Identification in gelada baboons (*Theropithecus gelada*) of a distinct simian T-cell lymphotropic virus type 3 with a broad range of Western blot reactivity

Sonia Van Dooren,[1] Vedapuri Shanmugam,[2] Vinod Bhullar,[2] Bharat Parekh,[3] Anne-Mieke Vandamme,[1] Walid Heneine[2] and William M. Switzer[2]

Correspondence
William M. Switzer
bis3@cdc.gov

[1]Rega Institute for Medical Research, Katholieke Universiteit Leuven, Leuven, Belgium

[2,3]HIV and Retrovirology Branch[2] and HIV Immunology and Diagnostics Branch[3], Division of AIDS, STD and TB Laboratory Research, National Center for HIV, STD and TB Prevention, Centers for Disease Control and Prevention, 1600 Clifton Rd, MS G-19, Atlanta, GA 30333, USA

Antibodies to simian T-cell lymphotropic virus (STLV) were found in serum or plasma from 12 of 23 (52·2 %) gelada baboons (*Theropithecus gelada*) captive in US zoos. A variety of Western blot (WB) profiles was seen in the 12 seroreactive samples, including human T-cell lymphotropic virus (HTLV)-1-like ($n=5$, 41·7 %), HTLV-2-like ($n=1$, 8·3 %), HTLV-untypable ($n=4$, 33·3 %) and indeterminate ($n=2$, 16·6 %) profiles. Phylogenetic analysis of *tax* or *env* sequences that had been PCR amplified from peripheral blood lymphocyte DNA available from nine seropositive geladas showed that four were infected with identical STLV-1s; these sequences clustered with STLV-1 from Celebes macaques and probably represent recent cross-species infections. The *tax* sequences from the five remaining geladas were also identical and clustered with STLV-3. Analysis of the complete STLV-3 genome (8917 bp) from one gelada, TGE-2117, revealed that it is unique, sharing only 62 % similarity with HTLV-1/ATK and HTLV-2/Mo. STLV-3/TGE-2117 was closest genetically to STLV-3 from an Eritrean baboon (STLV-3/PH969, 95·6 %) but more distant from STLV-3s from red-capped mangabeys from Cameroon and Nigeria (STLV-3/CTO-604, 87·7 %, and STLV-3/CTO-NG409, 87·2 %, respectively) and Senegalese baboons (STLV-3/PPA-F3, 88·4 %). The genetic relatedness of STLV-3/TGE-2117 to STLV-3 was confirmed by phylogenetic analysis of a concatenated *gag-pol-env-tax* sequence (6795 bp). An ancient origin of 73 628–109 809 years ago for STLV-3 was estimated by molecular clock analysis of third-codon positions of *gag-pol-env-tax* sequences. LTR sequences from five STLV-3-positive geladas were >99 % identical and clustered with that from a *Papio anubis* × *P. hamadryas* hybrid Ethiopian baboon, suggesting a common source of STLV-3 in these sympatric animals. LTR sequences obtained 20 years apart from a mother–infant pair were identical, providing evidence of both mother-to-offspring transmission and a high genetic stability of STLV-3. Since STLV-3-infected primates show a range of HTLV-like WB profiles and have an ancient origin, further studies using STLV-3-specific testing are required to determine whether STLV-3 infects humans, especially in regions of Africa where STLV-3 is endemic.

Received 4 September 2003
Accepted 16 October 2003

INTRODUCTION

Simian T-cell lymphotropic viruses (STLVs) are a diverse group of deltaretroviruses that naturally infect many species of Old World primates (Gessain *et al.*, 2002; Slattery *et al.*, 1999). STLVs are related to human T-cell lymphotropic virus types 1 and 2 (HTLV-1 and -2) and collectively HTLVs and STLVs are commonly referred to as primate T-cell lymphotropic viruses (PTLVs) (Slattery *et al.*, 1999). While STLV-1 has been found in over 20 species of African and Asian primates, STLV-2 has been detected in only two captive troops of bonobos (*Pan paniscus*; strains STLV-2/PP1664

The complete genome of STLV-3/TGE-2117 has been assigned the GenBank accession number AY217650. The GenBank accession numbers for the STLV-3 *tax* and LTR and STLV-1 *tax* and *env* sequences are AY241678–AY241693.

and STLV-2/panp) (Giri *et al.*, 1994; Goubau *et al.*, 1994; Liu *et al.*, 1994; Meertens *et al.*, 2001; Meertens & Gessain, 2003; Miyoshi *et al.*, 1982; Nerrienet *et al.*, 2001; Van Brussel *et al.*, 1997, 1999; Vandamme *et al.*, 1996). STLV-3 was first identified in 1994 in only a single captive Eritrean sacred baboon (*Papio hamadryas*; strain STLV-3/PH969) (Goubau *et al.*, 1994). More recently, STLV-3 has been found in wild-caught monkeys originating from different African regions including several Ethiopian sacred (*P. hamadryas*; strain PHA-7550) and hybrid baboons (*P. hamadryas* × *P. anubis*; strains HYB-2210 and -2220, respectively) (Takemura *et al.*, 2002), Cameroonian and Nigerian red-capped mangabeys (*Cercocebus torquatus torquatus*; strains CTO-604 and CTO-602 and strain CTO-NG409, respectively) (Meertens *et al.*, 2002, 2003), Cameroonian spot-nosed guenons (*Cercopithecus nictitans*; strains CNI-217 and CNI-227) (Van Dooren *et al.*, 2001) and Senegalese olive baboons (*P. papio*; strains PPA-F3, -F4, -F5 and -F9) (Meertens & Gessain, 2003).

Despite these recent reports, the virological and immunological characteristics of STLV-3 infection of non-human primates (NHPs) remain poorly understood. STLV-3 has the highest amino acid sequence identity to the HTLV-1 and HTLV-2 p24 Gag (85 %) and envelope transmembrane proteins (72–88 %, respectively) (Meertens *et al.*, 2002; Van Brussel *et al.*, 1997). Thus, sera from STLV-3-infected primates show cross-reactivity to the p24 Gag and transmembrane (GD21) proteins in Western blot (WB) assays containing HTLV-1 antigens and conserved HTLV-1 and HTLV-2 recombinant GD21 proteins (Goubau *et al.*, 1994; Meertens *et al.*, 2001, 2002; Meertens & Gessain 2003; Takemura *et al.*, 2002; Van Dooren *et al.*, 2001). Sera from STLV-3-infected animals may also react in WB assays with the HTLV-2-type specific recombinant envelope (Env) protein, K55; thus, STLV-3 seroreactivity has been referred to as HTLV-2-like, though phylogenetic analysis clearly demonstrates the uniqueness of STLV-3 (Goubau *et al.*, 1994; Meertens *et al.*, 2001, 2002; Meertens & Gessain 2003; Takemura *et al.*, 2002; Van Dooren *et al.*, 2001). Therefore, the use of HTLV-1-based serological assays cannot reliably distinguish between STLV-3 and STLV-1/2 infections.

The recognition from recent data that STLV-3 is not a rare infection also highlights the importance of defining its prevalence among different primate species, understanding its genetic diversity and evolution and determining whether humans exposed to NHPs are infected with STLV-3-like variants. However, little is known about the epidemiology and genetic diversity of STLV-3. Full-length STLV-3 genomes have been obtained only for strains PH969, CTO-604, CTO-NG409 and PPA-F3. For all other STLV-3 strains, only short sequences in the pX and LTR regions are available (Takemura *et al.*, 2002; Van Dooren *et al.*, 2001). Unlike STLV-1 and STLV-2, which are believed to be the simian equivalents of HTLV-1 and HTLV-2, it is not known whether STLV-3-like variants naturally infect humans. Limited testing of people with indeterminate HTLV serological results has found no evidence of STLV-3-like infections (Busch *et al.*, 2000; Vandamme *et al.*, 1997). However, phylogenetic studies of STLV-1 and HTLV-1 suggesting the occurrence of multiple interspecies transmissions of STLV (Gessain *et al.*, 2002; Mahieux *et al.*, 1998; Meertens *et al.*, 2001; Slattery *et al.*, 1999), combined with the ability of STLV-3 to grow in human cells *in vitro* (Goubau *et al.*, 1994), both suggest that humans may be at risk for cross-species infections with STLV-3. Zoonotic transmission of simian retroviruses to humans is not uncommon, since studies of workers occupationally exposed to NHPs have documented infection with a number of simian retroviruses, including simian immunodeficiency virus, simian foamy virus and simian type D retrovirus (Heneine *et al.*, 1998; Khabbaz *et al.*, 1994; Lerche *et al.*, 2001). Information on STLV-3 infection among captive primates may also help to assess the types of simian retroviruses that people may be occupationally exposed to. Thus, a better understanding of the genetic diversity of STLV-3 and its distribution among NHPs will help define the risks of transmitting these viruses to humans.

