sequence obtained by PCR-based genome walking using uncultured peripheral blood lymphocytes from an HTLV-3-infected person. The HTLV-3(2026ND) genome is 8,917 bp long and is genetically equidistant from HTLV-1 and HTLV-2, sharing about 62% identity. Phylogenetic analysis of all gene regions confirms this relationship and shows that HTLV-3 falls within the diversity of STLV-3, suggesting a primate origin. However, HTLV-3(2026ND) is unique, sharing only 87% to 92% sequence identity with STLV-3. SimPlot and phylogenetic analysis did not reveal any evidence of genetic recombination with either HTLV-1, HTLV-2, or STLV-3. Molecular dating estimates that the ancestor of HTLV-3 is as old as HTLV-1 and HTLV-2, with an inferred divergence time of 36,087 to 54,067 years ago. HTLV-3 has a prototypic genomic structure, with all enzymatic, regulatory, and structural proteins preserved. Like STLV-3, HTLV-3 is missing a third 21-bp transcription element found in the long terminal repeats of HTLV-1 and HTLV-2 but instead contains a unique activator protein-1 transcription factor upstream of the 21-bp repeat elements. A PDZ motif, like that in HTLV-1, which is important for cellular signal transduction and transformation, is present in the C terminus of the HTLV-3 Tax protein. A basic leucine zipper region located in the antisense strand of HTLV-1, believed to play a role in viral replication and oncogenesis, was also found in the complementary strand of HTLV-3. The ancient origin of HTLV-3, the broad distribution of STLV-3 in Africa, and the propensity of STLVs to cross species into humans all suggest that HTLV-3 may be prevalent and support the need for expanded surveillance for this virus.

Deltaretroviruses are a diverse group of human and simian T-lymphotropic viruses (HTLV and STLV, respectively) that until recently were composed of only two distinct human groups called HTLV types 1 and 2 (HTLV-1 and -2) (1, 13, 25, 26, 31, 43, 46). Phylogenetic analysis of simian T-lymphotropic virus type 1 (STLV-1) and global HTLV-1 sequences suggests that different STLV-1s were introduced into humans multiple times in the past (1, 13, 15, 25, 26, 31, 45). By convention these viruses are called HTLVs when found in humans, regardless of their suspected zoonotic origin (1, 13, 15, 25, 26, 31, 46). Although an STLV-2 has been identified in two troops of captive bonobos (*Pan paniscus*), the zoonotic relationship of this virus to HTLV-2 is less clear (10, 40). Recently, two new HTLVs, HTLV-3 and HTLV-4, were identified in primate hunters in Cameroon, expanding substantially the diversity of deltaretroviruses in humans. While the prevalence of HTLV-3 and HTLV-4 is currently not known, the independent identification of another HTLV-3 in Cameroon suggests that this infection may not be infrequent (7, 43).

Limited sequencing of small gene regions of the HTLV-3 genome showed that it shared the most phylogenetic relatedness to STLV-3, found in West-Central Africa, and therefore represents the first human virus in this lineage (43). It is now recognized that STLV-3, originally isolated from a captive hamadryas baboon (*Papio hamadryas*) (14) more than 10 years ago, has a wide geographic distribution across Africa, infecting many nonhuman primates (NHPs), including Ethiopian gelada (*Theropithecus gelada*), sacred baboons (*P. hamadryas*), hybrid baboons (*P. hamadryas* × *Papio anubis* hybrid) (35, 42), and Senegalese olive baboons (*Papio papio*) (20), Cameroonian and Nigerian red-capped mangabeys (*Cercocebus torquatus torquatus*) (21, 22), and Cameroonian spot-nosed guenons and agile mangabeys (*Cercopithecus nictitans* and *Cercocebus agilis*, respectively) (9, 41). Collectively, members of the HTLV groups and their STLV analogues are called primate T-lymphotropic viruses (PTLV); PTLV-1, PTLV-2, and PTLV-3 are composed of HTLV-1 and STLV-1, HTLV-2 and STLV-2, and HTLV-3 and STLV-3, respectively. The PTLV-4 group currently has only one member, HTLV-4, since a simian counterpart has yet to be identified. Thus, the origin of HTLV-4 is unclear (43).

* Corresponding author. Mailing address: Laboratory Branch, Division of HIV/AIDS Prevention, National Center for HIV, STD, and TB Prevention, Centers for Disease Control and Prevention, 1600 Clifton Rd., MS G-45, Atlanta, GA 30333. Phone: (404) 639-0219. Fax: (404) 639-1174. E-mail: bis3@cdc.gov.

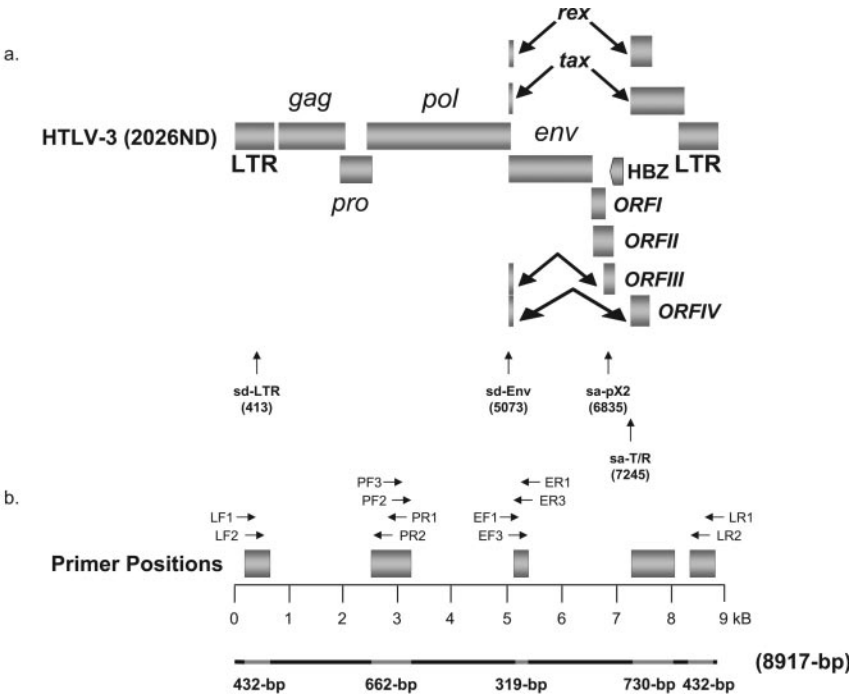


FIG. 1. (a) HTLV-3 genomic organization and (b) schematic representation of PCR-based genome-walking strategy. (a) Noncoding LTRs, coding regions for all major proteins (*gag*, group-specific antigen; *pro*, protease; *pol*, polymerase; *env*, envelope; *rex*, regulator of expression; *tax*, transactivator), HBZ, and 3′ genomic ORFs of unknown function are shown. Putative sd and sa sites are indicated. (b) Small provirus sequences (grey bars) were first amplified from each major gene region and the LTR by using generic primers as described in Materials and Methods. The complete provirus sequence was then obtained by using PCR primers located within each major gene region by genome walking as indicated by arrows and black bars.

Like human immunodeficiency virus (HIV), both HTLV-1 and HTLV-2 have spread globally and are human-pathogenic viruses (1, 13, 31, 46). HTLV-1 causes adult T-cell leukemia/lymphoma (ATL), HTLV-1 associated myelopathy/tropical spastic paraparesis (HAM/TSP), and other inflammatory diseases in <5% of those infected (13, 31, 46). HTLV-2 is less pathogenic than HTLV-1 and has been associated with a neurologic disease similar to HAM/TSP (1). The recent identification of HTLV-3 and HTLV-4 in only three persons limits an evaluation of the disease potential and secondary transmissibility of these novel viruses, which will require longitudinal epidemiologic studies (7, 43). However, complete genomic sequences of these viruses can provide basic information on their genetic structure and on whether important functional motifs involved in viral expression and HTLV-induced leukemogenesis are preserved (5, 11, 12, 24, 38, 47). In addition, determination of the viral sequence will be important for understanding the evolution and genetic relatedness of HTLV-3 to known HTLVs and STLVs and for development of improved diagnostic assays to better understand the epidemiology of these novel human viruses.

In this paper, we report the first full-length sequence of HTLV-3 and demonstrate that this virus is highly divergent and distinct from HTLV-1 and HTLV-2. We show that HTLV-3 clearly falls within the diversity of STLV-3 but is unique, sharing only 87 to 92% genomic identity. The observed low nucleotide substitution rate and conserved genomic structure of HTLV-3 demonstrate the genetic stability of this virus. Taken together, the finding that HTLV-3 is as old as the ancestor of HTLV-1 and HTLV-2 (25), the identification of a second HTLV-3 in Cameroon (7), and the wide distribution of STLVs across Africa (14, 20–22, 35, 42) suggest that HTLV-3 may be more common than previously imagined. We also found molecular features in HTLV-3 that are more similar to HTLV-1 than to HTLV-2, suggesting a pathogenic potential in HTLV-3-infected persons like that observed in HTLV-1 infection.

MATERIALS AND METHODS

DNA preparation and PCR-based genome walking. DNA was prepared from uncultured peripheral blood mononuclear cells available from a person designated 2026ND, identified in the original PTLV surveillance study in Cameroon reported in detail elsewhere (33, 44). DNA integrity was confirmed by β-actin PCR as previously described (43). All DNA preparation and PCR assays were performed in a laboratory where only human specimens are processed and tested according to recommended precautions to prevent contamination. To obtain the full-length genomic sequence of HTLV-3, we first PCR amplified small regions of each major coding region by using nested PCR and degenerate PTLV primers (Fig. 1). The *tax* (577-bp) and polymerase (*pol*) (709-bp) sequences were amplified by using primers and conditions provided elsewhere (43). Envelope (*env*) (371-bp) sequences were amplified using standard PCR conditions of 45°C annealing with the external primers PGENVF1 [5′-TGGATCCCGTGG(A/C)GI(C/T)TCCTIAA-3′] and PGENVR1 [5′-GT(A/G)TAIG(C/G)(A/G)(C/G)AIGTCCAIG(A/C)(T/C)TGG-3′] and the internal primers PFENVF2 [5′-AIAGACC(T/A)(C/T)CAAC(A/T)CCATGGGTAA-3′] and PGENVR2 [5′-G(A/C)(T/C)TGGCAICCIA(A/G)GTAIGGGCA-3′]. A 398-bp fragment of the long terminal repeat (LTR) was obtained using conserved STLV-3 primers as previously reported (43).

