JCM Accepted Manuscript Posted Online 27 January 2021
J Clin Microbiol doi:10.1128/JCM.02701-20

Novel genetic constructs for production of recombinant HTLV-1/2 antigens and evaluation of their reactivity to plasma samples from HTLV1-infected patients

Ueriton Dias de Oliveira,[a,b] Fred Luciano Neves Santos,[c] Bernardo Galvão-Castro,[d] Marco Aurelio Krieger,[e] Nilson Ivo Tonin Zanchin [a#]

[a]Laboratory of Structural Biology and Protein Engineering, Carlos Chagas Institute, Oswaldo Cruz Foundation, Curitiba, Paraná, Brazil
[b]Cellular and Molecular Biology Postgraduate Program, Federal University of Paraná, Curitiba, Paraná, Brazil
[c]Advanced Laboratory of Public Health, Gonçalo Moniz Institute (IGM), FIOCRUZ-BA, Salvador, Bahia, Brazil
[d]Integrated and Multidisciplinary HTLV Center, Bahiana School of Medicine and Public Health (EBMSP), Salvador, Bahia, Brazil
[e]Molecular Biology Institute of Paraná, Curitiba, Paraná, Brazil

Running title: New genetic constructs for HTLV recombinant antigens

#Address correspondence to Nilson I. T. Zanchin, nilson.zanchin@fiocruz.br

Abstract

HTLV-1 can cause life-threatening diseases for which there are no effective treatments. Prevention of HTLV-1 infection requires massive testing of pregnant women, blood for transfusion and, organs for transplantation as well as safe sex. In this context, serological assays are widely used for monitoring HTLV-1 infections. Despite the necessity of recombinant antigens to compose serological tests, there is little information available on procedures to produce recombinant HTLV1/2 antigens for serological diagnostic purposes. In this work, we tested a series of genetic constructions to select those more amenable for production in bacterial systems. To overcome the constraints to express sections of viral envelope proteins in bacteria, we have used the p24 segment of the gag protein as a scaffold to display the immunogenic regions of gp46 and gp21. Nine recombinant antigenic proteins derived from HTLV-1 and five derived from HTLV-2 were successfully purified. The HTLV-1 antigens showed high efficiency in discriminating HTLV-positive from HTLV-negative samples using ELISA. Interestingly, HTLV-1-positive samples showed a high level of cross-reaction with HTLV-2 antigens. This finding is explained by the high sequence conservation between the structural proteins of these two highly related viruses. In summary, the results presented in this work provide a detailed description of the methods used to produce recombinant HTLV-1 and HTLV-2 antigens and demonstrate that the HTLV-1 antigens show strong potential for serological diagnosis of HTLV1 infections.

Journal of Clinical Microbiology
JCM
Journal of Clinical Microbiology

Introduction

Four types of human T-cell leukemia viruses have already been identified (HTLV-1 to HTLV-4). HTLV-1 is an especially threatening oncogenic virus (1) directly associated with the development of adult T-cell leukemia (ATL) and, HTLV-associated myelopathy/tropical spastic paraparesis (HAM/TSP), uveitis and infective dermatitis (2, 3). Although HTLV-2 has been isolated from a patient with hairy cell leukemia (4) and patients with HAM/TSP (5, 6), a clinical correlation between this virus and the symptoms has not been established. HTLV-1 shows a wide geographic distribution, being found in all continents although with significant heterogeneity (7–11). In some endemic regions of high prevalence such as South America, West Africa, Japan, Iran and some locations in the USA and China, the estimates of infected people may reach over 1% of the population (12–18). HTLV-2 distribution usually co-localizes with HTLV-1 and has been reported to show a higher prevalence in intravenous drug users (11, 19) and in indigenous populations from the Americas (20, 21). HTLV-3 and HTLV-4 have been reported only for two Central African countries. They were initially isolated from non-human primate hunters living in rural areas of Cameroon (22, 23). HTLV-4 was subsequently isolated also from non-human primate hunters in Gabon (24). So far, HTLV-3 and HTLV-4 have not been associated with diseases (25).

HTLV transmission may occur via mother to child during breastfeeding, during sexual contact, transfusion of contaminated blood products, organ transplantation and contaminated injections. Both ATL and HAM/TSP are life-threatening diseases for which there is no effective therapy available (26). Their prevention depends upon continuous monitoring of pregnant women's health, blood for transfusion, organs for transplantation and, safe sex. In countries with high prevalence and large populations, such as Brazil and Japan, the estimates of infected people can reach up to ~800.000 and ~1 million,

Journal of Clinical Microbiology

respectively (7, 13, 17). Therefore, periodic monitoring of HTLV infections in these countries requires the screening of an extremely large number of samples. Serological screens are usually performed on automated systems based on ELISA (enzyme-linked immunosorbent assay), particle agglutination or, chemiluminescence methods. Confirmation diagnosis usually relies on tests based on antigen profiling by Western blotting, Line immunoassays or on detection of viral genome by nucleic acid technology assays.

Several commercial serological tests for HTLV screening are available with reported good performance for HTLV detection (27–34). However, despite the epidemiological importance of HTLV-1/2 infections, the literature is poor on information about the production of recombinant HTLV-1/2 antigens for serological diagnostic purposes. Most studies have been restricted to production of a few different segments of the Env protein. Initial studies have described the expression of truncated gp21 and gp46 protein segments (35–38), while subsequent studies have described the production of different gp21 and gp46 immunogenic regions in fusion with GST (glutathione-S-transferase) (39, 40). In addition, a fusion protein comprising sections of p24 (residues 14-139) and gp46 (residues 197-295) was also produced and shown to discriminate HTLV-1 positive from negative samples (41). The genetic constructs for the recombinant proteins mentioned above focused on the immunoreactive segments without considering the structural constraints of the native proteins. Furthermore, *Schistosoma japonicum* GST shares 82% amino acid identity to *S. mansoni* GST. Therefore, GST should be avoided as a fusion protein for epitope display since it may cause cross-reaction with schistosomiasis in endemic regions.