In this study we performed a serosurvey for STLV in gelada baboons (*Theropithecus gelada*) living in US zoos. We present here the serological and molecular characterization of the STLV found in these primates and describe the complete proviral sequence of a unique STLV-3 genome identified in one gelada baboon. In addition, we describe novel WB profiles in STLV-3-infected baboons and provide information regarding the genetic stability and transmission of STLV-3 among geladas housed at the same zoo.

METHODS

Specimens. Archived serum or fresh plasma samples from 23 captive gelada baboons (*T. gelada*) were available from five US zoos and all specimens were stored at −20 °C until testing. Fresh EDTA- or sodium citrate-treated whole-blood specimens were also obtained on an opportunistic basis for initial or follow-up testing in accordance with the animal care and use committees at each institution. Peripheral blood lymphocytes (PBLs) were obtained from 11 geladas by Ficoll-Hypaque centrifugation and DNA lysates were prepared as described previously (Switzer *et al.*, 1995). DNA was also extracted in a separate laboratory from frozen heparinized blood samples collected in 1981 from seven other animals (TGE-BK1–TGE-BK7) for PCR testing as described below. Demographic data for the gelada baboons were provided by each zoo.

Serological assays. Initial screening for STLV antibodies in serum and plasma samples was performed by using the Vironostika HTLV-1/2 microelisa system (Organon-Teknika) following the manufacturer's instructions. Seroreactive samples were then tested in a WB test (HTLV Blot 2.4, Genelabs Diagnostics) containing disrupted HTLV-1 virions, a gp21 recombinant protein (GD21) common to both HTLV-1 and HTLV-2 and two HTLV-type-specific recombinant Env proteins, MTA-1 and K55, which allow serological differentiation of HTLV-1 and HTLV-2, respectively.

PCR analysis. Generic nested PTLV *tax* and STLV-1 *env* or STLV-3-specific LTR PCR was performed on DNA lysates prepared from uncultured PBLs. The PTLV *tax* and STLV-3-specific LTR nested PCR was also performed on DNA extracted from frozen blood samples

Table 1. Sequences of primers used for amplifying partial *tax*, *env* and LTR regions of primate T-cell lymphotropic viruses

Name	Primer sequence (5′→3′)	Location*	Expected PCR product size (bp)	Annealing temp. (°C)	No. of cycles
PH1F	TTGTCATCAGCCCACTTCCCAGG	*tax*, 7243–7262, outer			
PH2R	AAGGAGGGGAGTCGAGGGATAAGG	*tax*, 7478–7455, outer	236	50	40
PH2F	CCCAGGTTTCGGGCAAAGCCTTCT	*tax*, 7257–7280, inner			
PH2R†	AAGGAGGGGAGTCGAGGGATAAGG	*tax*, 7478–7455, inner	222	50	40
PLTRF1	CCCCAAGACGAACCACAACCACCAACT	LTR, 17–43, outer			
PLTRR1	CAACTGTTTGCTTTCTTCCCTAGGGCT	LTR, 699–673, outer	683	50	35
PLTRF2	TCATCCGTCTGAGAGCCGTCTCGC	LTR, 62–86, inner	602	50	35
PLTRR2	AGCCCAGACTCTTTAGGACCCAACTCC	LTR, 663–637, inner			
P21EF1	GGACCCACTGCTTTGACCCCC	*env*, 6019–6039, outer			
P21ER1	ATTGCGTGCTTGGTTTACAGG	*env*, 6682–6662, outer	664	45	35
P21EF2	CATAACTCCCTCATCCTGCCC	*env*, 6069–6089, inner			
P21ER2	AGTAATGGGGGTATCTGACGC	*env*, 6639–6619, inner	571	45	35

*Positions of the *tax* and LTR primers are given according to the STLV-3 (strain PH969) genome; the *env* primer positions are given according to HTLV-1 (strain ATK).

†Primer PH2R was used with PH2F in a semi-nested PCR.

from seven other geladas. The primer sequences, expected product sizes and amplification conditions are given in Table 1. The generic PTLV *tax* primers used in the primary PCR amplification have been shown previously to amplify HTLV-1, HTLV-2 and divergent PTLVs, including STLV-2/Panp and STLV-3/PHA-PH969 (Busch *et al.*, 2000). In the current study, we also performed a semi-nested *tax* PCR using the primers PH2F and PH2R (Table 1) to increase the amount of specific amplified product necessary for direct sequence analysis. The STLV-1 *env* and STLV-3 LTR primers were based on the HTLV-1/ATK and STLV-3/PHA-PH969 sequences with GenBank accession numbers J02029 and Y07616, respectively. To obtain the complete sequence of the unique gelada STLV-3, we PCR amplified sequential overlapping regions of the genome as previously described (Meertens *et al.*, 2003). Nested PCR products were electrophoresed in 1·8 % agarose gels and visualized by ethidium bromide staining.

Sequence and phylogenetic analysis. PCR products were purified using the Qiaquick PCR purification kit (Qiagen) and sequenced in both directions using a BigDye terminator cycle kit (PE Biosystems) and an ABI 373 automated sequencer. BLAST searches at the National Center for Biotechnology Information website (http://www.ncbi.nlm.nih.gov/BLAST/) were done on the *env* and *tax* nucleotide sequences to identify genetically similar sequences in the GenBank database. Percentage nucleotide identity was determined with the GAP program in the Wisconsin sequence analysis package on a UNIX workstation (Womble, 2000). Comparison of full-length STLV-3 genomes was done using STLV-3/TGE-2117 as the query sequence and the Kimura two-parameter model and a transition/transversion ratio of 2·0 implemented in the program SimPlot (Lole *et al.*, 1999). The coding regions of the *gag*, polymerase (*pol*), *env* and *tax* genes were concatenated prior to phylogenetic analysis using the DAMBE program version 4.0.75 (Xia & Xie, 2001). Phylogenetic analysis of the *tax* (180 bp) and concatenated *gag-pol-env-tax* (6795 bp) sequences was performed with the software package PAUP* version 4.0b10 (Swofford, 1998) on nucleotide alignments made using ClustalX (Thompson *et al.*, 1994), which were manually edited in Se-Alv2.0.11 (http://evolve.zoo.ox.ac.uk/) according to the codon reading frames.

Neighbour-joining (NJ) and maximum-likelihood (ML) trees were constructed under the most appropriate evolutionary model determined with Modeltest 3-06 (Posada & Crandall, 1998). For the *gag-pol-env-tax* sequences, the general time-reversible model, allowing six different substitution rate categories, with gamma distributed rate heterogeneity and an estimated proportion of invariable sites, provided the best fit to the data. For the *tax* analysis, the transitional model allowing different rates among transitions including a gamma distributed rate heterogeneity seemed to be the best-fitting model. For the *env* sequences, the Tamura Nei evolutionary model was used with gamma distributed rate heterogeneity, taking into account a different substitution rate for transversions and purine and pyrimidine transitions. The NJ trees were constructed by iteratively optimizing the model parameters, followed by a bootstrap analysis of 1000 replicates. The ML trees were constructed starting from the NJ tree with optimized parameters using an heuristic search with the nearest neighbour interchange and the subtree pruning–regrafting branch-swapping algorithm (Rogers & Swofford, 1999). LTR sequences were aligned using the CLUSTAL W program (Thompson *et al.*, 1994) and distance-based LTR trees were generated using the Kimura two-parameter model in conjunction with the NJ method in the MEGA program (version 2.1) (Kumar *et al.*, 2001). Bootstrap replicates (1000) were used to test the reliability of the final topology of the LTR tree.

Molecular clock analysis and dating. The molecular clock hypothesis was tested using the likelihood ratio test with the likelihoods for the ML and clock-like ML trees obtained in PAUP* (Swofford, 1998). The clock was tested with the best-fitting evolutionary model, estimated in Modeltest and with the NJ optimized parameters as described above. The evolutionary rate and subsequently the divergence times were estimated based on a known divergence time point and on the branch lengths of the ML clock tree according to the formula: evolutionary rate (r) = branch length (bl)/divergence time (t).