HTLV-3(2026ND)-specific primers were then designed from sequences obtained in each of the four viral regions described above and were used in nested, long-template PCRs to fill in the gaps in the genome as depicted in Fig. 1 by

TABLE 1. Percent nucleotide and amino acid identities of HTLV-3(2026ND) with other PTLV prototypes

Region	% Nucleotide (amino acid) identity of HTLV-3(2026ND) with the following PTLV (strain):						
	HTLV-1 (ATK)	HTLV-2 (MoT)	STLV-2 (PP1664)	STLV-3 (PH969)	STLV-3 (PPAF3)	STLV-3 (CTO604)	STLV-3 (NG409)
Genome	61.6	62.9	62.6	86.7	92.0	88.4	90.6
LTR	48.7	43.7	41.4	86.2	91.1	86.9	86.9
gag	69.3 (83.2)	69.4 (80.5)	70.6 (80.7)	86.4 (95.5)	91.3 (97.6)	89.4 (96.2)	90.6 (96.7)
p19	(74.4)	(68.3)	(67.2)	(95.9)	(95.9)	(95.9)	(94.3)
p24	(90.1)	(90.1)	(90.6)	(98.1)	(99.1)	(98.6)	(99.1)
p15	(78.0)	(73.8)	(72.6)	(88.4)	(96.5)	(90.7)	(94.2)
pro	59.7 (62.6)	59.2 (66.7)	59.4 (59.3)	83.3 (87.0)	88.8 (91.5)	85.0 (89.3)	88.0 (90.4)
pol	62.2 (66.2)	63.9 (71.2)	63.5 (69.9)	86.1 (92.7)	92.6 (94.9)	88.4 (92.9)	92.0 (92.9)
env	65.9 (73.8)	69.0 (78.2)	67.1 (77.4)	88.1 (95.1)	92.3 (95.1)	88.4 (94.3)	91.2 (95.3)
SU	(68.4)	(70.7)	(69.7)	(92.7)	(97.1)	(92.4)	(94.0)
TM	(83.5)	(91.6)	(91.0)	(99.4)	(98.9)	(97.8)	(97.8)
rex	76.9 (61.9)	76.3 (60.6)	75.8 (63.5)	87.1 (88.5)	90.9 (94.5)	88.5 (94.0)	88.3 (92.3)
tax	75.4 (81.4)	73.1 (83.4)	72.3 (80.4)	90.2 (97.4)	94.0 (98.3)	91.4 (96.6)	92.8 (96.9)

using the Expand High Fidelity kit containing both *Taq* and *Tgo* DNA polymerases (Roche). The external primer sequences for the LTR-*pol* fragment are 2026LF1 (5′-GGTAAGATCCCACTGGGTCGAGC-3′) and 2026PR1 (5′-GAAGCCAGGTCTCGGGTGACG-3′), and the internal primer sequences are 2026LF2 (5′-CGCTCCCCTGGAGCTCTCTCG-3′) and 2026PR2 (5′-GCCACTTCCCATTGGGCTTTTTGACGG-3′). The external primer sequences for the *pol-env* fragment are 2026PF3 (5′-GCTCTCACCGATAAAGTAACAAACG-3′) and 2026ER1 (5′-GGTAGGAAGAGGCTCCTATGAACAG-3′), and the internal primer sequences are 2026PF2 (5′-CAGGACTGCATAACATACGAGACCCTCC-3′) and 2026ER3 (5′-CCTATGAACAGGGTGCATCGACTGGG-3′). The external primer sequences used to obtain about 3 kb of the 3′ end of the genome (*env-tax*-LTR) are 2026EF1 (5′-CCTAAGCCCCCCATGTCCAGAC-3′) and 2026LR1 (5′-CGAGAGAGCTCCAGGGGAGCG-3′), and the internal primer sequences are 2026EF3 (5′-CCTACTCCCTGTATGTATTCCCCCATTGG-3′) and 2026LR2 (5′-GCTCGACCCAGTGGGATCTTACCGAGTGG-3′).

PCR products were revealed on 1.5% agarose gels stained with ethidium bromide, purified with a Qiaquick PCR purification kit (QIAGEN), and sequenced in both directions with a BigDye terminator cycle kit and automated sequencers (Applied Biosystems). Selected PCR products were also cloned into the pCR4-TOPO vector using the TOPO TA cloning kit (Invitrogen), and recombinant plasmid DNA was prepared using the QIAGEN plasmid purification kit prior to automated sequencing.

Sequence and phylogenetic analysis. Percent nucleotide divergence was calculated using the GAP program in the Genetics Computer Group (GCG) Wisconsin package (44). Functional genetic motifs involved in viral expression, regulation, and HTLV-induced oncogenesis were examined by detailed comparison of the HTLV-3 genome with full-length PTLV sequences (5, 11, 12, 24, 38). The secondary structure of the LTR RNA was determined using the RNAstructure program, version 4.2 (19). Sequences were aligned using the Clustal W program (36), gaps were removed, and distance-based trees were generated using the Kimura two-parameter model together with the neighbor-joining method in the MEGA program (version 2.1) and maximum-likelihood (ML) analysis in the PAUP* program as described in detail elsewhere (34, 43). The reliability of the final topology of the trees was tested with 1,000 bootstrap replicates. Comparison of full-length PTLV genomes available at GenBank and determination of genetic recombination were done using HTLV-3(2026ND) as the query sequence and the F84 (ML) model, with a transition/transversion ratio of 2.0, implemented in the SimPlot program (17). Splice acceptor (sa) and splice donor (sd) sites were predicted using an artificial neural network implemented in the NetGene2 program (available at the Web server www.cbs.dtu.dk/services/NetGene2).

The ancestor of HTLV-3(2026ND) was dated by aligning full-length genomes from prototypical PTLV available at GenBank with HTLV-3(2026ND) using Clustal W. Sequence gaps were removed, and minor adjustments in the alignment were made manually. LTR sequences were excluded from the analysis, since this region does not align accurately in PTLV. The best-fitting evolutionary model for the aligned sequences was determined using Modeltest, version 3.6 (23). The general time-reversible model, allowing six different substitution rate categories, with gamma-distributed rate heterogeneity (1.9724) and an estimated proportion of invariable sites (0.3687), was determined to best fit the data. Little substitution saturation was observed in the 7,213-bp alignment ($P < 0.0001$) using the DAMBE program; therefore, these sequences were determined to be satisfactory for use in phylogenetic analyses. Likewise, using the best-fitting evolutionary model defined above, good phylogenetic signal in the alignment was also found with likelihood mapping analysis using the Tree-Puzzle program, version 5.2.

The molecular clock hypothesis, or constant rate of evolution, was tested using the likelihood ratio test with likelihoods for the ML and clock-like ML trees obtained in PAUP*. The clock was tested with the best-fitting evolutionary model estimated in Modeltest, and ML trees were constructed in PAUP* starting from the neighbor-joining tree that is iteratively optimized using two consecutive heuristic searches with nearest-neighbor interchange followed by a final heuristic search with the tree-bisection-reconnection algorithm. To adjust for rate heterogeneity among different PTLV taxa, clock-like ML trees were transformed into ultrametric trees using the nonparametric rate smoothing (NPRS) algorithm in the TreeEdit program (version 1.0a10 carbon) (27). The branches of the NPRS tree were scaled using a divergence time of 40,000 to 60,000 years ago (YA) for the Melanesian HTLV-1 lineage (HTLV-1mel) based on genetic and archaeological evidence suggesting that ancestors of indigenous Melanesians and Australians migrated from Southeast Asia during this time (16, 25, 26). Variation in age estimates (branch lengths) was determined in PAUP* with 100 bootstrap repetitions by enforcing topological constraints and using a heuristic search without branch swapping on the clock-like ML tree. Branch lengths in all 100 trees were calibrated as before, and average divergence times and confidence intervals ($\alpha = 0.05$) were calculated in Excel. The evolutionary rate was estimated based on a known divergence time point of 40,000 to 60,000 YA and on branch lengths of the ML clock-like tree according to the following formula: evolutionary rate (r) = branch length (bl)/divergence time (t) (42).

Nucleotide sequence accession number. The HTLV-3(2026ND) provirus sequence has GenBank accession number DQ093792.