Considering that efficient production of low-cost recombinant HTLV-1/2 antigens will facilitate epidemiological studies in countries with limited financial resources, we designed a series of constructions for bacterial expression of HTLV-1/2 antigens. Most constructs are based on the p24 protein as a scaffold to display the immunogenic regions of gp46 and

Journal of Clinical Microbiology

gp21 taking into consideration local structural features to favor soluble expression in bacterial systems. We describe the purification of nine recombinant antigenic proteins derived from HTLV-1 and five derived from HTLV-2. The recombinant HTLV-1 antigens showed high efficiency to distinguish HTLV-1-positive from HTLV-1-negative samples using ELISA. In addition, due to the high sequence similarity, HTLV-1-positive plasma samples showed high reactivity with HTLV-2 antigens. Overall, the present work provides detailed information on the methods to produce recombinant HTLV-1 and HTLV-2 antigens. Furthermore, it demonstrates that especially the HTLV-1 antigens present a high potential for serological diagnosis of HTLV-1 infections.

Materials and methods

Selection of antigenic regions and plasmid constructs

Selection of antigenic regions of HTLV-1/2 structural proteins was based on epitope mapping studies reported in the literature (Table 1). The coding sequences of the genetic constructs were acquired from GenScript (Piscataway, NJ – USA) as synthetic genes optimized for *Escherichia coli* expression. The synthetic genes encoding antigens 2-2, 2-4, 3-1 and 3-3 were acquired into the plasmid pUC57 and subcloned into the plasmid pET28a using the *Bam*HI and *Hind*III restriction enzymes. Expression vectors 2-5 and 2-7 were constructed by fusing the gp21 and gp46 coding sequences to the 3' end of gap p24 coding sequence into the expression vector 2-4. For these cloning, gp21 and gp46 coding sequences were amplified by PCR using primers gp21F (5' GGCTGAATTCAGTGGACGCTCCGGGCTATGACCCG 3'), gp21R (GATCAAGCTTTCAGCGAGCCCACTGGGACAGGCCCAG), gp46F (5' AACGCGGATCCGCGAATTCCGTCGACGCTCCAGGCTATGATCC 3') and gp46R (5' AACCGAAGCTTTTATCAACTTCCCGTAATCCGACCTGCAAC 3'). The PCR products were digested with the restriction enzymes *EcoR*I and *Hind*III and inserted into the respective sites of expression vector 2-4. The resulting expression vectors produced proteins with a poly-histidine tag in the N-terminal region. Except for construct 3-4, all other HTLV-2 synthetic genes were acquired in plasmid pET28a, cloned into the *Nco*I and *Xho*I restriction sites. The antigens expressed from these constructs contain a poly-histidine tag in the C-terminal region. The synthetic gene for HTLV-2 antigen 3-4 was designed as described by Christensen *et al.* (42) and cloned into the *Nco*I and *Xho*I restriction sites of pET28a. It encodes a 10-histidine residue segment in the N-terminal region. The cysteine residues of positions 107 and 136 were replaced by serine.

Journal of Clinical Microbiology

Expression of and purification of recombinant proteins

Expression vectors were transformed into the *E. coli* strain BL21-Star (DE3) (Invitrogen/ThermoFisher Scientific, USA) for production of the recombinant proteins. *E. coli* cultures were maintained in Lysogeny Broth (LB) supplemented with kanamycin (50 µg/ml) at 37°C under rotational agitation for 16 hours. For protein expression, fresh pre-cultures were diluted 1:20 in 1-2 liters of LB medium and incubated under the same conditions up to an OD_{600} of approximately 0.6-0.8. IPTG (isopropyl β-D-1-thiogalactopyranoside) was added to a final concentration of 500 µM and the cultures incubated for 4 hours at 37°C. Cells were collected by centrifugation and stored at -20°C. For purification of soluble antigens, the cell pellet was suspended in affinity chromatography buffer A (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole) in a ratio of ~3 ml of buffer per g of cells and purified using His-Trap HP 1 mL columns on an ÄKTA purifier system (GE Healthcare, USA). Proteins were eluted with a 20-500 mM imidazole gradient and analyzed SDS-PAGE.

Antigens 2-2, 2-5, 2-7, 2-14, 3-1, 3-5 and 3-7 were purified from inclusion bodies. Solubilization of inclusion bodies was tested using sequential washes with increasing urea concentrations. Initially, following lysis and centrifugation, the pellets of inclusion bodies were suspended in 50 mM Tris-HCl, pH 8.0, 300 mM NaCl and 2% Triton X-100 (v/v), and sonicated six times using a QSONICA Q7 ultrasonic homogenizer (QSONICA Newtown, CT - USA) with 30% amplitude for 30 sec with 1 min intervals on ice. The extract was centrifuged at 20,000 xg for 15 min at 4°C. The supernatant was transferred to a new tube and the pellet suspended in 50 mM Tris-HCl, pH 8.0, 300 mM NaCl and 1 M urea, and sonicated again as described above. The supernatant was transferred to a new tube and the pellet suspended in 50 mM Tris-HCl, pH 8.0, 300 mM NaCl and 4 M urea. After a new

Journal of Clinical Microbiology

round of sonication and centrifugation, the supernatant was transferred to a new tube and pellet suspended in 50 mM Tris-HCl, pH 8.0, 300 mM NaCl and 8 M urea. Efficiency of solubilization was evaluated by analysis using SDS-PAGE. Antigen 2-2, 2-5, 2-7, 2-14 and 3-1 were purified by affinity chromatography as described above from the 4 M urea fractions. Antigens 3-5 and 3-7 were purified from the 8 M urea fractions.

Mass spectrometry

Proteins were initially reduced and alkylated and in-gel digested with trypsin. Mass spectrometry was performed by liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) on a Thermo Scientific Easy-nLC 1000 system coupled to an LTQ Orbitrap XL ETD (mass spectrometry facility RPT02H PDTIS, Fiocruz Paraná). The ten most intense ions were sequentially isolated and fragmented in the linear ion trap using collision-induced dissociation at a target value of 30,000. Peak list picking, protein identification, and validation were obtained using the MaxQuant platform (version 1.5.2.8).

Ethics Statement.

Approval for this study was granted by the Institutional Review Board (IRB) for Human Research at the Gonçalo Moniz Institute (IGM), Oswaldo Cruz Foundation (FIOCRUZ), Salvador, Bahia (BA), Brazil (protocol no. 67809417.0.0000.0040).