RESULTS AND DISCUSSION

STLV-like antibodies were found in the sera or plasma of 12 of 23 (52·2 %) gelada baboons. Five of the 12 seroreactive baboon samples (41·7 %) showed an STLV-1/HTLV-1-like WB profile (MTA-1 positive), one sample (8·3 %) showed

Table 2. Demographic, Western blot and PCR results for captive *Theropithecus gelada* baboons

Animal	Sex	Birth date*	Relation†	Western blot pattern‡	STLV-3 LTR PCR§	PTLV *tax* PCR	STLV-1 *env* PCR
TGE-BK1	Female	1/1/69	Mother of TGE-2117	GD21, p19, gp21, p24, p32, p36, gp46, p53	Positive	Positive	ND
TGE-BK2	Female	2/3/74		GD21, p19, gp21, p24, p32, p36, gp46, p53	Positive	Positive	ND
TGE-BK3	Male	~1963	Father of TGE-2117	GD21, p19, gp21, p24, p32, p36, gp46, p53, K55	Positive	Positive	ND
TGE-BK4	Female	~1962		GD21, p19	Negative	Negative	ND
TGE-BK5	Female	11/21/76		GD21, gp21, p24, p32, p36, gp46, p53, MTA-1	Positive	Positive	ND
TGE-BK6	Female	3/9/78		Negative	Negative	Negative	ND
TGE-BK7	Female	11/11/76		Negative	Negative	Negative	ND
TGE-2117	Female	3/2/79	Offspring of TGE-BK1 and TGE-BK3	GD21, gp21, p24, p32, p36, gp46, p53	Positive	Positive	ND
TGE-5077	Male	2/15/89	Mate of TGE-2117	Negative	ND	Negative	ND
TGE-4623	Male	4/18/94	Offspring of TGE-2117	Negative	NA	NA	NA
TGE-5191	Male	12/14/80	Offspring of TGE-2117 and TGE-5077	Negative	ND	Negative	ND
TGE-2118	Female	9/26/76		GD21, gp21, p24, p32, p36, gp46, p53	NA	NA	NA
TGE-3819	Female	1/17/91	Offspring of TGE-2118	Negative	ND	Negative	ND
TGE-650206	Female	~1963		Negative	NA	NA	NA
TGE-791603	Female	6/11/79		Negative	NA	NA	NA
TGE-811	Male	8/26/85		Negative	ND	Negative	ND
TGE-778	Male	1/4/78		GD21	ND	Negative	ND
TGE-106	Female	~1972		Negative	ND	Negative	ND
TGE-598276	Female	12/14/80		Negative	NA	NA	NA
TGE-598275	Male	11/24/85		GD21, gp21, p24, p32, p36, gp46, p53, MTA-1	ND	Positive	Positive
TGE-1884	Female	8/3/82		GD21, gp21, p24, p32, p36, gp46, p53, MTA-1	ND	Positive	Positive
TGE-2045	Female	11/5/83		GD21, gp21, p24, p32, p36, gp46, p53, MTA-1	ND	Positive	Positive
TGE-5415	Male	~1976	Mate of TGE-2045	GD21, gp21, p24, p32, p36, gp46, p53, MTA-1	ND	Positive	Positive

*Birth dates for wild-born animals are estimated to the year.

†Baboons with known pedigree or breeding information are shown.

‡K55, HTLV-2-type-specific peptide; MTA-1, HTLV-1-type-specific peptide.

§ND, Not done; NA, sample not available.

an STLV-2/HTLV-2-like WB pattern (K55 positive) and four samples (33·3 %) showed reactivity to all HTLV proteins (GD21, gp21, p24, p32, p36, gp46 and p53) except the type-specific peptides and were considered positive but untypable (Table 2, Fig. 1). The two other samples (16·6 %) reacted weakly with p19 and/or GD21 and thus were classified as indeterminate (Table 2, Fig. 1).

To characterize the STLV infections further, generic nested PTLV *tax* and STLV-1 *env* or STLV-3-specific LTR PCR was performed on DNA from 18 geladas including nine seropositive, two seroindeterminate and seven seronegative animals (Table 2). Proviral *tax* and *env* or LTR sequences were amplified from DNA samples from one gelada with the HTLV-2-like WB pattern (TGE-BK3), three with the untypable STLV/HTLV seroreactivity (TGE-BK1, TGE-BK2 and TGE-2117) and five with HTLV-1-like WB profiles (TGE-BK5, TGE-598275, TGE-1884, TGE-2045 and TGE-5415) (Table 2). Sequence analysis of the 174 bp *tax* sequences showed that TGE-BK3, TGE-BK1, TGE-BK2, TGE-2117 and TGE-BK5 were identical to the same *tax* region of STLV-3/PH969. The finding of STLV-3 in a group of animals with such a broad range of HTLV WB profiles is unusual and to our knowledge is the first report of an HTLV-1-like WB profile in samples from STLV-3-infected primates.

The *tax* and *env* sequences from the four remaining geladas with the HTLV-1-like WB profiles (TGE-598275, TGE-1884, TGE-2045 and TGE-5415) were identical and were found by BLAST analysis to be most similar to STLV-1 from mandrills and Celebes macaques (STLV-1/mnd and STLV-1/TE4, respectively). To characterize these STLV-1s further, we performed phylogenetic analysis of *env* sequences amplified from PBL DNA of these animals. This analysis showed that the gelada *env* sequences clustered with STLV-1 found naturally to infect Celebes macaques (*Macaca tonkeana*) (Fig. 2) (Ibrahim *et al.*, 1995). These four geladas were housed at the same zoo at one time on an island with Celebes crested and stump-tailed macaques (*M. nigra* and *M. arctoides*, respectively). Celebes macaques consist of seven primate species, including *M. tonkeana* and *M. nigra*, and thus the four gelada baboons were most likely infected

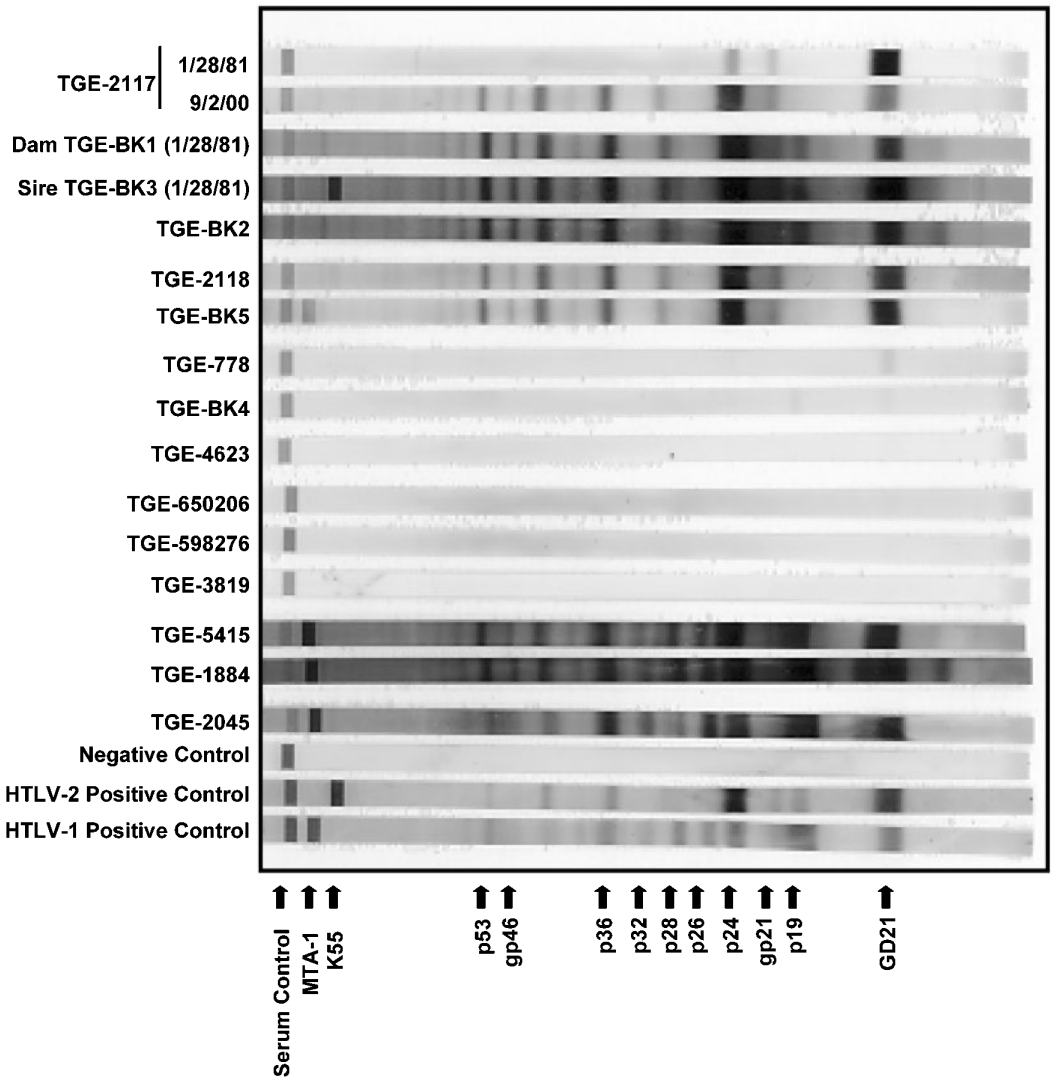


Fig. 1. Western blot serological pattern of selected gelada baboons.

via cross-species transmission from STLV-1-infected *M. nigra*. However, samples were not available from the *M. nigra* to investigate this hypothesis. None of the baboons in this study other than these four geladas were reported to have been in contact with other primate species, including other baboons.