RESULTS

Comparison of the HTLV-3(2026ND) provirus genome with prototypical PTLV. Primers corresponding to small sequences in each of the three major genes of PTLV and in the LTR were used to generate the complete genome of HTLV-3(2026ND) as depicted in Fig. 1. Comparison of the resulting sequences with the genomes of other PTLV demonstrated that the complete provirus genome of HTLV-3(2026ND) is 8,917 bp. Comparison of HTLV-3(2026ND) with prototypical PTLV genomes demonstrates that this new human virus is equidistant from the HTLV-1 (61.6% identity) and PTLV-2 (62.9% identity) groups across the genome. Sequence analysis also confirms that HTLV-3 has the closest nucleotide and protein sequence identity to STLV-3 (87 to 92% identity) (Table 1). The greatest genetic divergence between the PTLV groups was seen in the LTR region (52 to 59%), while the greatest intergroup identity was observed within the highly conserved regu-

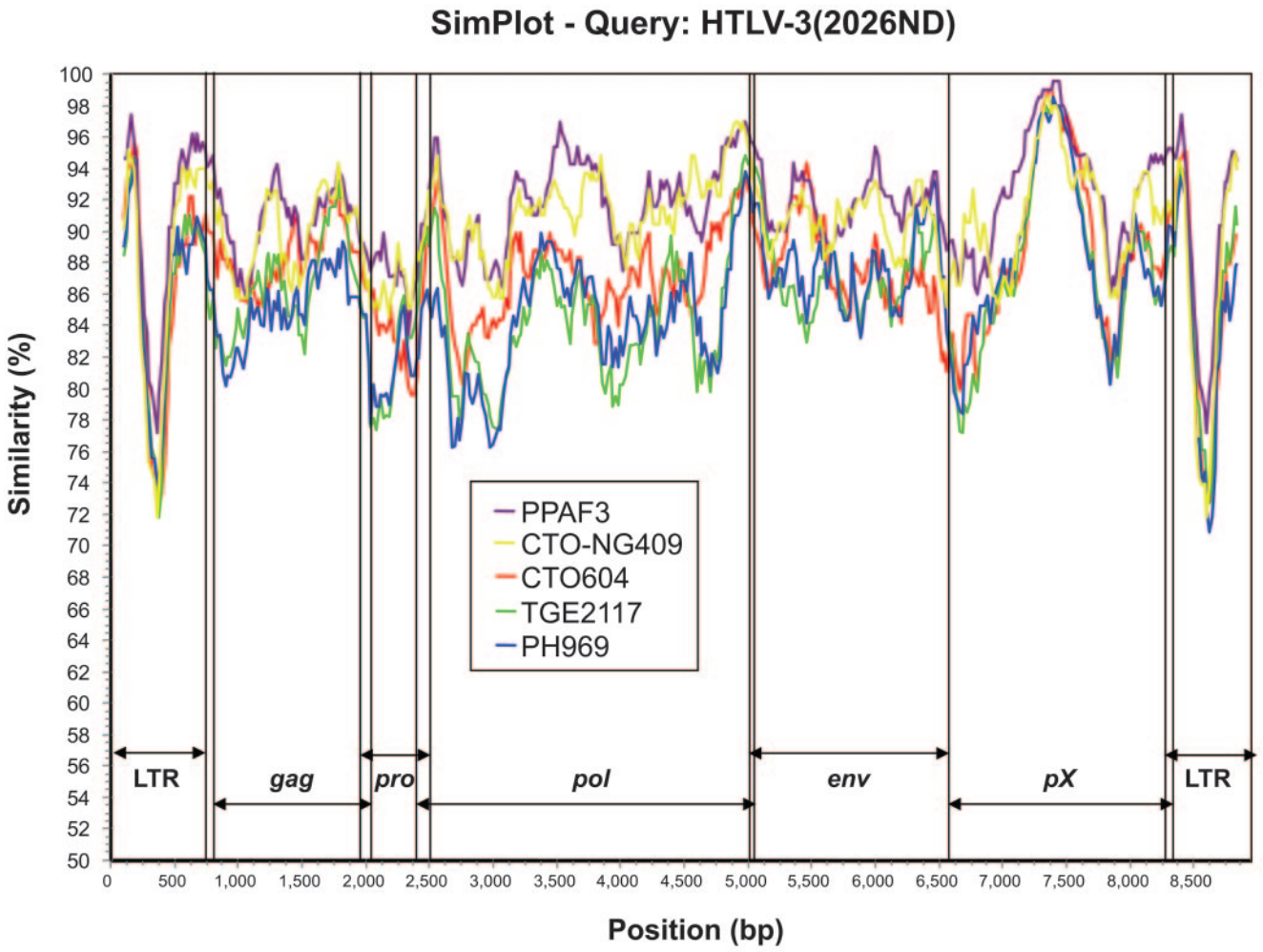


FIG. 2. Similarity plot analysis of the full-length HTLV-3(2026ND) and STLV-3 genomes using a 200-bp window size in 20-step increments on gap-stripped sequences. The F84 (maximum-likelihood) model was used with a transition-to-transversion ratio of 2.0.

latory genes, *tax* and *rex* (72 to 77%). Interestingly, within the PTLV-3 group, HTLV-3(2026ND), which was identified in a hunter from Cameroon, was unique but shared the greatest overall sequence identity with STLV-3(PPAF3) (92%) from a Senegalese baboon rather than STLV-3(CTO604) (88.4%), identified in red-capped mangabeys, also from Cameroon. This relationship was highlighted further by comparing HTLV-3(2026ND) with all available full-length STLV-3 genomes in a similarity plot analysis, where the highest identity was seen in the highly conserved *tax* gene (Fig. 2). As seen within other PTLV groups, no clear evidence of genetic recombination of HTLV-3(2026ND) with STLV-3 or STLV-1/HTLV-1 and STLV-2/HTLV-2 provirus sequences (data not shown) was observed using bootscanning analysis in the SimPlot program. We did not directly compare HTLV-3(2026ND) to the recently reported second strain of HTLV-3(Pyl43) because only two short sequences were available in GenBank and this virus has been shown to be nearly identical to STLV-3(CTO604) in these regions (7).

Phylogenetic analysis. The genetic relationship of HTLV-3(2026ND) to STLV-3 was confirmed by using aligned full-length prototype sequences excluding the LTR region (Fig. 3a). Phylogenetic analysis inferred three major PTLV groups with very high bootstrap support (100%), with HTLV-1, HTLV-2, and HTLV-3 each clustering in separate clades (Fig. 3a). Within the PTLV-3 phylogroup, HTLV-3(2026ND) formed a separate lineage but clustered with high bootstrap support with STLV-3 from West Central Africa (strains CTO604, CTO-NG409, and PPA-F3), suggesting a primate origin for this human infection. The relationship of HTLV-3 to STLV-3 was supported further by phylogenetic inference of identical tree topologies using an alignment of sequences from each major gene region (Fig. 3b to d). The phylogenetic stability seen across the PTLV genome also demonstrates further the absence of major recombination events occurring in PTLV despite evidence of dual infections in humans and primates (6, 9), compared to other retroviruses such as HIV, which undergo frequent recombination (37).

Dating the origin of HTLV-3(2026ND) and other PTLV. The finding of HTLVs in three distinct clades suggests an ancient, independent evolution of these viruses. Hence, additional molecular analyses were performed to estimate the divergence times of the PTLV lineages. Although others have reported finding a clock-like behavior of STLV-3 sequences (20–22, 25),

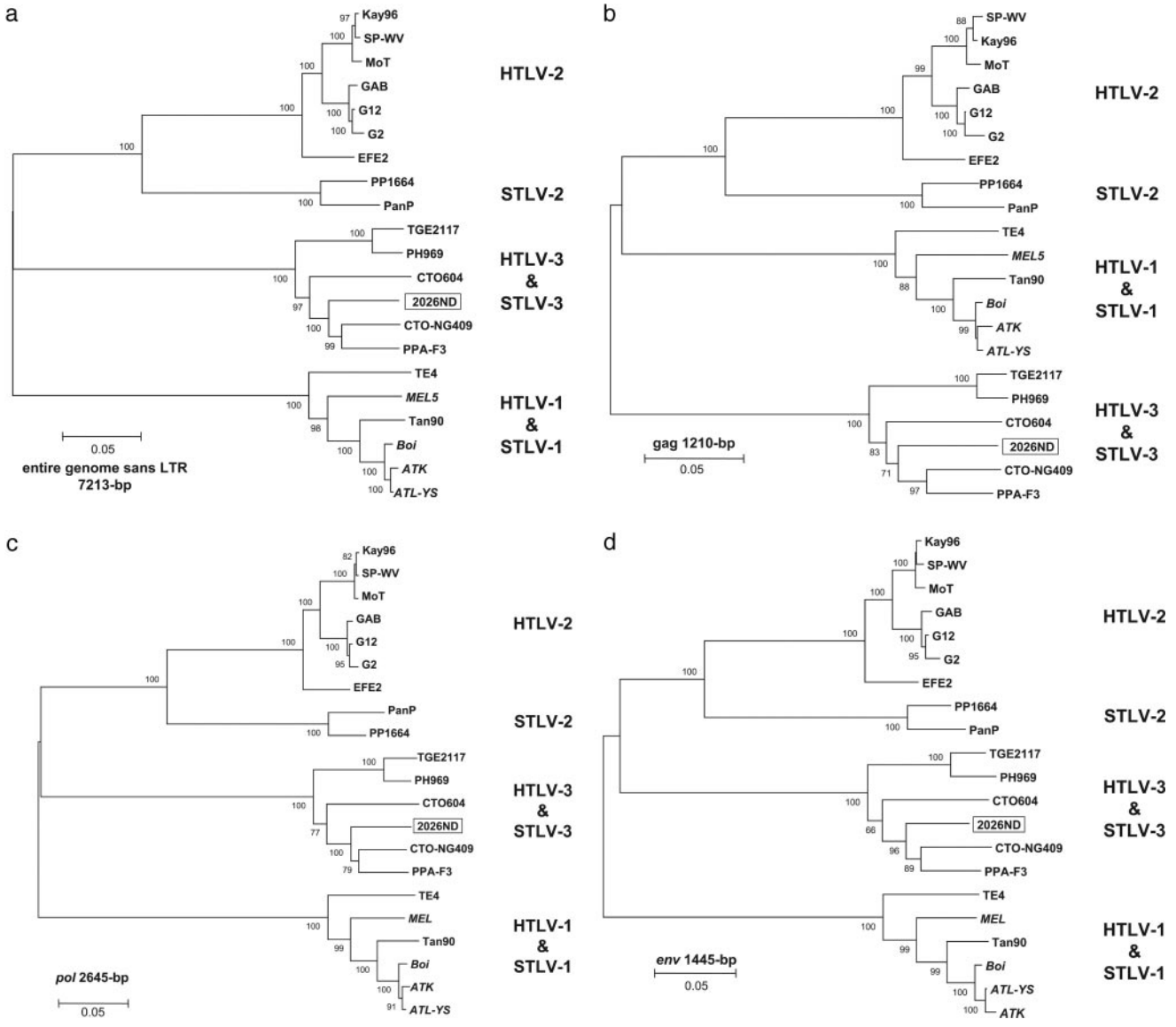


FIG. 3. Phylogenetic relationship of HTLV-3(2026ND) to other PTLVs. (a) Entire genome without LTR; (b) *gag*; (c), polymerase (*pol*); (d) envelope (*env*). Sequences generated in the current study are boxed; HTLV-1 sequences are italicized. Support for the branching order was determined by 1,000 bootstrap replicates; only values of 60% or more are shown. Branch lengths are proportional to the evolutionary distance (scale bar) between the taxa.