Sample collections

We used anonymized human plasma obtained from the biorepository of the Integrated and Multidisciplinary HTLV Center (CHTLV) at Bahiana School of Medicine and Public Health (EBMSP), Salvador, Brazil. Based on an expected error of 2.5%, sensitivity and specificity of 99% and a 95% confidence interval, the estimated minimum sample set for this study

was 61 samples from non-infected and 61 from HTLV-infected individuals. The study was carried out using plasma from 94 HTLV-1-infected individuals previously screened for antibodies against HTLV-1/2 using an enzyme-linked immunosorbent assay (Ortho HTLV-1/HTLV-2 Ab-Capture ELISA systems; Ortho-Clinical Diagnostic, Raritan, NJ, USA). All these reactive samples were retested by Western blotting (HTLV Blot, version 2.4; Genelabs Diagnostics, Singapore). The HTLV-1-positive samples were obtained from 23 patients diagnosed with HAM/TSP (16 female and 7 males), 62 asymptomatic (47 female and 15 males) and 9 individuals (6 female and 3 male) without information available on their clinical state. We also included 94 HTLV-negative sera obtained from the Hematology and Hemotherapy Foundation of the State of Bahia (HEMOBA Foundation), which returned negative in tests for Chagas disease, hepatitis B virus, hepatitis C virus, HIV-1/2, HTLV-1/2 and syphilis infections. In addition to these samples, 15 samples from HIV-positive patients, as previously defined by their serological diagnoses (reference standard tests), were kindly provided by the National Institute of Health Quality Control (INCQS - FIOCRUZ/RJ) and incorporated into the present sera sample set to evaluate cross-reactivity.

ELISA

In-house monoplex ELISA were performed in 96-well plates (Nunc MaxiSorp, EUA) with deposition of 50 ng of antigen in each well in a 100 µl volume in a carbonate-bicarbonate buffer (50 mM pH 9.6) and incubated for 16 h at 4°C. For multiplex ELISA with combination of antigens, 50 ng of each antigen was used in each well. Four combinations were tested aiming to cover all immunogenic regions of the Env and gag proteins as follows: (1) Ag 2-2 + Ag 2-16, (2) Ag 2-5 + Ag 2-7 + Ag 2-16, (3) Ag 2-5 + Ag 2-14 + 2-16 and (4) Ag 2-5 + Ag 2-15 + Ag 2-16).

Journal of Clinical Microbiology

After antigen deposition, unbound antigens were washed five times with 300 µl PBST (10 mM sodium phosphate pH 7.2, 150 mM NaCl, Tween-20 0.05% v/v) using a HydroFlex plate washer (TECAN, Switzerland) and the plates blocked with PBST containing 5% low-fat milk for 1 h at room temperature. The wells were washed again with PBST and incubated with 100 µl of the plasma samples diluted 1:200 in PBST at 37°C for 1 h. The wells were washed as described above and incubated with 100 µl of HRP-conjugated goat anti-human IgG (Biomanguinhos, FIOCRUZ/RJ, Brazil) diluted 1:20,000 in PBS at 37°C for 1 h with mild agitation. After five new washes, the assays were revealed with 100 µl TBM plus (tetramethyl-benzidine; Ken-En-Tec Diagnostics A/S, Taastrup, Denmark) at room temperature for 10 min in the dark. The reactions were stopped with 50 µl of 5 N H_2SO_4 and the absorbance at 450 nm was determined on a Synergy H1 hybrid multi-mode microplate reader (Biotek, Winooski, VT, USA).

Data analysis.

Data were analyzed using a scatter plot graphing software (GraphPad Prism version 8, San Diego, CA, USA). Continuous variables were determined as geometric mean ± standard deviation (SD). Data set normality was determined using the Shapiro-Wilk test, followed by the Student's t-test, and when homogeneity assumption was not confirmed, the Wilcoxon signed-ranks test was applied. A 5% level of significance was adopted for all statistical testing (p-value < 0.05). Cut-off point analysis was used to identify the optimal value of optical density (OD) to differentiate between negative and positive samples. The threshold value was established by area under the ROC curve (AUC) calculation. AUC values were also used to evaluate the global accuracy for each antigen, which can be classified as outstanding (1.0), elevated (0.82–0.99), moderate (0.62–0.81) or low (0.51–0.61) (60). All results were expressed by plotting the values as a reactivity index (RI). RI values were

Journal of Clinical Microbiology

calculated as sample's OD divided by CO and the results interpreted as follows: negative (RI < 1.0), positive (RI ≥ 1.0) and grey zone (0.90 ≤ RI ≤ 1.10). The HTLV-1 and HTLV-2 recombinant antigen performance parameters were determined using a dichotomous approach and compared regarding sensitivity (Se), specificity (Sp), and accuracy (Ac). Confidence intervals (CI) were employed with a confidence level of 95%. Imprecision assessments were based on Cohen's Kappa coefficient (κ) (Landis and Koch, 1977), which was interpreted as follows: perfect (κ = 1.0), almost perfect (1.0 < κ > 0.80), substantial (0.80 ≤ κ > 0.60), moderate (0.60 ≤ κ > 0.40), fair (0.40 ≤ κ > 0.20), slight (0.20 ≤ κ > 0) or poor (κ = 0). A flowchart (Appendix A: Fig. S1) and a checklist (Appendix A: Table S1) have been provided according to the Standards for Reporting of Diagnostic Accuracy Studies (STARD) guidelines (61).

Sequence identification:

GenBank accession numbers of the p19, p24 and envelope protein sequences used for definition of cloning sequences and/or multiple sequence alignment: HTLV-1/p19, P03362.3; HTLV-1/p24, X91888.1; HTLV-1/Env, NP_057865-1. HTLV-2/p19, PDB code 1JVR; HTLV-2/p24, NP_954567.1; HTLV-2/Env, CAA61545-1. HTLV-3 strain Lobak18, EU649782.1; HTLV-3 strain Cam2013OK, GQ463602.1; HTLV-4 isolate GabL14, KU863535.1. GenBank accession numbers of HTLV-1 subtype sequences: a, L36905.1; b, JX507077; c, KF242505 and g, KU214243.