The complete proviral sequence of the STLV-3-infected TGE-2117 gelada baboon was then determined to characterize this virus genetically. The entire genome of this novel STLV-3 (called STLV-3/TGE-2117) was 8917 bp in length with a genomic organization very similar to that of other PTLVs and included the structural, enzymic and regulatory proteins. Interestingly, as with other STLV-3s, the LTR (695 bp) was smaller than in HTLV-1 (756 bp) and HTLV-2 (764 bp), having two and not three of the 21 bp repeats (Meertens *et al.*, 2002, 2003, Meertens & Gessain, 2003). Comparison of the complete genome of STLV-3/TGE-2117 showed a nucleotide sequence similarity of only 62 % to both HTLV-1/ATK and HTLV-2/Mo (Kalyanaraman *et al.*, 1982; Poiesz *et al.*, 1980). In contrast, this new STLV-3 was closer genetically to other STLV-3s, with an overall nucleotide identity ranging from 87·2 % to 95·6 % for STLV-3/CTO-NG409 and STLV-3/PH969, respectively (Table 3). The regions of similarity between the five full-length STLV-3 genomes was also determined by similarity plot (SimPlot) analysis using STLV-3/TGE-2117 as the query sequence (Fig. 3). SimPlot analysis confirmed the high sequence identity of STLV-3/TGE-2117 to STLV-3/PH969 and also showed that regions of the *tax* gene were the most conserved while LTR sequences were the most divergent.

Comparison of the major STLV-3 nucleotide and protein sequences showed that STLV-3/TGE-2117 was most similar to STLV-3/PH969 but was equidistant from the STLV-3 from West/Central Africa (Table 3). Overall, the protease

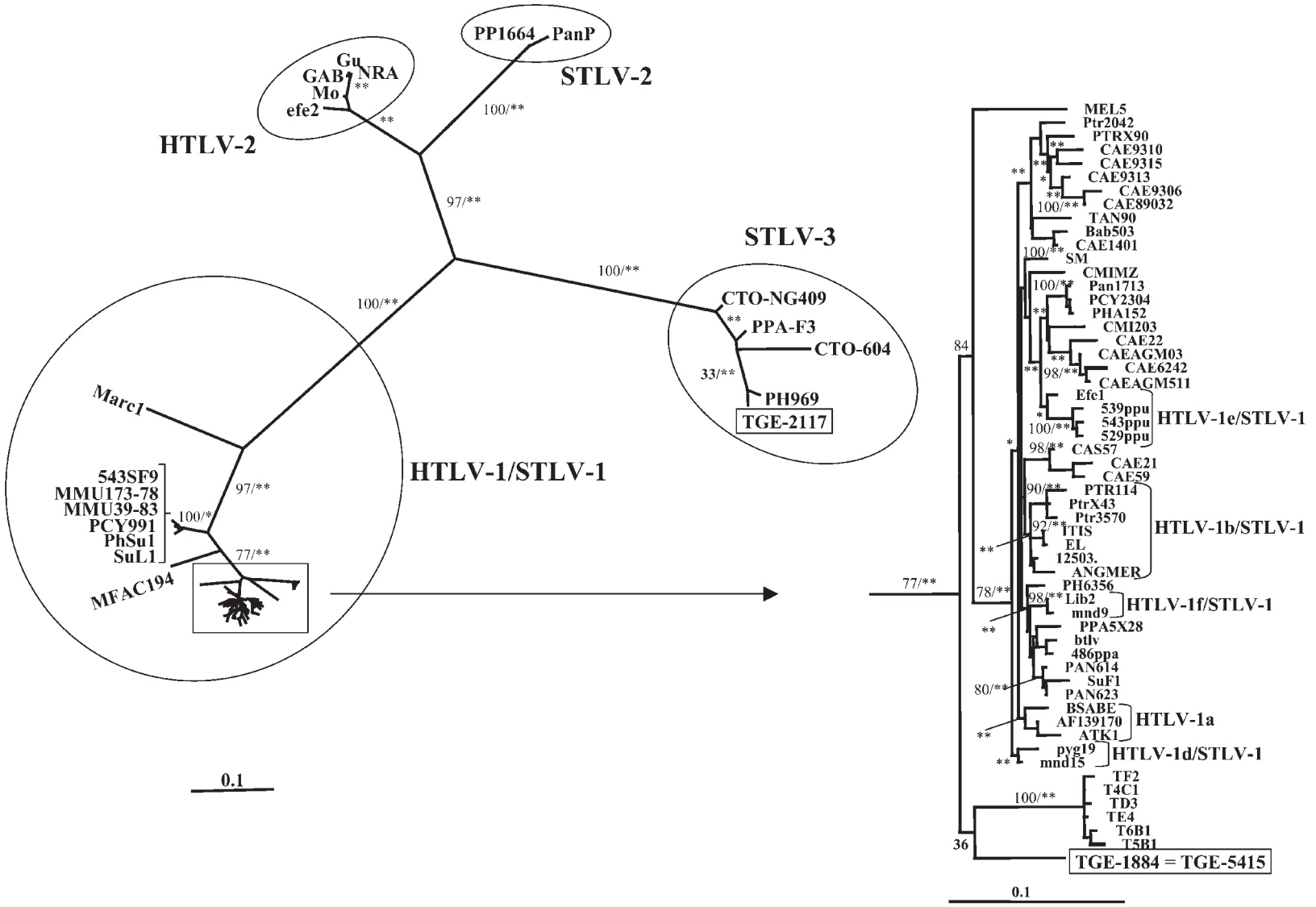


Fig. 2. Unrooted phylogenetic tree generated by maximum-likelihood analysis of STLV-1 *env* (529 bp) sequences from two gelada baboons with other available primate and human T-cell lymphotropic virus (PTLV and HTLV, respectively) sequences. Bootstrap analysis of 1000 replicates is shown on the tree branches. Statistical support for the branches was also calculated with a likelihood ratio test and the significance is expressed in *P* values as indicated on the branches (*$P<0{\cdot}05$; **$P<0{\cdot}01$). A sub-branch of the PTLV-1 radiation is shown as an exploded view in the cladogram format to illustrate better the relationship of the gelada *env* sequences with other PTLV-1 sequences.

Table 3. Nucleotide and protein sequence percentage similarity between STLV-3 strain TGE-2117 and prototypic STLV-3 strains

Protein comparisons are in parentheses.

Strain	Overall	LTR	*gag* (Gag)	*pro* (Pro)	*pol* (Pol)	*env* (Env)	*rex* (Rex)	*tax* (Tax)
PH969	95·6	93·1	96·9 (99·5)	95·1 (98·8)	95·9 (98·2)	94·7 (97·6)	92·4 (98·9)	97·7 (100·0)
CTO-604	87·7	86·4	88·2 (96·2)	85·8 (91·5)	87·6 (93·5)	85·9 (94·5)	89·1 (90·1)	91·8 (96·9)
CTO-NG409	87·2	86·1	87·4 (94·6)	84·3 (91·5)	86·5 (93·7)	87·7 (94·1)	88·3 (89·0)	90·4 (97·4)
PPA-F3	88·4	87·6	89·0 (96·5)	85·6 (91·5)	87·2 (94·1)	89·2 (96·1)	87·8 (89·6)	90·5 (97·4)

(*pro*) and *rex* genes of the STLV-3s shared the least nucleotide and protein sequence identity while the *tax* gene of the STLV-3s had the highest nucleotide and protein identity. Interestingly, the Tax protein of STLV-3/TGE-2117 was highly conserved and showed 100 % identity to that of STLV-3/PH969. The gp46 region of STLV-3/TGE-2117 was only 64 % and 57 % identical to the HTLV-1/2-type-specific MTA-1 and K55 peptides, respectively, despite a broad reactivity of sera from STLV-3-infected geladas to these peptides. In comparison, the type-specific regions of the