we were unable to confirm these results. Instead, an alignment of full-length PTLV genomes without LTR sequences suggested that PTLV evolved at different rates. However, reliable retrovirus divergence times can be obtained by using NPRS of the sequences to relax the stringency of a clock assumption, followed by time calibration of the tree using a value of 40,000 to 60,000 YA for the origin of the Melanesian HTLV-1 (27, 34, 42). By using these dates and methods, the mean evolutionary rate for PTLV was estimated to be 1.12×10^{-6} (confidence interval, 6.82×10^{-7} to 1.56×10^{-6}) substitutions/site/year, which is consistent with rates determined previously both with and without enforcing a molecular clock (16, 20–22, 25, 42). The mean evolutionary rate for HTLV-3(2026ND) is estimated to be 9.94×10^{-7} (confidence interval, 6.04×10^{-7} to 1.38×10^{-6}). The PTLV ancestor was estimated to have originated about 630,000 to 947,000 YA, confirming an ancient evolution of the primate deltaretroviruses (Fig. 4) (25). The separation of PTLV-1 and PTLV-2 occurred about 579,077 to 867,458 YA, while HTLV-2 and STLV-2 diverged around 191,621 to 286,730 YA (Fig. 4). The PTLV-3 progenitor was estimated to have appeared between 63,294 and 94,700 YA, with the ancestor of HTLV-3(2026ND) occurring about 36,087 to 54,067 YA, suggesting an ancient origin of this virus (Fig. 4).

Genomic organization of HTLV-3(2026ND) and identification of conserved functional motifs. Despite the fact that HTLV-3(2026ND) is genetically equidistant from HTLV-1 and HTLV-2, its genomic structure was similar to that of other PTLV and included the structural, enzymatic, and regulatory proteins all flanked by LTRs (Fig. 1). Like that of STLV-3, the HTLV-3(2026ND) LTR (697 bp) was smaller than those of

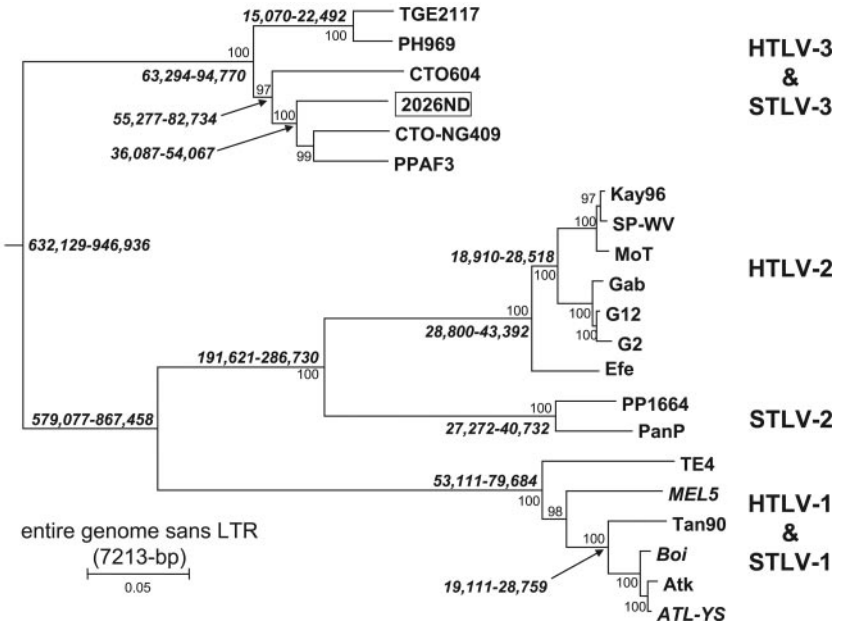


FIG. 4. Estimated divergence dates for the most recent common ancestor of HTLV-3(2026ND) and other PTLVs. Divergence dates are provided for each major node of a neighbor-joining tree rooted with PTLV-1 as the outgroup; estimates are provided as ranges using as calibration points 40,000 and 60,000 YA as the time of separation of the Melanesian HTLV-1 (MEL5) sequence from other PTLV-1 strains. Bootstrap analysis of 1,000 replicates is shown on the tree branches; only values of >60% are shown.

HTLV-1 (756 bp) and HTLV-2 (764 bp), by having two rather than the typical three 21-bp transcription regulatory repeat sequences in the U3 region of HTLV-1 and HTLV-2 (Fig. 5a) (19, 20, 38, 41). Other regulatory motifs, such as the polyadenylation signal, TATA box, and cap site, were all conserved in the HTLV-3(2026ND) LTR (Fig. 5a). Interestingly, an activator protein-1 (AP-1) site [TGA(C/G)T(C/A)A] was also identified at position 58 of the LTR upstream of the first 21-bp repeat element (Fig. 5a). The AP-1 site was found to be conserved in the LTR of STLV-3(PPAF3) and STLV-3(NG409) but is not present in other STLV-3s due to a G-to-A mutation at position 4 of the motif (data not shown). The AP-1 site is

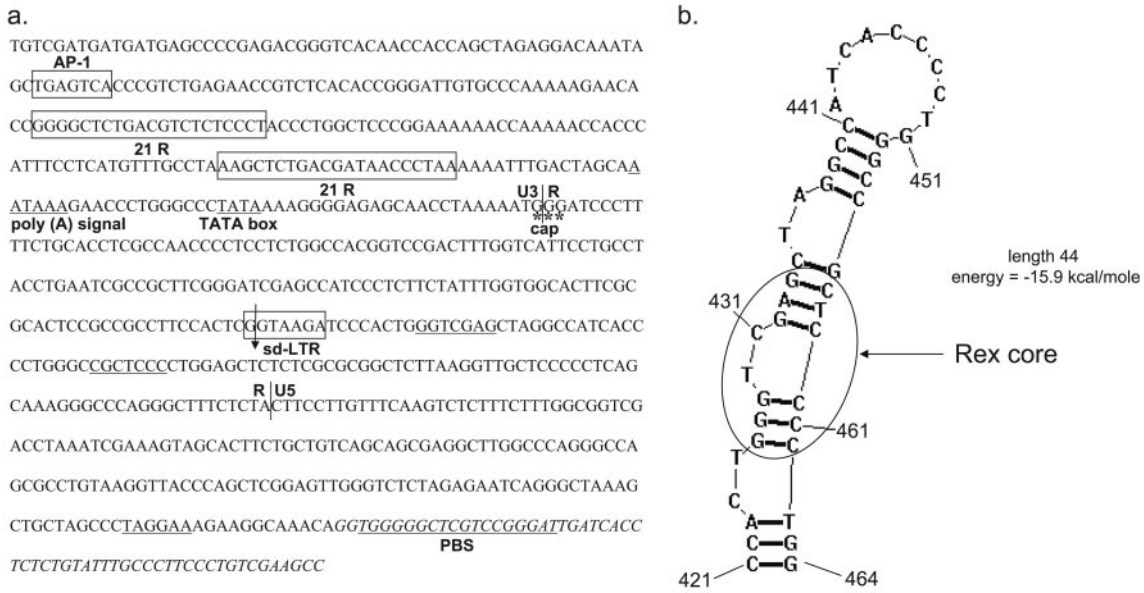


FIG. 5. (a) Nucleotide sequence of the HTLV-3(2026ND) LTR and pre-*gag* region. The U3-R-U5 locations (vertical lines), the AP-1 motif, approximate cap site (cap), polyadenylation [poly(A)] signal, TATA box, predicted splice donor site (sd-LTR), and two 21-bp repeat elements (21 R)are indicated. In the R and U5 regions, the predicted Rex core elements and nuclear riboprotein A1 binding sites are underlined. The pre-*gag* region and primer binding site (PBS) (underlined) are italicized. (b) Plot of predicted RNA stem-loop secondary structure of HTLV-3(2026ND) LTR region. The position of the Rex responsive element (RexRE) core is indicated.

also absent from prototypical HTLV-1 and PTLV-2 LTRs. Secondary-structure analysis of the LTR RNA sequence predicted a stable stem-loop structure from nucleotides 421 to 464 (Fig. 5b), similar to that shown to be essential for Rex responsiveness of viral gene expression in both HTLV-1 and HTLV-2.

Translation of predicted protein open reading frames (ORFs) across the viral genome identified all major Gag, Pro (protease), Pol, and Env proteins, as well as the regulatory proteins, Tax and Rex (Fig. 1). Translation of the overlapping *gag* and *pro* and *pro* and *pol* ORFs occurs by one or more successive −1 ribosomal frameshifts that align the different ORFs. The conserved-slippage nucleotide sequence 6(A)–8 nt–6(G)–11 nt–6(C) (where nt stands for nucleotide) is present in the Gag-Pro overlap, while a point mutation in the Pro-Pol overlap slippage sequence (GTTAAAC compared to TTTA-AAC in HTLV-1 and HTLV-2) was observed in HTLV-3(2026ND). Importantly, the asparagine codon (AAC) crucial for the slippage mechanism was unaffected.