Results

Definition of the genetic constructs containing the immunogenic regions of HTLV1/2 structural proteins

Definition of the constructions followed the necessity to have antigenic regions relevant for serodiagnosis and genetic constructs compatible with expression in *E. coli*. As shown in Table 1, for the HTLV-1 Env protein two immunogenic segments were identified in gp46, ranging approximately from residues 175 to 201 and 240 to 262, and one in gp21, from residues 360 to 405. For HTLV-2, the immunogenic regions of the Env protein are located between residues 172-208 in gp46 and 370-400 in gp21. Immunogenic regions have been identified also in the C-terminal of gag p19 (residues 100-130) for both HTLV-1 and HTLV-2. A summary of the genetic constructs with the limits of the amino acid region in each clone is shown in Table 2 and Figures 1B and 1C, whereas Figure 1A shows a representation of an HTLV virus genome indicating the respective position of the structural proteins.

For expression HTLV-2 p19, clone 3-4 was designed based on the construct described by Christensen et al. (42) including residues 2-137 with an additional 10-residue histidine tag in the N-terminal region. It yielded high levels of expression of soluble p19 as expected. Surprisingly, expression clone 2-1 encoding HTLV-1 p19 in *E. coli* did not yield any detectable protein. Construct 2-4 for expression of HTLV-1 gag p24 was designed based on its NMR structure (62). Construct 3-3 of HTLV-2 gag p24 was based on its structural similarity with HTLV-1 p24. Both contain a histidine tag in the N-terminal region and were expressed at high levels as soluble proteins. Construct 2-16 was designed to overcome the limitation of insoluble expression of HTLV-1 p19. This construct contains the immunogenic region comprising residues 103 to 130 of HTLV-1 p19 fused to the C-terminal region of HTLV-1 p24 and was expressed as a soluble protein.

For the Env protein, an initial set of constructions was made considering the immunogenic regions. These constructs (2-2 and 3-1) correspond to the gp46-gp21

Journal of Clinical Microbiology

284 segments of HTLV-1 and HTLV-2, respectively, and produced only insoluble proteins
285 extremely hard to purify. To improve the soluble expression and facilitate purification, a new
286 set of constructs was designed using p24 as a scaffold for the immunogenic regions of
287 gp46 and gp41. Clones 2-5, 2-11, 2-12 and 2-13 contain different segments of HTLV-1
288 gp21 fused to the C-terminal of HTLV-1 p24 (Table 2). The limits of gp21 segments in
289 clones 2-12 and 2-13 were determined taking into account the structural information
290 described for the gp21 ectodomain (residues 335-445) crystallized in fusion with the
291 maltose-binding protein (63, 64). Both 2-12 (gp21 residues 338-421) and 2-13 (gp21
292 residues 404-456) recombinant antigens expressed at high levels as soluble proteins and
293 cover most of gp21 protein and the immunogenic regions found in clone 2-2. Clones 2-7, 2-
294 14 and 2-15 contain different segments of HTLV-1 gp46 fused to HTLV-1 p24 (Table 2).
295 The immunogenic regions comprising residues 175-210 and 240-262 are found,
296 respectively, in clones 2-14 and 2-15. Both showed high expression but only clone 2-15
297 was soluble. Similarly, segments of HTLV-2 gp21 and gp46 were fused to p24 to facilitate
298 expression and purification (Table 2). Clones 3-5 and 3-6 contain different segments of
299 HTLV-2 gp21 whose limits were determined by sequence comparison with the crystal
300 structure HTLV-1 gp21 (64). Both were expressed as insoluble proteins. Clone 3-7 contains
301 HTLV-2 p24 fused to residues 162-208 comprising the immunogenic region of HTLV-2
302 gp46 (172 and 208).
303
304 **Purification of recombinant antigens**
305 A total of nine HTLV-1 and five HTLV-2 antigens were purified. Antigens 2-4, 2-12, 2-
306 13, 2-15 and 2-16 were expressed as soluble proteins and purified by immobilized metal
307 affinity chromatography (IMAC) (Fig. 1C). HTLV-1 antigens 2-2, 2-5, 2-7 and 2-14 were
308 isolated from inclusion bodies using buffer containing 4 M urea and purified by IMAC
309 (Fig.1D). Antigen 2-11 was also found in inclusion bodies and its purification was not

Journal of Clinical Microbiology

efficient. Since it contains p24, which corresponds to antigen 2-4 and, its p21 (360-421) segment is also found in clone 2-12, we decided not to proceed with purification of this antigen. Antigen 2-1 did not show any expression therefore, its purification was not possible. HTLV-2 antigens 3-3 and 3-4 were purified by IMAC from the soluble fraction (Fig. 1D). HTLV-2 antigen 3-1 was purified by IMAC from inclusion bodies solubilized in buffer containing 4 M urea and antigens 3-5 and 3-7 were purified by IMAC after solubilization of inclusion bodies in 8 M urea (Fig. 1D). HTLV-2 antigen 3-6 purification was not possible. It differs from antigen 3-5 (p24 + gp21 334-436) by having a shorter gp21 segment (gp21 366–436). From this result, we can assume that deletion of gp21 residues 334 to 365 affects the expression of this truncated protein in *E. coli*. The extra high molecular mass bands seen for two HTLV-1 (Ag 2-5 and 2-7) were analyzed by mass spectrometry. This analysis confirmed that these bands correspond to the respective 2-5 and 2-7 antigens (data not shown). Their slower mobility in SDS-PAGE is most probably due to formation of cysteine dimers, which are not completely reduced during sample preparation for SDS-PAGE.

Evaluation of HTLV-1 antigen performance

Nine recombinant HTLV-1 antigens were initially tested individually using ELISA. ROC curves were generated for a panel of 94 samples from individuals diagnosed as HTLV1-infected patients in parallel with 94 negative control plasma samples (Table S2). The area under the ROC curves (AUC) values was extremely high for all antigens (Fig. 2), ranging from 98.4 (Ag 2-15) to 99.9 (Ag 2-14). Considering a 95% confidence interval, all antigens showed similar performance parameters. These results demonstrate excellent diagnostic accuracy for all HTLLV-1 antigens.