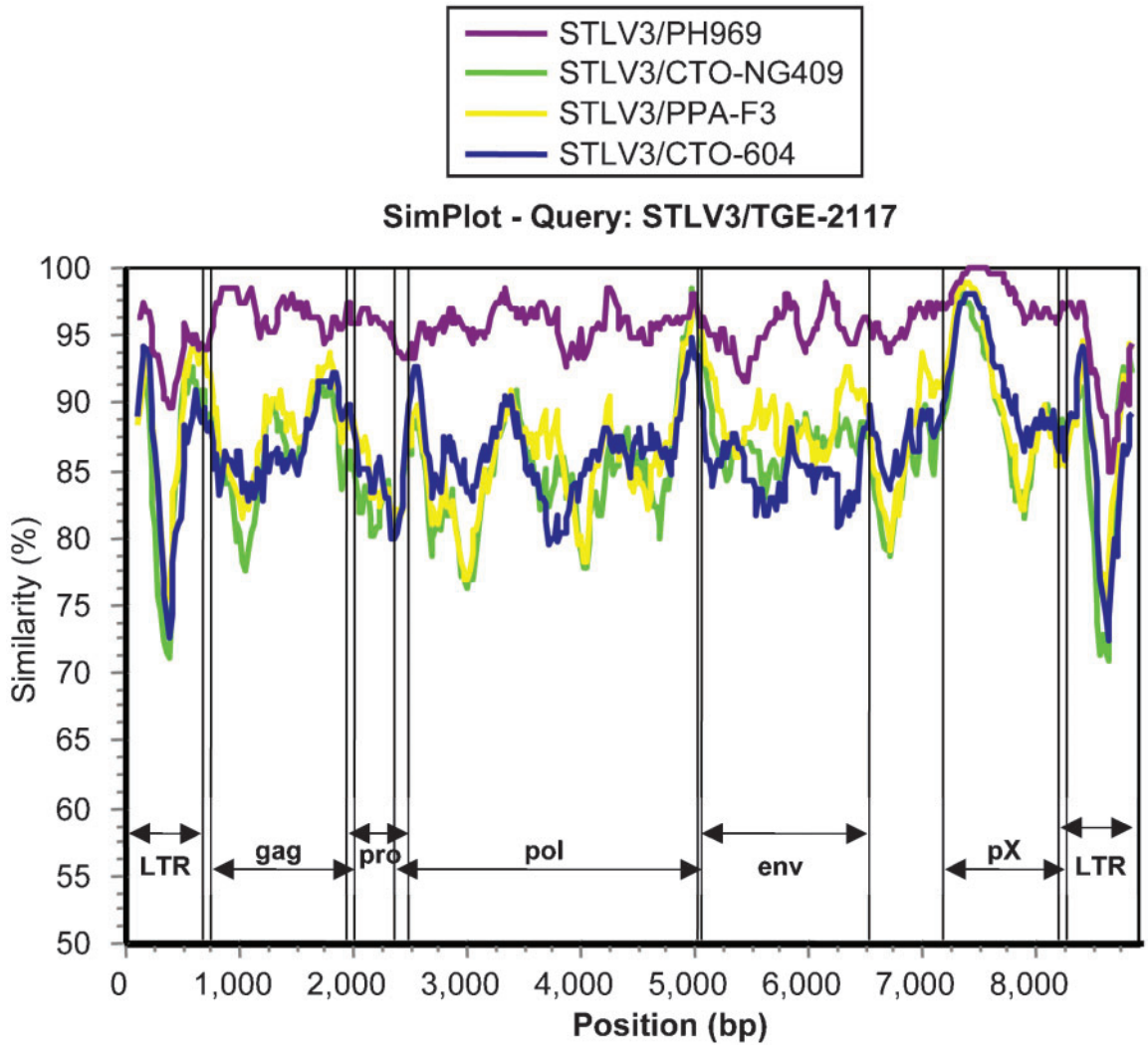
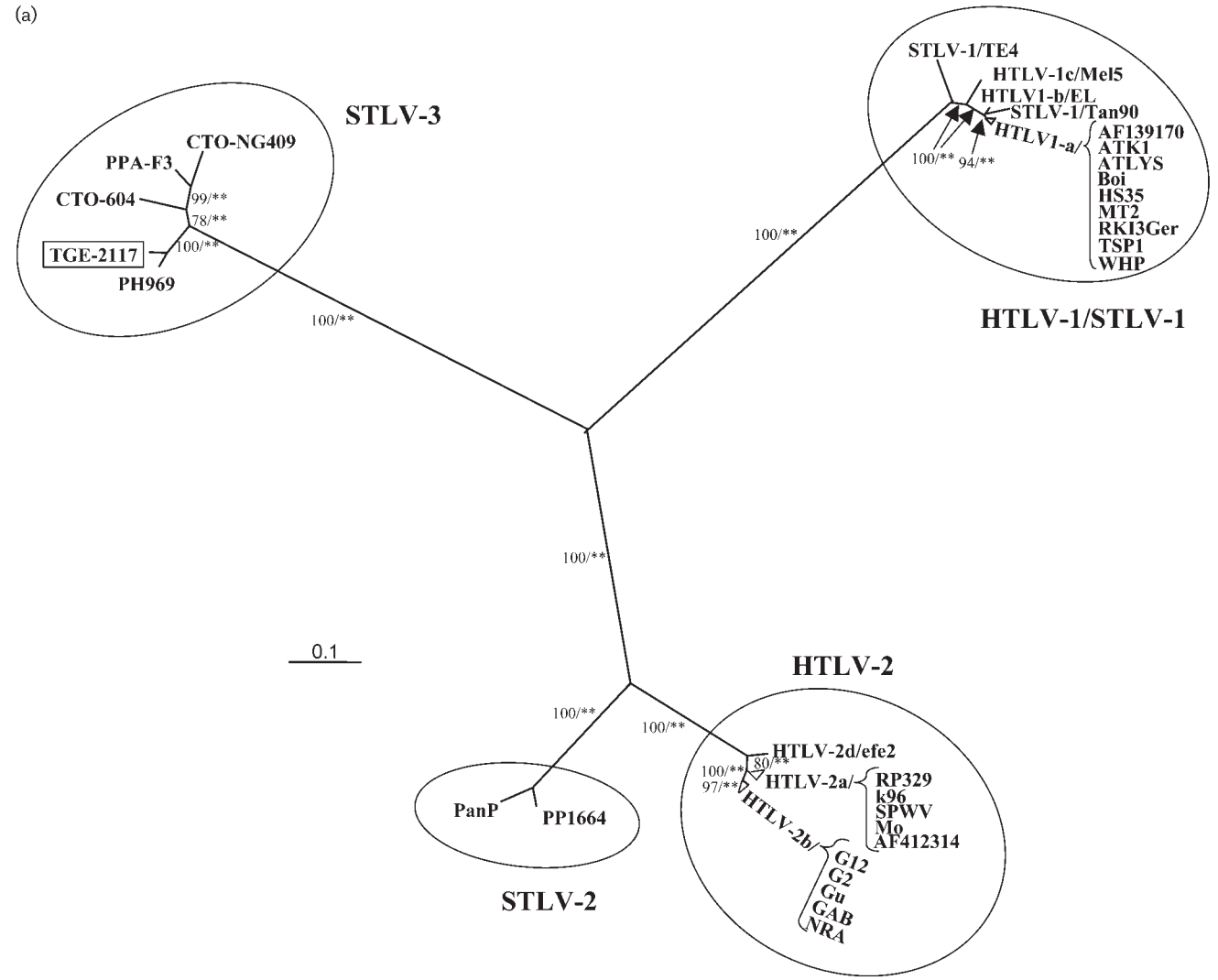


Fig. 3. Similarity plot analysis of the full-length genome of STLV-3/TGE-2117 with other STLV-3s using a 200 bp window size in 20 bp step increments on gap-stripped sequences. The Kimura two-parameter model was used with a transition-to-transversion ratio of 2·0.



STLV-3/TGE-2117 gp46 amino acid sequence were 91–100 % identical to these same regions in all four other full-length STLV-3 genomes. Ample amounts of DNA were not available from gelada TGE-BK5 for PCR and sequence analysis of the gp46 region to determine whether a genetic mechanism exists for the HTLV-1-like WB profile seen in this STLV-3-infected baboon.

Phylogenetic analysis of concatenated *gag-pol-env-tax* sequences confirmed that TGE-2117 was infected with an STLV-3-like virus that clustered tightly with strain PH969 (Fig. 4a). A 180 bp *tax* sequence of all available STLV-3 sequences showed that TGE-2117 was closest to PH969 and Hyb2210 (Fig. 4b). In both trees, the novel STLV-3/TGE-2117 was clearly related to the Ethiopian and Eritrean baboon strains, with strong bootstrap support of 100·0 in the *gag-pol-env-tax* trees (Fig. 4a).