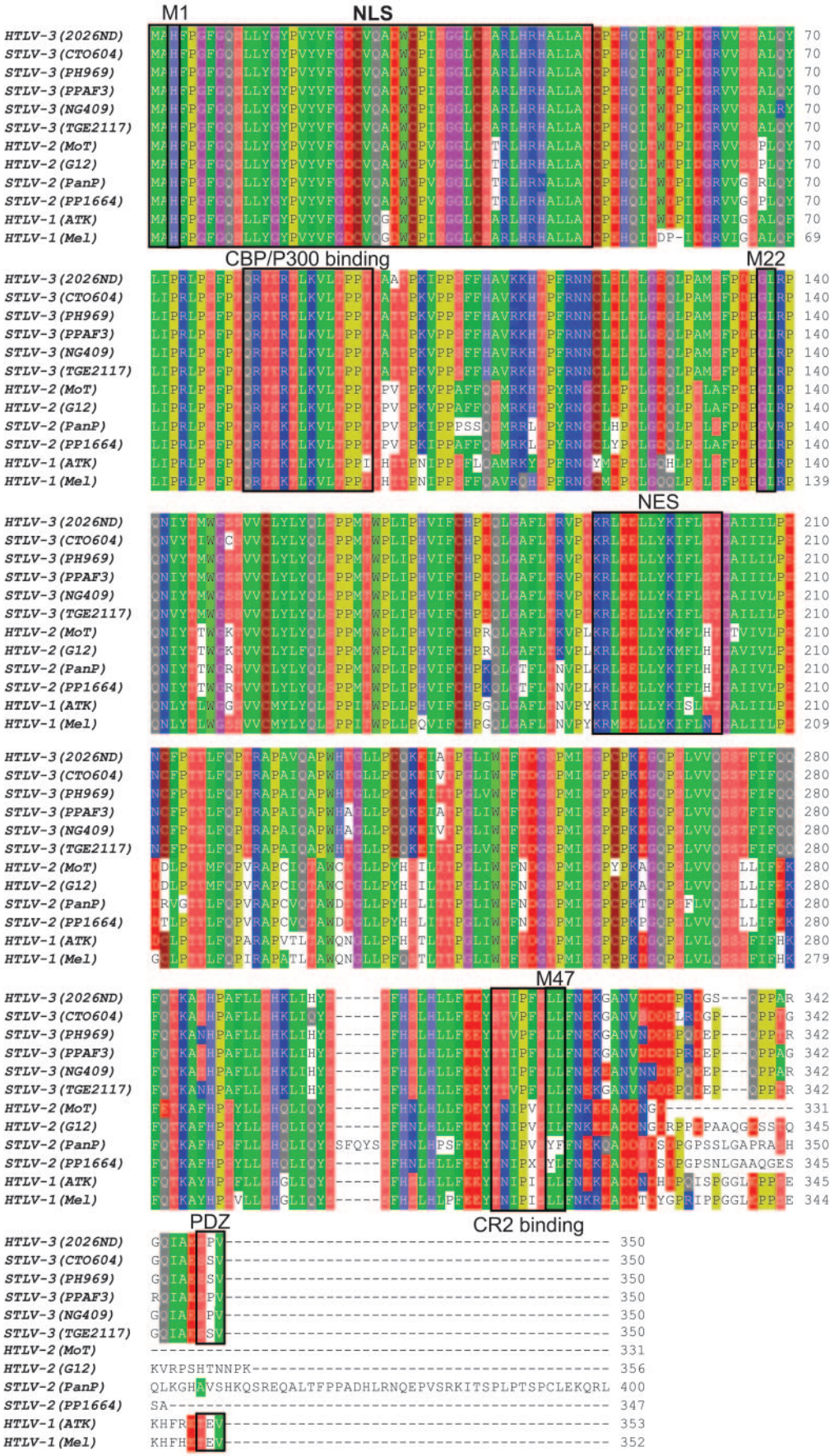
The structural and group-specific precursor Gag protein consisted of 422 amino acids (aa) and is predicted to be cleaved into the three core proteins p19 (matrix), p24 (capsid), and p15 (nucleocapsid), similar to the situation with HTLV-1, HTLV-2, and STLV-3. Across PTLV, Gag is one of the most conserved proteins, with identities ranging from 81% and 83% for HTLV-1 and PTLV-2 to 95% for STLV-3. This conservation supports the observed cross-reactivity to Gag seen with PTLV-3 antisera in Western blot (WB) assays using HTLV-1 antigens (20–22, 35, 42, 43). The Gag capsid protein showed >90% identity to HTLV-1, while the matrix and nucleocapsid proteins were more divergent, showing <78% identity to PTLV-1 and PTLV-2 (Table 1). Thus, seroreactivity to the matrix and nucleocapsid proteins may be useful to discriminate the three major PTLV groups in serologic assays.

The predicted size of the HTLV-3(2026ND) Env polyprotein is 489 aa, similar to the sizes of the Env proteins of STLV-3(PPA-F3) and STLV-3(NG409) but slightly shorter than those found in other STLV-3s due to sequence variation at the carboxyl terminus of the surface (SU) protein [314 aa versus 315 aa for STLV-3(PH969) and STLV-3(TGE2117) and 316 aa for STLV-3(CTO-604)]. In contrast, the transmembrane (TM) protein (175 aa) was highly conserved across all PTLV, supporting the use of the recombinant HTLV-1 TM protein (GD21) on WB strips to identify divergent PTLV. Despite the weak reactivity of anti-HTLV-3(2026ND) antibodies to the HTLV-1 type-specific SU peptide (MTA-1) (43) spiked onto WB strips, there was only 70.8% identity of the HTLV-3(2026ND) SU to MTA-1, which is similar to the 68.8% identity of the HTLV-2 SU to MTA-1 and which allows serologic discrimination of HTLV-2 from HTLV-1 in this region. Likewise, the HTLV-3(2026ND) and HTLV-1 SUs share only 72.1% and 67.4% identity, respectively, to the HTLV-2 type-specific SU peptide (K55). These results suggest that high antibody reactivity to either type-specific peptide, MTA-1 or K55, may not be expected in plasma and sera from HTLV-3-infected persons.

The HTLV-1 and HTLV-2 Tax proteins (Tax1 and Tax2, respectively) transactivate initiation of viral gene expression from the promoter located in the 5′ LTR and are thus essential for viral replication (11, 47). Tax1 and Tax2 have also been shown to be important for T-cell immortalization, while the HTLV-3 Tax (Tax3) has not yet been characterized (11, 47). Hence, we compared the sequence of Tax3 with those of prototypic HTLV-1, PTLV-2, and STLV-3s to determine if motifs associated with specific Tax functions were preserved between each group. Alignment of predicted Tax3 sequences shows excellent conservation of the critical functional regions, including the nuclear localization signal, cAMP response element (CREB) binding protein (CBP)/P300 binding motifs, and nuclear export signal (Fig. 6). Three sets of amino acids (M1, M22, M47) shown to be important for Tax1 transactivation and activation of the nuclear factor κβ (NF-κβ) pathway are also highly conserved in Tax3 (Fig. 6) (32). The C-terminal transcriptional activating domain (CR2), essential for CBP/p300 binding, was also conserved within TAX3, except for two mutations, N to T and I/V to F, at positions 2 and 5 of the motif, respectively (Fig. 6). However, similar mutations in the CR2 binding domain of the STLV-3 Tax have been shown recently to retain its ability to bind CBP and, to a lesser extent, p300 with no deleterious effect on transactivation of the viral promoter (8).

Although important functional motifs are highly conserved in PTLV, phenotypic differences between HTLV-1 and HTLV-2 Tax proteins have led to speculation that these differences account for the different pathologies associated with both HTLVs (11). Recently, the C terminus of Tax1, but not Tax2, has been shown to contain a conserved PDZ binding domain present in cellular proteins involved in signal transduction and induction of interleukin-2-independent growth required for T-cell transformation (24, 38). The presence of a PDZ domain in PTLV-1 and its absence in PTLV-2 suggest that this motif may contribute to the phenotypic differences between these two viral groups. The consensus PDZ domain has been defined as (S/T)XV-COOH, where X is any amino acid. Examination of the PTLV-3 Tax sequences showed that both HTLV-3 and STLV-3 have predicted PDZ domains with the consensus sequence S(P/S)V, compared to T(E/D)V in PTLV-1 (Fig. 6).

Besides Tax and Rex, two additional ORFs encoding four proteins ($p27^{I}$, $p12^{I}$, $p30^{II}$, and $p13^{II}$, where I and II indicate ORFI and ORFII, respectively) have been identified in the pX region of HTLV-1 and are important in viral infectivity and replication, T-cell activation, and cellular gene expression (5). Analysis of the pX region of HTLV-3(2026ND) revealed a total of four additional putative ORFs (named I to IV, respectively) encoding predicted proteins of 96, 122, 72, and 118 aa (Fig. 1a). Since none of the potential ORFs started with methionine start codons, we determined potential splice junctions in the HTLV-3 genome to ascertain the availability of these ORFs to predict novel proteins via complex splicing mechanisms. Prediction of splice junction positions in HTLV-3 identified only two donor sites with high confidence, at nucleotide 413 in the LTR (sd-LTR) and at nucleotide 5073 in Env (sd-Env), similar to those seen in STLV-3 (Fig. 1a) (39). Many splice acceptor sites were found throughout the genome but with low confidence (data not shown). Nonetheless, two of these splice acceptor sites are conserved in STLV-3, in the pX region at nucleotide 6835 (sa-pX2) and in Tax/Rex at nucleotide 7245 (sa-T/R) (20, 21, 39). A 45-aa protein is then predicted using the sd-Env and sa-pX2 in ORFIII (Fig. 1); this protein is 20 aa shorter than the 65-aa RORFII protein iden-


M1
NLS
CBP/P300 binding
M22
NES
M47
PDZ
CR2 binding
HTLV-3(2026ND)
STLV-3(CTO604)
STLV-3(PH969)
STLV-3(PPAF3)
STLV-3(NG409)
STLV-3(TGE2117)
HTLV-2(MoT)
HTLV-2(G12)
STLV-2(PanP)
STLV-2(PP1664)
HTLV-1(ATK)
HTLV-1(Mel)