HTLV-1 antigens were also assayed in a multiplex format to evaluate if a combination of a larger number of immunogenic regions could improve the overall

performance of the assay. Four combinations were tested as follows: Ag 2-2 + Ag 2-16; Ag 2-5 + 2-7 + 2-16; Ag 2-5 + Ag 2-14 + Ag 2-16, and Ag 2-5 + Ag 2-15 + Ag 2-16. Similar with the results obtained for individual antigens, the performance parameters for all combinations were extremely high (Fig. 3), confirming the excellent discrimination power and high diagnostic values of the antigens.

The efficiency of the antigens can be evaluated also by the number of samples that fall in the grey zone. Considering a grey zone established as the cut off value +/- 10% (RI values of 1.0 ± 0.10), only 4.3% of the samples (3 HTLV-1-negative + 5 HTLV-1-positive samples) were found in the grey zone for Ag 2-14. Similarly, only 3.7% of the samples (2 HTLV-1-negative + 5 HTLV-1-positive samples) fell in the grey zone for antigen combination 2-5, 2-14 and 2-16. This furthers shows that these antigens can properly discriminate between HTLV-1 positive and negative samples.

Evaluation of HTLV-2 antigen reactivity with HTLV-1-positive plasma samples

Five recombinant HTLV-2 antigens were tested in ELISA against the same panel of plasma samples including 94 HTLV1-positive and 94 negative controls. The AUC values were extremely high and similar for all antigens (Fig. 4), ranging from 94.9 (Ag 3-4) to 99.7 (Ag 3-5). Ag 3-5 shows higher parameters relative to Ag 3-3, Ag 3-4 and Ag 3-7 considering a 95% confidence interval. These performances can be explained by the high degree of amino acid sequence similarity between HTLV-1 and HTLV-2 structural proteins and will be addressed in the discussion.

Evaluation of cross-reaction of HIV-1-positive sera with HTLV-1/2 antigens

The HTLV-1 and HTLV-2 recombinant antigens produced in this work were assayed with a panel of 15 HIV-1-positive sera provided by the National Institute of Health Quality

Journal of Clinical Microbiology

Control (INCQS - FIOCRUZ/RJ). The HIV-1-positive samples presented reactivity similar with HTLV-1-negative samples (Fig. 5) except for one sample testing positive for Ag 2-4 and another for Ag 2-12. Interestingly, antigen 2-14, which showed the lowest number of samples in the gray zone with the HTLV-1-positive samples did not show any cross-reaction with HIV-1-positive samples, indicating a high selectivity towards HIV-1 antibodies.

Among the HTLV-2 recombinant antigens, only Ag 3-5 did not show cross-reaction with HIV-1 samples. Ag 3-1 showed cross-reaction with one sample and Ag 3-3, Ag 3-4 and Ag 3-7 showed cross-reaction with HIV-1 samples, being positive for 6, 3 and 4 sera, respectively, out of a total of 15 sera.

Discussion

HTLV-1 is a particularly dangerous virus since 5-10% of infected people develop HTLV-1-related life-threatening diseases (1) for which there is no efficient treatment. Therefore, reducing the medical burden caused by HTLV-related disease has relied on prevention of transmission. Prevention of mother-to-child transmission and screening of contaminated blood products for transfusions requires large scale screens that can be easier performed by using serological methods. Serological methods depend on antigens to capture antibodies raised against the virus and the type and quality of the antigens is critical for efficiency of the screen assays.

Production of viral antigens using tissue culture is usually very expensive and time consuming. Therefore, bacterial systems can be an alternative for cost-effective production of recombinant antigen of pathogenic organisms (65). In the case of HTLV-1/2, the immunogenic regions have been mapped mainly to three segments of the viral envelope protein (gp46 regions 175-201 and 240-262 and gp21 360-405) and the matrix (p19) and capsid (p24) proteins. In the case of p24 from HTLV-1 and p19 from HTLV-2, we took

advantage of previous studies showing high yields of these proteins using *E. coli* expression systems (42, 62). The genetic constructs that we designed also resulted in high levels of soluble expression for these two proteins. However, a similar HTLV-1 p19 construct did not produce any protein while HTLV-2 p24, which shares 87% amino acid identity in the clones used in this work, was produced at high levels as soluble protein. Structural predictions indicate global conservation of secondary and three-dimensional structure of HTLV-1 p19. However, the overall amino acid identity of HTLV-1 and HTLV-2 p19 is only 59%. This lower overall conservation when compared to p24 may account for HTLV-1 p19 ineffective expression in *E. coli*.

The constructs containing only the immunogenic segments of gp46-gp21 Env protein are highly insoluble when expressed in *E. coli*. Their purification under denaturing conditions is possible but not efficient. As an alternative, we have used constructions where the immunogenic regions of gp46 and gp21 were fused to the C-terminus of p24 or the p24 C-terminal domain. In the case of HTLV-1, two fusions with full-length p24 (Ag 2-15 and 2-16) and two fusions with the C-terminal domain (Ag 2-12 and 2-13) were expressed as soluble proteins. Antigens 2-5, 2-7, 2-14 were expressed in inclusion bodies but their purification was straightforward and resulted in higher yields than the segments of the Env protein alone (Ag 2-2 and 3-1). Similar results were obtained for the HTLV-2 p24 fusions with segments of the gp46 and gp21.

The ELISA results show that all HTLV-1 antigens present excellent efficiency to discriminate HTLV-1-postive from negative samples. By considering a 95% confidence interval there is no significant difference between the antigens. However, antigens 2-5 and 2-14, containing full-length p24 fused to either gp21 338-421 or gp46 162-222 residues, respectively, were the ones showing the highest specificity, further confirming that these gp21 and gp46 segments are critical for HTLV-1 detection in immunoassays. Accordingly,

Journal of Clinical Microbiology

lowest kappa scores are observed for Ag 2-4 and Ag 2-16, corresponding to p24 alone and p24 fused to the antigenic region of p19, respectively. This indicates that the gag proteins alone present a somewhat poorer discrimination efficiency. Individually, antigen 2-14 also shows the highest sensibility and the lowest number of samples in the gray zone, which are features to be considered to select a smaller set of antigens for further studies. ELISA using antigen combinations produced results similar with the ones obtained for the individual antigens. The mix containing 2-5, 2-14 and Ag 2-16, which comprises p24 and the immunogenic regions of gp21, gp46 an p19, respectively, also showed a reduced number of samples in the gray zone. Since this was seen for Ag 2-14 alone, we can assume the Ag 2-14 contributes to improve discrimination between positive and negative samples.