Phylogenetic analysis of the LTR has been used before to differentiate between subtypes of HTLV-1 and HTLV-2 because it is one of the most divergent regions within PTLV genomes (Meertens *et al.*, 2002, 2003, Meertens & Gessain 2003; Switzer *et al.*, 1995; Takemura *et al.*, 2002; Vidal *et al.*, 1994). The LTR of STLV-3/TGE-2117 was 7–14 % divergent from the four other full-length STLV-3s and 4–8 % divergent from Ethiopian baboon (*Papio* species) strains Hyb2210, PHA-7556 and PHA-7550, demonstrating further the uniqueness of this new virus (data not shown). We show by phylogenetic analysis of the LTR sequences that STLV-3 is composed of two clades that are represented by geographically distinct strains from East and West/Central Africa, respectively (Fig. 5). This finding supports predictions based on analysis of *env* sequences of only the first two reported full-length STLV-3 genomes (PH969 and CTO-604) (Merteens *et al.*, 2002; Van Brussel *et al.*, 1997). Our data also confirm that STLV-3s, like STLV-1s, cluster together in highly supported clades corresponding to their geographical origin rather than their host phylogeny (Slattery *et al.*, 1999).

Although others have estimated divergence times of STLV-3 strains using the molecular clock-like behaviour of the third

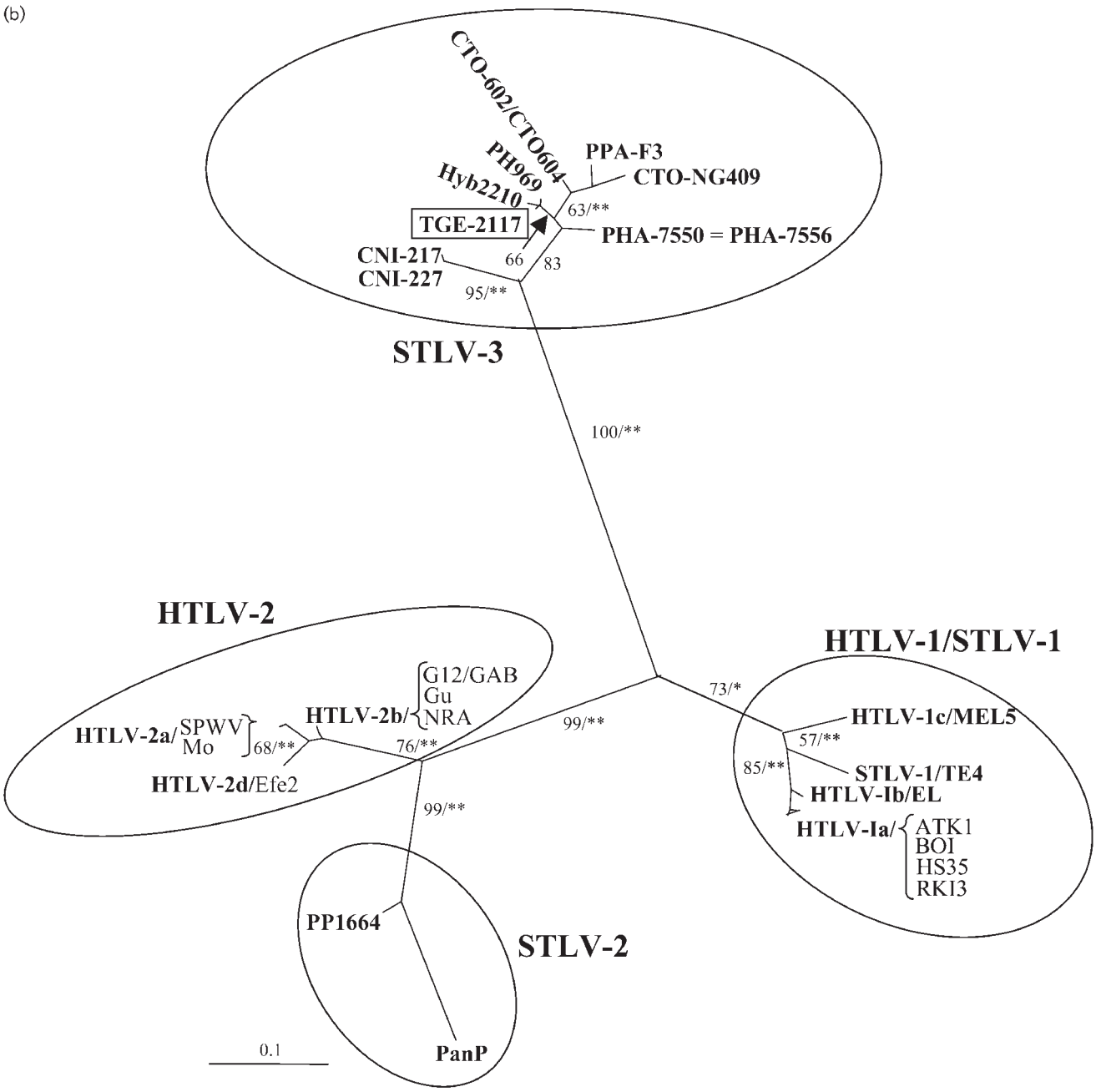


Fig. 4. Unrooted phylogenetic trees generated by maximum-likelihood analysis of the concatenated *gag-pol-env-tax* sequences (6795 bp) (a) and *tax* (180 bp) (b) with other available primate and human T-cell lymphotropic virus (PTLV and HTLV, respectively) sequences. Bootstrap analysis of 1000 replicates is shown on the tree branches. Statistical support for the branches was also calculated with a likelihood ratio test and the significance is expressed in *P* values as indicated on the branches (*$P<0{\cdot}05$; **$P<0{\cdot}01$).

codon position (3rd cp) of the *env* polyprotein alignment, we were unable to confirm the clock-like behaviour for the *env* sequences when using all available full-length STLV-3 genomes (Meertens *et al.*, 2002; Meertens & Gessain, 2003; Meertens *et al.*, 2003; Salemi *et al.*, 2000). We found that the STLV-3s evolved at different rates by using a similar 3rd cp analysis of the *env*, *tax*, or *gag-pol-env-tax* sequences and hence the molecular clock was disturbed. None the less, we enforced the molecular clock, resulting in a less reliable estimate and broader confidence intervals, but allowing us to estimate the date of the different tree nodes using a value of 40 000–60 000 years ago (ya) as a starting date for

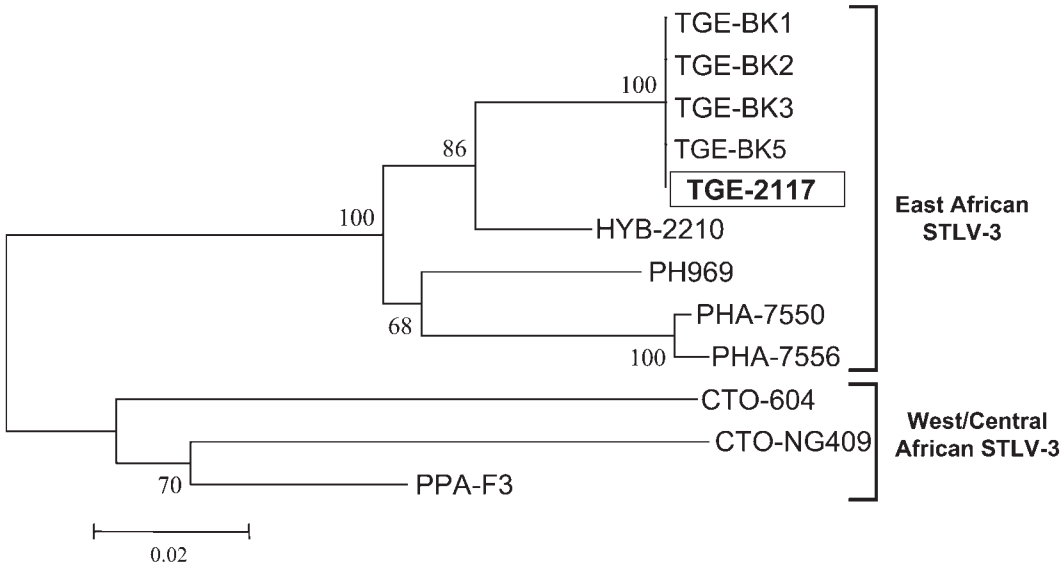


Fig. 5. Unrooted phylogenetic tree generated by neighbour-joining analysis of the LTR (399 bp) of STLV-3/TGE-2117 with other available primate and human T-cell lymphotropic virus (PTLV and HTLV, respectively) sequences. Bootstrap analysis of 1000 replicates is shown on the tree branches; only values >60 are shown.

the origin of the Melanesian HTLV-1. The evolutionary rate for STLV-3/TGE-2117 was estimated at around $0{\cdot}79–1{\cdot}18 \times 10^{-6}$ and $1{\cdot}72–2{\cdot}56 \times 10^{-6}$ nucleotide substitutions per site per year for the *gag-pol-env-tax* and *env* 3rd cp ML trees, respectively. Using this value, we estimated that the origin of all STLV-3s is 73 628–109 809 ya for the *gag-pol-env-tax* concatamer and 85 634–128 451 ya for the *env* 3rd cp, suggesting an ancient origin. The origin of the East African and the West/Central African STLV-3 subgroups occurred about 20 450–30 650 and 61 860–92 710 ya, respectively, based on the *gag-pol-env-tax* concatamer data. These divergence dates are slightly lower than those reported elsewhere (Meertens *et al.*, 2002; Meertens & Gessain, 2003) and may be due to the use of a constant-rate model in previous studies, which has been shown to overestimate divergence times (Yang, 1996).