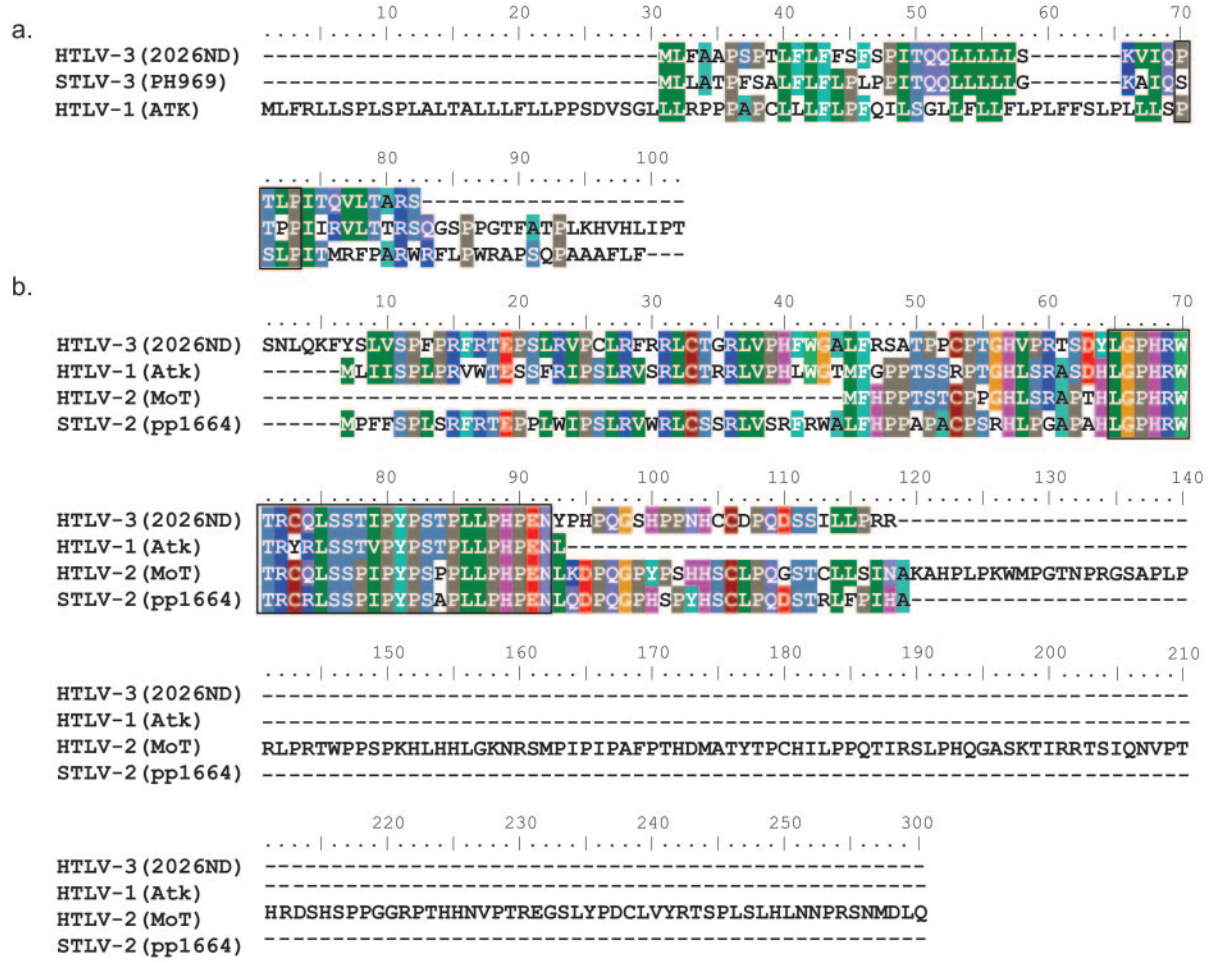


FIG. 7. Comparison of predicted accessory protein sequences of selected PTLV. (a) HTLV-3(2026ND) ORFIII compared to STLV-3(PH969) RORFII and HTLV-1 ORFI ($p12^{I}$). The location of the conserved predicted SH3-binding domain (PXXP) is boxed. (b) HTLV-3(2026ND) ORFIV aligned with HTLV-1 ORFII ($p13^{II}$), HTLV-2 ORFII ($p28^{XII}$), and STLV-2 ORFII. The highly conserved region is boxed.

tified in STLV-3(PH969) (Fig. 7a) (20, 21, 39). The HTLV-3 ORFIII protein shared 64.4% and 46.7% identity with the STLV-3 RORFII and HTLV-1 $p12^{I}$ proteins, respectively. The predicted HTLV-3 ORFIII protein was leucine rich, like that seen in the leucine zipper motifs of the HTLV-1 $p12^{I}$, but contained only one of the four SH3-binding domains (PxxP) seen in $p12^{I}$ (Fig. 7a) (5). The sa-T/R is used with the sd-Env to generate the Tax and Rex proteins and possibly the HTLV-3 ORFIV protein via complex splicing mechanisms (Fig. 1). The predicted HTLV-3 ORFIV protein shared the highest identity to the $p13^{II}$ protein of HTLV-1 (70.1%), followed by the ORFII proteins of HTLV-2 (63.5%) and STLV-2 (59.8%). Interestingly, 23 of 29 (79.3%) amino acids in the HTLV-3 ORFIV (positions 64 to 91) were identical among the ORFIIs of HTLV-1 and HTLV-2/STLV-2, suggesting a conserved functionality of this motif (Fig. 7b). In contrast to the HTLV-3 ORFII and ORFIV proteins, both the predicted HTLV-3 ORFI and ORFII proteins did not share significant sequence identity with any PTLV accessory proteins but instead shared weak sequence identity with only miscellaneous cellular proteins available in GenBank (data not shown). Analysis of alternatively spliced mRNA expression in viable cells or tissue culture, and/or in vitro characterization, will be required to investigate the expression and functionality of these putative accessory proteins.

A novel protein termed the HTLV-1 basic leucine zipper (bZIP) factor (HBZ) was recently found to be encoded on the complementary strand of the viral RNA genome between the *env* and *tax/rex* genes (12). Although HBZ was originally reported to be exclusive to PTLV-1 (12), we now show that it is conserved among PTLV (Fig. 8), including HTLV-3(2026ND), emphasizing the potential importance of this protein in viral

FIG. 6. Comparison of predicted Tax amino acid sequences of PTLV. Shown in boxes are known functional motifs: NLS, nuclear localization signal; CBP/P300, cAMP response element (CREB) binding protein; NES, nuclear export signal; CR2, C-terminal transcriptional activating domain; PDZ, PDZ binding motif; M1, M22, and M47, motifs important for Tax transactivation and NF-κβ activation (31).

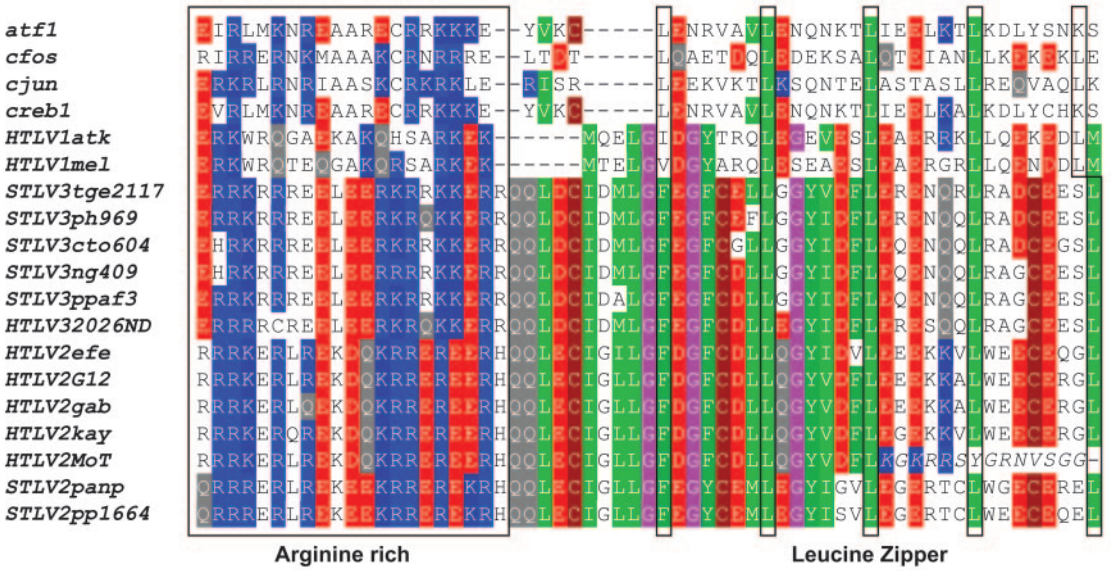


FIG. 8. Comparison of predicted amino acid sequences of PTLV and cellular bZIP transcription factors. Conserved arginine-rich and leucine zipper regions of the bZIP proteins are boxed. The frameshift mutation of the HTLV-2(MoT) leucine zipper region is italicized.

replication, persistence, and oncogenesis (2, 12). The carboxyl terminus of the HBZ ORF contains a 21-aa arginine-rich region that is relatively conserved in PTLV and known cellular bZIP transcription factors, followed by a leucine zipper region that possesses 5 or 4 conserved leucine heptads in HTLV-1 and all other PTLV, respectively (Fig. 8). HTLV-1 has 5 leucine heptads similar to that found in mammalian bZIP proteins, while PTLV-2 and PTLV-3 have 4 leucine heptads followed by a leucine octet (Fig. 8). Of all PTLV with full-length genomes available in GenBank, only HTLV-2(MoT) did not have the full complement of leucine heptads. HTLV-2(MoT) was limited to the initial 3 leucine motifs due to a single-nucleotide deletion at position 6823 that caused a frameshift in the predicted HBZ sequence. Resequencing of the HTLV-2(MoT) provirus genome directly from peripheral blood lymphocyte DNA, and not a tissue culture isolate, is necessary to verify the HBZ ORF mutation in this virus.