An extremely high level of reactivity was observed when the HTLV-1-positive plasma samples were tested with HTLV-2 antigens. Although this might be striking at first glance, the results can be explained by the high overall sequence conservation of p24 and the immunogenic regions of p19 and gp21 (Figure 6). p24 from HTLV-1 and HTLV-2 share 87% amino acid identity over the segment used in the clones of this study. p19 is more divergent showing overall sequence identity of 59% but the immunogenic region in the C-terminal is conserved. Similarly, the immunogenic region in the gp21 segment of the Env protein is totally conserved between the two viruses. This conservation explains the high sensibility and specificity of Ag 3-5 for HTLV-1-positive samples. Therefore, any attempt to discriminate between the two viruses using antibody capture methods should avoid the use of p24 and the conserved epitopes from pg21 and p19. The immunogenic regions of the gp46 segment of the Env protein are more divergent, which can explain the differences observed for Ag 3-5 and Ag 3-7 (Figure 6).

Antigen conservation and diversity must be considered also for the HTLV-3 and HTLV-4 types which, at least in Central Africa, show overlapping geographic distribution

Journal of Clinical Microbiology

with the other HTLV types. Sequence alignment shows that the four types present similar ratio of sequence conservation (Figure 6). p24 is highly conserved among the four types, with amino acid identity ranging from 85% to 92% (Figure 6). p19 and the envelope proteins show lower sequence conservation with amino acid identity ranging from 57% to 77% f and from 66% to 81%, respectively (Figure 6). Interestingly, the HTLV-3 strain Lobak18 (66) and the HTLV-3 isolate Cam2013OK (67) show highly divergent p19 and Env sequences. Therefore, they were kept separate in this analysis (Figure 6). p19 presents more sequence diversity in the C-terminal region while the envelope protein shows more diversity in the N-terminal region and in the gp46 segment. However, gp46 also contains regions conserved in the four types which comprise amino acids 220-240, 293-304 and 311-332. The gp21 segment of the envelope protein is highly conserved in all four subtypes (Figure 6). In HTLV-1 and HTLV-2, this segment was shown to contain immunodominant epitopes. The combined conservation and immunogenicity of p24 and the conserved segments in gp46 and gp21 of all four HTLV types, indicates that most probably antibodies raised against HTLV-3 and HTLV-4 can be detected in immunoassays using antigens from HTLV-1 and HTLV-2.

A frequent problem reported to HTLV diagnosis involves the high ratio of false positive results (68–70), which can be caused by different reasons. False-positive results in quantitative PCR assays for viral genome detection may be due to viral load, defective provirus and genome sequence variation (69, 71). In the case of serological assays, antibody levels depend on infection stage and seroconversion as well as the nature of the antigens combined with the type of assay used (69–71). This type of problem can be minimized and overcome by using antigens with high discrimination capacity. In our assays, the average signal of the negative samples were relatively high and most recombinant antigens showed a high number of samples in the grey zone, which corresponds to the

Journal of Clinical Microbiology

reactivity index +/- 10%. However, we should point out that two assays showed a very low number of samples in the grey zone. This was the case of antigen 2-14, showing only 4.3% of the samples in the grey zone and of the combination antigens 2-5, 2-14 and 2-16 with all immunogenic regions of p19, p24, gp46 and gp21 (Ag), showing 3.7% of samples in the grey zone.

Another concern involving HTLV-1 serodiagnosis involves possible antigen cross reaction with HIV antibodies. Only two recombinant HTLV-1 antigens produced in this work showed cross reaction, each with a single sample of HIV-1-infected patients. The other seven antigens discriminated properly the HIV-1-positive samples and three of them (Ag 2-7, Ag 2-13 and 2-14) showed no sample in the gray zone, indicating low potential for these antigens to cross react with HIV-1 antibodies. On the other hand, among the HTLV-2 antigens only Ag 3-5 did not show any cross reaction with HIV-1 samples. Antigens 3-3 (p24), 3-4 (p19) and 3-7 (p24 + gp46 162-208) showed unexpectedly high number of sera with reactivity index above 1, indicating cross reaction with HIV-1 antibodies. This cannot be explained on basis of sequence similarity only. Sequence analysis using PSI-BLAST do not show any amino acid identity between p19 of HTLV-1 and HTLV -2 and the HIV-1 matrix protein. Similarly, no relevant amino acid identity was detected between the Env proteins of HTLV-1 and HIV-1. The Env protein of HIV-1 closest to HTLV-1 shows 31% of identify over a stretch of only 70 residues. The envelope protein of the known HTLV-1 subtypes is highly conserved, sharing 95% to 99% amino acid identity and showing similar low sequence identity to HIV-1. HIV-1 p24 shows short regions of similarity with low degree of conservation, with the closest HIV-1 p24 sequence reaching only 33% identity over a stretch of 84 residues (28/84) with HTLV-1 and 38% over a segment of 69 residues (26/69) with HTLV-2. Despite low sequence conservation, one cannot rule out the existence of

Journal of Clinical Microbiology

structural epitopes, which can account for some of the cross reaction of HIV-1 antibodies with natively structured parts of the antigens.

In summary, we describe the production and serological evaluation of a large set of recombinant antigens containing the immunogenic regions of HTLV-1 and HTLV-2 structural proteins. The paper presents a detailed compilation of the immunogenic regions of these proteins and how they were combined to generate the genetic constructs. It presents an extensive description of the expression and purification procedures to isolate the recombinant antigens for serological assays. The results obtained in the assays using HTLV-1-positive sera revealed a high potential for utilization of the HTLV-1 recombinant antigens for HTLV-1 diagnosis. Our data suggest that the HTLV-1 recombinant proteins are eligible to enter phase II studies.

Acknowledgements

The authors acknowledge the FIOCRUZ Technical Platform program through the Platforms for Protein Purification and Characterization and Proteomics at the Carlos Chagas Institute. This study benefited from FINEP (grant agreement 01.11.0286.00) and BNDES (grant agreement 11.2.1328.1) funding. Nilson I. T. Zanchin is a CNPq research fellow (grants nr. 312195/2015-0 and 304167/2019-3).