Since STLV-3 is the most recently identified PTLV, very little is known about its transmission or genomic stability in infected animals. Sera archived in 1981 from five seropositive geladas from the same zoo included a captive-born mother (TGE-BK1), a wild-born father (TGE-BK3), their female offspring (TGE-2117) who was born in 1979 and two other females (TGE-BK2 and TGE-BK5). LTR sequences from the mother and TGE-2117 were identical and clustered together with 100 % bootstrap support by phylogenetic analysis (Fig. 5), suggesting a mother-to-offspring transmission of STLV-3 infection in this pair. The LTR sequences from two other STLV-3-infected female geladas (TGE-BK2 and TGE-BK5) were essentially identical to male gelada TGE-BK3, having >99 % nucleotide identity, suggesting a horizontal transmission. We also examined the genetic stability of STLV-3 over time by analysis of LTR sequences obtained from blood specimens from TGE-2117 and her parents that were collected 20 years apart. Interestingly, the LTR sequences for both parental baboons and the descendant TGE-2117 were identical, suggesting a high genetic stability of the STLV-3 genomes similar to that seen for other PTLVs (Slattery *et al.*, 1999). Viral persistence by oligoclonal expansion of infected lymphocytes within animals is believed to confer the high genetic stability of the PTLVs (Gabet *et al.*, 2003). This finding contrasts with the genetic mutability of other exogenous retroviruses, such as the human immunodeficiency virus, which has mutation rates 3–4 logs higher (Sharp *et al.*, 1994).

This study reports the serological and molecular characterization of a novel STLV-3 in Ethiopian *T. gelada* baboons and the full-length sequence analysis of this distinct STLV-3 found in animal TGE-2117. This report is the first to find and characterize STLV-3 and STLV-1 infection in this baboon species. The absence of STLV infection of geladas observed by others (Takemura *et al.*, 2002) may be explained by differences in the sensitivities of the antibody screening assays used in each study and/or the sampling of gelada specimens originating from different Ethiopian locations (Takemura *et al.*, 2002). Interestingly, the STLV-3 LTR sequences obtained from five gelada baboons in the current study were more genetically related to an STLV-3 LTR sequence found in a wild-born Ethiopian *P. hamadryas* and *P. anubis* hybrid (Hyb2210) than to those found in either Eritrean or Ethiopian hamadryas baboons. These results suggest that the STLV-3s in the gelada and hybrid baboons are distinct from those found in hamadryas baboons and that the origin of the STLV-3s in the hybrid baboons may be from the gelada and not from infected

hamadryas baboons as previously proposed (Takemura *et al.*, 2002).

Although fossil evidence suggests that the extant Theropithecus baboons once radiated across the African continent and exceeded the combined ranges of the *Papio* species, the gelada's current habitat is restricted to the high plateaus of Ethiopia (Pickford, 1993). In this region, gelada baboons are commonly sympatric with anubis and hamadryas baboons but occupy different ecological niches. In contrast, anubis and hamadryas baboons commonly interbreed, as reported elsewhere (Takemura *et al.*, 2002). None the less, these overlapping habitats may have provided opportunities for cross-species transmission of STLV-3 infections to or from each of these baboon species.

Questions about whether humans are infected with STLV-3-like viruses are raised by the widespread distribution of STLV-3 across Africa and its probable ancient existence (Goubau *et al.*, 1994; Meertens *et al.*, 2001, 2002; Meertens & Gessain, 2003; Takemura *et al.*, 2002; Van Dooren *et al.*, 2001), the ability of STLV-3 to grow in human cells *in vitro* (Goubau *et al.*, 1994) and evidence for multiple interspecies transmissions that have occurred among the PTLV-1s (Gessain *et al.*, 2002; Slattery *et al.*, 1999). The presence of STLV-3-like viruses in primates such as geladas that are commonly hunted for their coats or shot while raiding crops may support the possibility of such infection in humans (Pickford, 1993). In addition, the human population density on the Ethiopian plateau is among the highest in sub-Saharan Africa, thus placing the gelada in coexistence with humans living in this region and putting humans at risk for cross-species infections with STLV and other simian retroviruses. The finding by us and others of STLV-1 and STLV-3 infections in captive baboons also suggests that people working with these primates may be at risk for occupational exposure to these viruses (Goubau *et al.*, 1994, Meertens & Gessain, 2003; Takemura *et al.*, 2002).

HTLV-1, HTLV-2 and HTLV-indeterminate WB profiles have all been observed in Ethiopian populations (Buckner *et al.*, 1992; Vrielink *et al.*, 1995), but it is not known if any of these reactive samples are associated with STLV-3-like infections. While limited studies of 24 African and 325 non-African human specimens with HTLV-seroindeterminate WB results have detected no STLV-3-like infection (Busch *et al.*, 2000; Vandamme *et al.*, 1997), additional studies of larger numbers of human samples with not only indeterminate but also HTLV-1- and -2-like WB profiles will be necessary to assess fully the rate of infection with STLV-3-like viruses. Our finding of one STLV-3-infected gelada baboon with an HTLV-1-like WB profile and a previous report of an STLV-1-infected baboon with an HTLV-2-like serotype (Mahieux *et al.*, 2000) also raises questions about the accuracy of viral typing of PTLV infections by the currently used HTLV-1- and HTLV-2-type-specific Env peptides spiked onto HTLV-1 WB strips. Our results also highlight the need for better STLV-3-specific diagnostic assays to facilitate screening for STLV-3-like infections in both humans and NHPs.

ACKNOWLEDGEMENTS

We thank the Zoological Society of San Diego's Center for Reproduction of Endangered Species for providing DNA and serum samples from gelada baboons for this study and also the veterinary and clinical laboratory staff at the Los Angeles, Oakland, Gladys Porter and Folsom Children's Zoos for providing us with additional gelada specimens.

Use of trade names is for identification only and does not imply endorsement by the US Department of Health and Human Services, the Public Health Service, or the Centers for Disease Control and Prevention.

REFERENCES

Buckner, C., Roberts, C. R., Foung, S. K. H. & 7 other authors (1992). Immune responsiveness to the recombinant envelope epitopes of human T lymphotropic virus types I and II in diverse geographic populations. *J Infect Dis* **166**, 1160–1163.

Busch, M., Switzer, W. M., Khabbaz, R. F., Kaplan, J. E., Heneine, W. & the Retrovirus Epidemiology Donor Study Group. (2000). Absence of evidence of infection with divergent simian T-cell lymphotropic viruses in blood donors with seroindeterminate human T-cell lymphotropic type 1 results. *Transfusion* **40**, 443–449.

Gabet, A. S., Gessain, A. & Wattel, E. (2003). High simian T-cell leukemia virus type 1 proviral loads combined with genetic stability as a result of cell-associated provirus replication in naturally infected, asymptomatic monkeys. *Int J Cancer* **107**, 74–83.

Gessain, A., Meertens, L. & Mahieux, R. (2002). Molecular epidemiology of human T cell leukemia/lymphoma viruses type 1 and type 2 (HTLV-1/2) and related simian retroviruses (STLV-1, STLV-2 and STLV-L/3). In *The Molecular Epidemiology of Human Viruses*, 1st edn, pp. 121–165. Edited by T. Leitner. Boston: Kluwer Academic Publishers.

Giri, A., Markham, P., Digilio, L., Hurteau, G., Gallo, R. C. & Franchini, G. (1994). Isolation of a novel simian T-cell lymphotropic virus from *Pan paniscus* that is distantly related to the human T-cell leukemia/lymphotropic virus types I and II. *J Virol* **68**, 8392–8395.