DISCUSSION

Here we report the first complete nucleotide sequence and genomic characterization of HTLV-3 and show that the genome of this novel virus is genetically equidistant from HTLV-1 and HTLV-2 but clearly falls within the diversity of STLV-3. These results support further the HTLV-3 nomenclature proposed for this virus and demonstrate that HTLV-3 originated from STLV-3 (7, 43). Nonetheless, our finding that HTLV-3(2026ND) is distinct from all STLV-3s, sharing only 86 to 92% identity across its entire genome, suggests that the specific primate counterpart of HTLV-3(2026ND) has not yet been identified. This contrasts with the genetic relatedness of the second HTLV-3(Pyl43), reported recently in a Bakola pygmy from Cameroon, which is nearly identical to STLV-3 found in a red-capped mangabey from Cameroon and thus probably represents a recent zoonotic infection (7). The similarity seen between these two sequences is similar to that observed in linked STLV-3 transmission pairs (42). HTLV-3(2026ND) is also distinct from HTLV-3(Pyl43), demonstrating multiple introductions of different STLV-3-like viruses in persons in Cameroon following exposure to diverse STLV-3s prevalent in NHPs in this region (9, 20–22). Similar repeated and historical cross-species infections of humans with various STLV-1 strains led to the emergence and dissemination of several HTLV-1 subtypes in West-Central Africa (13, 26, 31, 43). Thus, it is possible that HTLV-3(2026ND) may similarly represent a strain circulating within humans living in this geographic region. This hypothesis is consistent with our finding that the HTLV-3(2026ND) predecessor originated over 30 millennia ago, at a time when we and others estimate that the ancestors of both HTLV-1 and HTLV-2 appeared. The inference of an ancient HTLV-3(2026ND) ancestor, combined with the wide geographic distribution of STLVs and a history of STLVs crossing into humans, and the recent finding of another HTLV-3 in an African pygmy (7, 13, 20–22, 35, 42) all imply that HTLV-3 infection is likely to be more prevalent than previously appreciated. Additional information on the diversity of HTLV-3 and STLV-3 will be needed to determine whether HTLV-3(2026ND) is truly an ancient infection that has become endemic in humans or represents a more recent primary zoonotic transmission from a primate infected with a very old, divergent STLV-3.

The inferred ancient history of the ancestral HTLVs and the recent finding of STLV-like infections in African hunters collectively indicate that cross-species transmission of STLVs to humans is both an ancient and a contemporary phenomenon dependent on behavior that exposes humans to NHPs (7, 43). The ancient origin of HTLV contrasts with that reported for HIV, which is believed to have entered into humans from simian immunodeficiency virus-infected NHPs within the last century (15, 29). Thus, HTLVs appear to have had a long period of viral evolution and adaptation in humans, possibly

resulting in the observed lower pathogenicity of HTLV compared to HIV.

A more precise determination of the origin and distribution of HTLV-3 infection will require further studies, such as expanded surveillance of both humans and NHPs. However, serosurveys for HTLV-3 and STLV-3 may be complicated by several factors, including different WB profiles for each HTLV-3 strain, one HTLV-1-like and the other showing indeterminate seroreactivity; variable WB profiles observed for STLV-3-infected NHPs; and high rates of HTLV indeterminate seroreactivity in some regions (7, 9, 14, 18, 20–22, 35, 42, 43). Thus, additional diagnostic tools are required to determine the level of penetration of HTLV-3 into the general population and to search for the potential primate origin of HTLV-3(2026ND). Screening for HTLV-3 will be facilitated by the development and application of diagnostic serologic and molecular assays based on the sequences reported here. For example, since the Gag matrix and nucleocapsid regions and the envelope surface protein are relatively conserved within PTLV-3 and are divergent from PTLV-1 and PTLV-2, it may be possible to use them in serologic assays to differentiate the three PTLV groups.

Changes in the molecular structure and sequences of viruses have been proposed to play a role in the increased transmissibility and pathogenesis of viruses following cross-species transmission and adaptation to a new host. Thus, we examined in detail the genetic structure and sequence of HTLV-3 to determine if important functional motifs involved in viral expression and HTLV-induced leukemogenesis are conserved (5, 11, 12, 24, 38, 47). All enzymatic, regulatory, and structural proteins are well conserved in HTLV-3(2026ND), including conserved functional motifs in Tax that are important for viral gene expression and T-cell proliferation. These results, combined with the observed genetic stability of HTLV-3, suggest the absence of dramatic adaptive changes following cross-species transmission of PTLVs. Nonetheless, we did observe several important molecular features of the HTLV-3 genome that are either similar to or distinct from those of other HTLVs. For example, identification of a PDZ domain, known to be important in cellular signal transduction and T-cell transformation (24, 38, 45), in the Tax protein of HTLV-3(2026ND), similar to that seen in HTLV-1 but not HTLV-2 (11), suggests that the HTLV-3 Tax may be more phenotypically similar to the HTLV-1 than to the HTLV-2 Tax. The high amino acid identity of the PTLV-3 Tax proteins, combined with the ability of STLV-3 to transform human cells in vitro, also suggests that the HTLV-3 Tax may function similarly to the HTLV-1 Tax (14). Indeed, recent studies demonstrated that the Tax proteins of HTLV-3(Pyl43) and STLV-3(CTO602) are expressed in vivo and that the STLV-3 Tax is a transactivator in vitro, suggesting that lymphoproliferative disorders may occur in STLV-3 and HTLV-3 infections (8). However, whether the presence of a PDZ domain in HTLV-3 is associated with specific cellular and/or clinical outcomes, as it is for HTLV-1, will require further investigation.

In contrast to the similarity observed between the HTLV-1 and HTLV-3 *tax* genes, the HTLV-3(2026ND) LTR is less conserved, having only two of the typical three 21-bp Tax-responsive elements identified in HTLV-1 and HTLV-2 that are responsible for basal viral transcription levels. Like that of STLV-3, the HTLV-3(2026ND) LTR is missing the TATA-distal 21-bp repeat element (20–22, 39, 42). Although others have shown that deletion of the middle, rather than the distal, 21-bp element is more critical for the loss of basal HTLV-1 transcription levels (3), additional studies are needed to determine what effect the absence of a 21-bp element will have on HTLV-3(2026ND) gene expression and replication. Aside from the identification of a unique AP-1 site, all of the remaining functional elements in the LTR were more conserved, including the stem-loop structure necessary for Rex-responsive control of viral expression in HTLV-1 and -2.

Recently, a novel HBZ protein was found to be encoded on the complementary strand of the viral RNA genome between the *env* and *tax/rex* genes, and HBZ was shown to negatively regulate viral replication and to enhance viral infectivity and persistence (2, 12). Protein translation on the minus-strand RNA is a unique feature of HTLV-1 not previously seen in retroviruses (12). The recent finding of HBZ mRNA expression in ATL patients suggests a role of HBZ mRNA in the survival of leukemic cells in vivo and in HTLV-1-associated oncogenesis (28). Although HBZ was originally reported to be exclusive to PTLV-1 (12), we now provide evidence for a putative HBZ region among all PTLV, including HTLV-3(2026ND), further demonstrating the potential importance of the HBZ protein and mRNA in viral replication and oncogenesis. For example, the HBZ protein has also been reported to bind to AP-1 regulatory elements, including the transcription factors JunB and c-Jun, to modulate their transcriptional activity (4), which may then disturb AP-1 regulation of many cellular processes, including cell proliferation, transformation, and death (30). The discovery of an AP-1 site in the LTR of HTLV-3(2026ND), combined with the ability of Tax1 and HBZ to activate cellular transcription through AP-1 sites in vitro (4, 47), is important and suggests an alternative mechanism for regulation of viral expression and replication not previously known for HTLV. Thus, additional studies are required to confirm the potential effect of the predicted PTLV HBZ proteins, in conjunction with the AP-1 site in the LTR of HTLV-3(2026ND), on viral expression and leukemogenesis.

In summary, we have shown that the novel HTLV-3 genome is genetically stable and has an ancient origin. We have also demonstrated that the HTLV-3 genome contains many of the functional motifs important for viral expression and pathology attributed to HTLV-1. Additional studies are needed to further characterize the unique molecular features of HTLV-3 identified here, to determine whether HTLV-3 has become endemic in humans, and to better understand the public health importance of this novel human virus.

ACKNOWLEDGMENTS

We thank Debbie Kuehl for assistance with data input and Susan Marriott for critical review of the manuscript.

N.D.W. is supported by a National Institutes of Health (NIH) Director's Pioneer Award Program (grant number DP1-OD000370) and an International Research Scientist Development Award from the NIH Fogarty International Center (K01 TW00003-1).

Use of trade names is for identification only and does not imply endorsement by the U.S. Department of Health and Human Services, the Public Health Service, or the Centers for Disease Control and Prevention. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