References

1. Tagaya Y, Gallo RC. 2017. The exceptional oncogenicity of HTLV-1. Front Microbiol 8.
2. Gessain A, Francis H, Sonan T, Sonan T, Giordano C, Akani F, Piquemal M, Caudie C, Malone G, Essex M, De-Thé G. 1986. HTLV-I and tropical spastic paraparesis in Africa. Lancet 328:698.
3. Osame M, Usuku K, Izumo S, Ijichi N, Amitani H, Igata A, Matsumoto M, Tara M. 1986. HTLV-I associated myelopathy, a new clinical entity. Lancet 327:1031–1032.
4. Kalyanaraman VS, Sarngadharan MG, Robert-Guroff M, Miyoshi I, Blayney D, Golde D, Gallo RC. 1982. A new subtype of human T-cell leukemia virus (HTLV-II) associated with a T-cell variant of hairy cell leukemia. Science (80-) 218:571–573.
5. Hjelle B, Torrez-Martinez N, Mills R, Appenzeller O, Jahnke R, Alexander S, Ross G, Hjelle B. 1992. Chronic neurodegenerative disease associated with HTLV-II infection. Lancet 339:645–646.
6. Jacobson S, Lehky T, Nishimura M, Robinson S, McFarlin DE, Dhib-Jalbut S. 1993. Isolation of HTLV-II from a patient with chronic, progressive neurological disease clinically indistinguishable from HTLV-I-associated myelopathy/tropical spastic paraparesis. Ann Neurol 33:392–396.
7. Gessain A, Cassar O. 2012. Epidemiological aspects and world distribution of HTLV-1 infection. Front Microbiol 3:388.
8. European Centre for Disease Prevention and Control. 2015. Geographical distribution of areas with a high prevalence of HTLV-1 infection - Publications Office of the EU.
9. González-Alcaide G, Ramos JM, Huamaní C, De Mendoza C, Soriano V. 2016. Human T-lymphotropic virus 1 (HTLV-1) and human T-lymphotropic virus 2 (HTLV-2): Geographical research trends and collaboration networks (1989-2012). Rev Inst Med

Journal of Clinical Microbiology

Trop Sao Paulo 58.

10. Willems L, Hasegawa H, Accolla R, Bangham C, Bazarbachi A, Bertazzoni U, Carneiro-Proietti AB de F, Cheng H, Chieco-Bianchi L, Ciminale V, Coelho-dos-Reis J, Esparza J, Gallo RC, Gessain A, Gotuzzo E, Hall W, Harford J, Hermine O, Jacobson S, Macchi B, Macpherson C, Mahieux R, Matsuoka M, Murphy E, Peloponese JM, Simon V, Tagaya Y, Taylor GP, Watanabe T, Yamano Y. 2017. Reducing the global burden of HTLV-1 infection: An agenda for research and action. Antiviral Res 137:41–48.

11. Martinez MP, Al-Saleem J, Green PL. 2019. Comparative virology of HTLV-1 and HTLV-2. Retrovirology 16:21.

12. Dourado I, Alcantara LCJ, Barreto ML, Teixeira MDG, Galvão-Castro B. 2003. HTLV-I in the General Population of Salvador, Brazil: A City with African Ethnic and Sociodemographic Characteristics. J Acquir Immune Defic Syndr 34:527–531.

13. Watanabe T. 2011. Current status of HTLV-1 infection. Int J Hematol 94:430–434.

14. Xie J, Ge S, Zhang Y, Lin Y, Ni H, Zhang J, Chen C. 2015. The Prevalence of Human T-Lymphotropic Virus Infection among Blood Donors in Southeast China, 2004-2013. PLoS Negl Trop Dis 9:e0003685.

15. Satake M, Iwanaga M, Sagara Y, Watanabe T, Okuma K, Hamaguchi I. 2016. Incidence of human T-lymphotropic virus 1 infection in adolescent and adult blood donors in Japan: a nationwide retrospective cohort analysis. Lancet Infect Dis 16:1246–1254.

16. Nunes D, Boa-Sorte N, Grassi MFR, Taylor GP, Teixeira MG, Barreto ML, Dourado I, Galvão-Castro B. 2017. HTLV-1 is predominantly sexually transmitted in Salvador, the city with the highest HTLV-1 prevalence in Brazil. PLoS One 12:e0171303.

17. Rosadas C, Malik B, Taylor GP, Puccioni-Sohler M. 2018. Estimation of HTLV-1

556 vertical transmission cases in Brazil per annum. PLoS Negl Trop Dis 12:e0006913.
557 18. Chen X, Liu F, Fu X, Feng Y, Zhang D, Liu H, Xi G, Hu J. 2019. Prevalence of human
558 T-cell lymphotropic virus type-1 infection among blood donors in mainland China: a
559 systematic review and meta-analysis of the last 20 years. Expert Rev Hematol
560 12:579–587.
561 19. Murphy EL, Cassar O, Gessain A. 2015. Estimating the number of HTLV-2 infected
562 persons in the world. Retrovirology 12:1–1.
563 20. Paiva A, Casseb J. 2015. Origin and prevalence of human T-lymphotropic virus type
564 1 (HTLV-1) and type 2 (HTLV-2) among indigenous populations in the Americas. Rev
565 Inst Med Trop Sao Paulo 57:1–13.
566 21. Mata EC, Bezerra RM, Proietti Júnior AA, Pamplona LK, Gomes LO, Corrêa VC,
567 Caluff JS, Borges GS, Casseb J, Kanzaki L. 2018. HTLV-1/2 prevalence in two
568 Amazonian communities. J virus Erad 4:174–178.
569 22. Calattini S, Chevalier SA, Duprez R, Bassot S, Froment A, Mahieux R, Gessain A.
570 2005. Discovery of a new human T-cell lymphotropic virus (HTLV-3) in Central Africa.
571 Retrovirology 2:30.
572 23. Wolfe ND, Heneine W, Carr JK, Garcia AD, Shanmugam V, Tamoufe U, Torimiro JN,
573 Prosser AT, LeBreton M, Mpoudi-Ngole E, McCutchan FE, Birx DL, Folks TM, Burke
574 DS, Switzer WM. 2005. Emergence of unique primate T-lymphotropic viruses among
575 central African bushmeat hunters. Proc Natl Acad Sci U S A 102:7994–7999.
576 24. Richard L, Mouinga-Ondémé A, Betsem E, Filippone C, Nerrienet E, Kazanji M,
577 Gessain A. 2016. Zoonotic Transmission of Two New Strains of Human T-
578 lymphotropic Virus Type 4 in Hunters Bitten by a Gorilla in Central Africa, p. 800–803.
579 *In* Clinical Infectious Diseases. Oxford University Press.
580 25. Gessain A, Rua R, Betsem E, Turpin J, Mahieux R. 2013. HTLV-3/4 and simian