Goubau, P., Van Brussel, M., Vandamme, A. M., Liu, H. F. & Desmyter, J. (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 U S A* **91**, 2848–2852.

Heneine, W., Switzer, W. M., Sandstrom, P. & 9 other authors (1998). Identification of a human population infected with simian foamy viruses. *Nat Med* **4**, 403–407.

Ibrahim, F., De Thé, G. & Gessain, A. (1995). Isolation and characterization of a new simian T-cell leukemia virus type 1 from naturally infected Celebes macaques (*Macaca tonkeana*): complete nucleotide sequence and phylogenetic relationship with the Australo-Melanesian human T-cell leukemia virus type 1. *J Virol* **69**, 6980–6993.

Kalyanaraman, V. S., Sarngadharan, M. G., Robert-Guroff, M., Miyoshi, I., Golde, D. & Gallo, R. C. (1982). A new subtype of human T-cell leukemia virus (HTLV-II) associated with a T-cell variant of hairy cell leukemia. *Science* **218**, 571–573.

Khabbaz, R. F., Heneine, W., George, J. R. & 7 other authors (1994). Brief report: infection of a laboratory worker with simian immunodeficiency virus. *N Eng J Med* **330**, 172–177.

Kumar, S., Tamura, K., Jakobsen, I. B. & Nei, M. (2001). MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**, 1244–1245.

Lerche, N. W., Switzer, W. M., Yee, J. L., Shanmugam, V., Rosenthal, A. N., Chapman, L. E., Folks, T. M. & Heneine, W. (2001). Evidence of infection with simian type D retrovirus in persons occupationally exposed to nonhuman primates. *J Virol* **75**, 1783–1789.

Liu, H.-F., Vandamme, A. M., Van Brussel, M., Desmyter, J. & Goubau, P. (1994). New retroviruses in human and simian T-lymphotropic viruses. *Lancet* **344**, 265–266.

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W. & Ray, S. C. (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.

Mahieux, R., Chappey, C., Georges-Courbot, M. C., Dubreuil, G., Mauclere, P., Georges, A. & Gessain, A. (1998). Simian T-cell lymphotropic virus type 1 from *Mandrillus sphinx* as a simian counterpart of human T-cell lymphotropic virus type 1 subtype D. *J Virol* **72**, 10316–10322.

Mahieux, R., Chappey, C., Meertens, L., Mauclère, P., Lewis, J. & Gessain, A. (2000). Molecular characterization and phylogenetic analyses of a new simian T cell lymphotropic virus type 1 in a wild-caught African baboon (*Papio anubis*) with an indeterminate STLV type 2-like serology. *AIDS Res Hum Retroviruses* **16**, 2043–2048.

Meertens, L. & Gessain, A. (2003). Divergent simian T cell lymphotropic virus type 3 in wild-caught *Papio hamadryas papio* from Senegal: widespread distribution of STLV-3 in Africa. *J Virol* **77**, 782–789.

Meertens, L., Rigoulet, J., Mauclere, P. & 8 other authors (2001). Molecular and phylogenetic analyses of 16 novel simian T cell leukemia virus type 1 from Africa: close relationship of STLV-1 from *Allenopithecus nigroviridis* to HTLV-1 subtype B strains. *Virology* **287**, 275–285.

Meertens, L., Mahieux, R., Mauclère, P., Lewis, J. & Gessain, A. (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.

Meertens, L., Shanmugam, V., Gessain, A., Beer, B. E., Tooze, Z., Heneine, W. & Switzer, W. M. (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.

Miyoshi, I., Yoshimoto, S., Fujishita, M., Taguchi, H., Kubonishi, I., Niiya, K. & Minezawa, M. (1982). Natural adult T-cell leukemia virus infection in Japanese monkeys [letter]. *Lancet* ii, 658.

Nerrienet, E., Meertens, L., Kfutwah, A., Foupouapouognigni, Y. & Gessain, A. (2001). Molecular epidemiology of simian T-lymphotropic virus (STLV) in wild-caught monkeys and apes from Cameroon: a new STLV-1, related to human T-lymphotropic virus subtype F, in a *Cercocebus agilis*. *J Gen Virol* **82**, 2973–2977.

Pickford, M. (1993). Climatic change, biogeography and Theropithecus. In *Theropithecus: the Rise and Fall of a Primate Genus*,1st edn, pp.227–243. Edited by N. G. Jablonski. Cambridge: Cambridge University Press.

Poiesz, B. J. W., Ruscetti, F., Gazdar, A. F., Bunn, P. A., Minna, J. D. & Gallo, R. C. (1980). Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci U S A* **77**, 7415–7419.

Posada, D. & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Rogers, J. S. & Swofford, D. L. (1999). Multiple local maxima for likelihoods of phylogenetic trees: a simulation study. *Mol Biol Evol* **16**, 1079–1085.

Salemi, M., Desmyter, J. & Vandamme, A. M. (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.

Sharp, P. M., Robertson, D. L., Gao, F. & Hahn, B. H. (1994). Origins and diversity of human immunodeficiency viruses. *AIDS* **8**, S27–S42.

Slattery, J. P., Franchini, G. & Gessain, A. (1999). Genomic evolution, patterns of global dissemination, and interspecies transmission of human and simian T-cell leukemia/lymphotropic viruses. *Genome Res* **9**, 525–540.

Switzer, W. M., Pieniazek, D., Swanson, P., Samdal, H. H., Soriano, V., Khabbaz, R., Kaplan, J. E., Lal, R. B. & Heneine, W. (1995). Phylogenetic relationship and geographic distribution of multiple human T-cell lymphotropic virus type II subtypes. *J Virol* **69**, 621–632.

Swofford, D. L. (1998). PAUP*. Phylogenetic analysis using parsimony (* and other methods), version 4. Sunderland, MA: Sinauer Associates.

Takemura, T., Yamashita, M., Shimada, M. K., Ohkura, S., Shotake, T., Ikeda, M., Miura, T. & Hayami, M. (2002). High prevalence of simian T-lymphotropic virus type L in wild Ethiopian baboons. *J Virol* **76**, 1642–1648.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (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.

Van Brussel, M., Goubau, P., Rousseau, R., Desmyter, J. & Vandamme, A. M. (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.

Van Brussel, M., Salemi, M., Liu, H. F., Goubau, P., Desmyter, J. & Vandamme, A. M. (1999). The discovery of two new divergent STLVs has implications for the evolution and epidemiology of HTLVs. *Rev Med Virol* **9**, 155–170.

Vandamme, A. M., Liu, H. F., Van Brussel, M., De Meurichy, W., Desmyter, J. & Goubau, P. (1996). The presence of a divergent T-lymphotropic virus in a wild-caught pygmy chimpanzee (*Pan paniscus*) supports an African origin for the human T-lymphotropic/simian T-lymphotropic group of viruses. *J Gen Virol* **77**, 1089–1099.

Vandamme, A. M., Van Laethem, K., Liu, H. F. & 7 other authors (1997). Use of a generic polymerase chain reaction assay detecting human T-lymphotropic virus (HTLV) types I, II and divergent simian strains in the evaluation of individuals with indeterminate HTLV serology. *J Med Virol* **52**, 1–7.

Van Dooren, S., Salemi, M., Pourrut, X., Peeters, M., Delaporte, E., Van Ranst, M. & Vandamme, A. M. (2001). Evidence for a second simian T-cell lymphotropic virus type 3 in *Cercopithecus nictitans* from Cameroon. *J Virol* **75**, 11939–11941.

Vidal, A. U., Gessain, A., Yoshida, M., Tekaia, F., Garin, B., Guillemain, B., Schulz, T., Farid, R. & De Thé, G. (1994). Phylogenetic classification of human T cell leukaemia/lymphoma virus type I genotypes in five major molecular and geographical subtypes. *J Gen Virol* **75**, 3655–3666.

Vrielink, H., Sisay, Y., Reesnink, H. W., Woerdeman, M., Winkel, C., de Leeuw, S. J., Lelie, P. N. & van der Poel, C. L. (1995). Evaluation of a combined lysate/recombinant antigen anti-HTLV-I/II ELISA in

high and low endemic areas of HTLV-I/II infection. *Transfusion Med* **5**, 135–137.

Womble, D. D. (2000). GCG: the Wisconsin Package of sequence analysis programs. *Methods Mol Biol* **132**, 3–22.

Xia, X. & Xie, Z. (2001). DAMBE: data analysis in molecular biology and evolution. *J Heredity* **92**, 371–373.

Yang, Z. (1996). Among-site rate variation and its impact on phylogenetic analyses. *Tree* **11**, 367–372.