REFERENCES

1. **Araujo, A., and W. W. Hall.** 2004. Human T-lymphotropic virus type II and neurological disease. Ann. Neurol. **56:**10–19.
2. **Arnold, J., B. Yamamoto, M. Li, A. J. Phipps, I. Younis, M. D. Lairmore, and P. L. Green.** 2006. Enhancement of infectivity and persistence in vivo by HBZ, a natural antisense coded protein of HTLV-1. Blood **107:**3976–3982.
3. **Barnhart, M. K., L. M. Connor, and S. J. Marriott.** 1997. Function of the human T-cell leukemia virus type 1 21-base-pair repeats in basal transcription. J. Virol. **71:**337–344.
4. **Basbous, J., C. Arpin, G. Gaudray, M. Piechaczyk, C. Devaux, and J.-M. Mesnard.** 2003. The HBZ factor of human T-cell leukemia virus type 1 dimerizes with transcriptional factors JunB and c-Jun and modulates their transcriptional activity. J. Biol. Chem. **278:**43620–43627.
5. **Bindhu, M., A. Nair, and M. D. Lairmore.** 2004. Role of accessory proteins of HTLV-1 in viral replication, T cell activation, and cellular gene expression. Front. Biosci. **9:**2556–2576.
6. **Brites, C., W. Harrington, Jr., C. Pedroso, E. Martins Netto, and R. Badaro.** 1997. Epidemiological characteristics of HTLV-I and II co-infection in Brazilian subjects infected by HIV-1. Braz. J. Infect. Dis. **1:**42–47.
7. **Callatini, S., S. A. Chevalier, R. Duprez, S. Bassot, A. Froment, R. Mahieux, and A. Gessain.** 2005. Discovery of a new human T-cell lymphotropic virus (HTLV-3) in Central Africa. Retrovirology **2:**30.
8. **Chevalier, S., L. Meertens, C. Pise-Masison, S. Calattini, H. Park, A. A. Alhaj, M. Zhou, A. Gessain, F. Kashanchi, J. Brady, and R. Mahieux.** 13 Mar. 2006, posting date. The tax protein from the primate T-cell lymphotropic virus type 3 is expressed in vivo and is functionally related to HTLV-1 Tax rather than HTLV-2 Tax. Oncogene [Online.] doi:10.1038/sj.onc.1209472.
9. **Courgnaud, V., S. Van Dooren, F. Liegeois, X. Pourrut, B. Abela, S. Loul, E. Mpoudi-Ngole, A. Vandamme, E. Delaporte, and M. Peeters.** 2004. Simian T-cell leukemia virus (STLV) infection in wild primate populations in Cameroon: evidence for dual STLV type 1 and type 3 infection in agile mangabeys (*Cercocebus agilis)*. J. Virol. **78:**4700–4709.
10. **Digilio, L., A. Giri, N. Cho, J. Slattery, P. Markham, and G. Franchini.** 1997. The simian T-lymphotropic/leukemia virus from *Pan paniscus* belongs to the type 2 family and infects Asian macaques. J. Virol. **71:**3684–3692.
11. **Feuer, G., and P. L. Green.** 2005. Comparative biology of human T-cell lymphotropic virus type 1 (HTLV-1) and HTLV-2. Oncogene **24:**5996–6004.
12. **Gaudray, G., F. Gachon, J. Basbous, M. Biard-Piechaczyk, C. Devaux, and J. M. Mesnard.** 2002. The complementary strand of the human T-cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription. J. Virol. **76:**12813–12822.
13. **Gessain, A., and R. Mahieux.** 2000. Epidemiology, origin and genetic diversity of HTLV-1 retrovirus and STLV-1 simian affiliated retrovirus. Bull Soc. Pathol. Exot. **93:**163–171.
14. **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.
15. **Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp.** 2000. AIDS as a zoonosis: scientific and public health implications. Science **287:**607–614.
16. **Lemey, P., O. G. Pybus, S. Van Dooren, and A. M. Vandamme.** 2005. A Bayesian statistical analysis of human T-cell lymphotropic virus evolutionary rates. Infect. Gen. Evol. **5:**291–298.
17. **Lole, K. S., R. C. Bollinger, R. S. Paranjape, D. Gadkari, S. S. Kulkarni, N. G. Novak, R. Ingersoll, H. W. Sheppard, and S. C. Ray.** 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J. Virol. **73:**152–160.
18. **Mahieux, R., P. Horal, P. Mauclere, O. Mercereau-Puijalon, M. Guillotte, L. Meertens, E. Murphy, and A. Gessain.** 2000. Human T-cell lymphotropic virus type 1 gag indeterminate Western blot patterns in Central Africa: relationship to *Plasmodium falciparum* infection. J. Clin. Microbiol. **38:**4049–4057.
19. **Mathews, D. H., J. Sabina, M. Zuker, and D. H. Turner.** 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. **288:**911–940.
20. **Meertens, L., and A. Gessain.** 2003. Divergent simian T-cell lymphotropic virus type 3 (STLV-3) in wild-caught *Papio hamadryas papio* from Senegal: widespread distribution of STLV-3 in Africa. J. Virol. **77:**782–789.
21. **Meertens, L., R. Mahieux, P. Mauclere, J. Lewis, and A. Gessain.** 2002. Complete sequence of a novel highly divergent simian T-cell lymphotropic virus from wild-caught red-capped mangabeys (*Cercocebus torquatus*) from Cameroon: a new primate T-lymphotropic virus type 3 subtype. J. Virol. **76:**259–268.
22. **Meertens, L., V. Shanmugam, A. Gessain, B. E. Beer, Z. Tooze, W. Heneine, and W. M. Switzer.** 2003. A novel, divergent simian T-cell lymphotropic virus type 3 in a wild-caught red-capped mangabey (*Cercocebus torquatus torquatus*) from Nigeria. J. Gen. Virol. **84:**2723–2727.
23. **Posada, D., and K. A. Crandall.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14:**817–818.
24. **Rousset, R., S. Fabre, C. Desbios, F. Bantignies, and P. Jalinot.** 1998. The C-terminus of the HTLV-1 Tax oncoprotein mediates interaction with the PDZ domain of cellular proteins. Oncogene **6:**643–654.
25. **Salemi, M., J. Desmyter, and A. M. Vandamme.** 2000. Tempo and mode of human and simian T-lymphotropic virus (HTLV/STLV) evolution revealed by analyses of full-genome sequences. Mol. Biol. Evol. **17:**374–386.
26. **Salemi, M., S. Van Dooren, and A. M. Vandamme.** 1999. Origin and evolution of human and simian T-cell lymphotropic viruses. AIDS Rev. **1:**131–139.
27. **Sanderson, M. J.** 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics **19:**301–302.
28. **Satou, Y., J.-I. Yasunaga, M. Yoshida, and M. Matsuoka.** 2006. HTLV-1 basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells. Proc. Natl. Acad. Sci. USA **103:**720–725.
29. **Sharp, P. M., E. Bailes, F. Gao, B. E. Beer, V. M. Hirsch, and B. H. Hahn.** 2000. Origins and evolution of AIDS viruses: estimating the time-scale. Biochem. Soc. Trans. **28:**275–282.
30. **Shaulian, E., and M. Karin.** 2002. AP-1 as a regulator of cell life and death. Nat. Cell Biol. **4:**E131–E136.
31. **Slattery, J. P., G. Franchini, and A. Gessain.** 1999. Genomic evolution, patterns of global dissemination, and interspecies transmission of human and simian T-cell leukemia/lymphotropic viruses. Genome Res. **9:**525–540.
32. **Smith, M. R., and W. C. Greene.** 1990. Identification of HTLV-I tax transactivator mutants exhibiting novel transcriptional phenotypes. Genes Dev. **4:**1875–1885.
33. **Switzer, W. M., D. Pieniazek, P. Swanson, H. H. Samdal, V. Soriano, R. F. Khabbaz, J. E. Kaplan, R. B. Lal, and W. Heneine.** 1995. Phylogenetic relationship and geographic distribution of multiple human T-cell lymphotropic virus type II subtypes. J. Virol. **69:**621–632.
34. **Switzer, W. M., M. Salemi, V. Shanmugam, F. Gao, M. E. Cong, C. Kuiken, V. Bhullar, B. E. Beer, D. Vallet, A. Gautier-Hion, Z. Tooze, F. Villinger, E. C. Holmes, and W. Heneine.** 2005. Ancient co-speciation of simian foamy viruses and primates. Nature **434:**376–380.
35. **Takemura, T., M. Yamashita, M. K. Shimada, S. Ohkura, T. Shotake, M. Ikeda, T. Miura, and M. Hayami.** 2002. High prevalence of simian T-lymphotropic virus type L in wild Ethiopian baboons. J. Virol. **76:**1642–1648.
36. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
37. **Thomson, M. M., and R. Najera.** 2005. Molecular epidemiology of HIV-1 variants in the global AIDS pandemic: an update. AIDS Rev. **7:**210–224.
38. **Tsubata, C., M. Higuchi, M. Takahashi, M. Oie, Y. Tanaka, F. Geyjo, and M. Fujii.** 2005. PDZ domain-binding motif of human T-cell leukemia virus type 1 Tax oncoprotein is essential for the interleukin 2 independent growth induction of a T-cell line. Retrovirology **2:**46.
39. **Van Brussel, M., P. Goubau, R. Rousseau, J. Desmyter, and A. M. Vandamme.** 1997. Complete nucleotide sequence of the new simian T-lymphotropic virus, STLV-PH969 from a Hamadryas baboon, and unusual features of its long terminal repeat. J. Virol. **71:**5464–5472.
40. **Van Brussel, M., M. Salemi, H. F. Liu, J. Gabriels, P. Goubau, J. Desmyter, and A. M. Vandamme.** 1998. The simian T-lymphotropic virus STLV-PP1664 from *Pan paniscus* is distinctly related to HTLV-2 but differs in genomic organization. Virology **243:**366–379.
41. **Van Dooren, S., M. Salemi, X. Pourrut, M. Peeters, E. Delaporte, M. Van Ranst, and A. M. Vandamme.** 2001. Evidence for a second simian T-cell lymphotropic virus type 3 in *Cercopithecus nictitans* from Cameroon. J. Virol. **75:**11939–11941.
42. **Van Dooren, S., V. Shanmugam, V. Bhullar, B. Parekh, A. M. Vandamme, W. Heneine, and W. M. Switzer.** 2004. Identification in gelada baboons (*Theropithecus gelada*) of a distinct simian T-cell lymphotropic virus type 3 with a broad range of Western blot reactivity. J. Gen. Virol. **85:**507–519.
43. **Wolfe, N. D., W. Heneine, J. K. Carr, A. D. Garcia, V. Shanmugam, U. Tamoufe, J. N. Torimiro, A. T. Prosser, M. Lebreton, E. Mpoudi-Ngole, F. E. McCutchan, D. L. Birx, T. M. Folks, D. S. Burke, and W. M. Switzer.** 2005. Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. Proc. Natl. Acad. Sci. USA **102:**7994–7999.
44. **Womble, D. D.** 2000. GCG: the Wisconsin Package of sequence analysis programs. Methods Mol. Biol. **132:**3–22.
45. **Xie, L., B. Yamamoto, A. Haoudi, O. J. Semmes, and P. L. Green.** 2006. PDZ binding motif of HTLV-1 Tax promotes virus-mediated T-cell proliferation in vitro and persistence in vivo. Blood **107:**1980–1988.
46. **Yamashita, M., E. Ido, T. Miura, and M. Hayami.** 1996. Molecular epidemiology of HTLV-1. J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. **13**(Suppl. 1)**:**S124–S131.
47. **Yoshida, M.** 2001. Multiple viral strategies of HTLV-1 for dysregulation of cell growth control. Annu. Rev. Immunol. **19:**475–496.