foamy retroviruses in humans: Discovery, epidemiology, cross-species transmission and molecular virology. Virology 435:187–199.

26. Martin F, Tagaya Y, Gallo R. 2018. Time to eradicate HTLV-1: an open letter to WHO. Lancet 391:1893–1894.

27. Berini CA, Susana Pascuccio M, Bautista CT, Gendler SA, Eirin ME, Rodriguez C, Pando MA, Biglione MM. 2008. Comparison of four commercial screening assays for the diagnosis of human T-cell Lymphotropic virus types 1 and 2. J Virol Methods 147:322–327.

28. Qiu X, Hodges S, Lukaszewska T, Hino S, Arai H, Yamaguchi J, Swanson P, Schochetman G, Devare SG. 2008. Evaluation of a new, fully automated immunoassay for detection of HTLV-I and HTLV-II antibodies. J Med Virol 80:484–493.

29. Kapprell H-P, Stieler M, Oer M, Goller A, Hausmann M, Schochetman G, Devare SG, Qiu X. 2010. Evaluation of a new third-generation ARCHITECT rHTLV-I/II assay for blood screening and diagnosis. Diagn Microbiol Infect Dis 67:61–9.

30. Malm K, Kjerstadius T, Andersson S. 2010. Evaluation of a new screening assay for HTLV-1 and -2 antibodies for large-scale use. J Med Virol 82:1606–1611.

31. Malm K, Kragsbjerg E, Andersson S. 2015. Performance of Liaison XL automated immunoassay platform for blood-borne infection screening on hepatitis B, hepatitis C, HIV 1/2, HTLV 1/2 and Treponema pallidum serological markers. Transfus Med 25:101–105.

32. Laperche S, Sauleda S, Piron M, Mühlbacher A, Schennach H, Schottstedt V, Queirós L, Uno N, Yanagihar K, Imdahl R, Hey A, Klinkicht M, Melchior W, Muench P, Watanabei T. 2017. Evaluation of sensitivity and specificity performance of elecsys HTLV-I/II assay in a multicenter study in Europe and Japan. J Clin Microbiol 55:2180–

2187.

33. da Silva Brito V, Santos FLN, Gonçalves NLS, Araujo THA, Nascimento DSV, Pereira FM, Boa-Sorte NCA, Grassi MFR, Caterino-de-Araujo A, Galvão-Castro B. 2018. Performance of commercially available serological screening tests for human T-cell lymphotropic virus infection in Brazil. J Clin Microbiol 56:e00961-18.

34. Yun SG, Kim S wook, Sohn JY, Cho Y. 2019. Evaluation of Elecsys HTLV-I/II assay in comparison with ARCHITECT rHTLV-I/II assay with Korean samples. J Clin Lab Anal 33:e22909.

35. Samuel KP, Lautenberger JA, Jorcyk CL, Josephs S, Wong-Staal F, Papas TS. 1984. Diagnostic potential for human malignancies of bacterially produced HTLV-I envelope protein. Science (80-) 226:1094–1097.

36. Chen YM, Lee TH, Samuel KP, Okayama A, Tachibana N, Miyoshi I, Papas TS, Essex M. 1989. Antibody reactivity to different regions of human T-cell leukemia virus type 1 gp61 in infected people. J Virol 63:4952–4957.

37. Chen YMA, Essex M. 1991. Identification of a Recombinant HTLV-II Envelope Protein for Serological Detection of HTLV-II Carriers. AIDS Res Hum Retroviruses 7:453–457.

38. Lipka JJ, Miyoshi I, Hadlock KG, Reyes GR, Chow TP, Blattner WA, Shaw GM, Hanson CV, Gallo D, Chan L FS. 1992. Segregation of human T cell lymphotropic virus type I and II infections by antibody reactivity to unique viral epitopes - PubMed. J Infect Dis 165:268–272.

39. Varma M, Rudolph DL, Knuchel M, Switzer WM, Hadlock KG, Velligan M, Chan L, Foung SKH, Lal RB. 1995. Enhanced specificity of truncated transmembrane protein for serologic confirmation of human T-Cell lymphotropic virus type 1 (HTLV-1) and HTLV-2 infections by Western blot (immunoblot) assay containing recombinant

envelope glycoproteins. J Clin Microbiol 33:3239–3244.

40. Sokol L, Agrawal D, Loughran TP. 2005. Characterization of HTLV envelope seroreactivity in large granular lymphocyte leukemia. Leuk Res 29:381–387.

41. Kuga T, Yamasaki M, Sekine S, Fukui M, Yokoo Y, Itoh S, Yoshida M, Hattori T, Takatsuki K. 1988. A gag-env Hybrid Protein of Human T-cell Leukemia Virus Type I and Its Application to Serum Diagnosis. Japanese J Cancer Res 79:1168–1173.

42. Christensen AM, Massiah MA, Turner BG, Sundquist WI, Summers MF. 1996. Three-dimensional structure of the HTLV-II matrix protein and comparative analysis of matrix proteins from the different classes of pathogenic human retroviruses. J Mol Biol 264:1117–1131.

43. Horal P, Hall WW, Svennerholm BO, Lycke J, Jeansson S, Rymo L, Kaplan MH, Vahlne A. 1991. Identification of type-specific linear epitopes in the glycoproteins gp46 and gp21 of human T-cell leukemia viruses type I and type II using synthetic peptides. Proc Natl Acad Sci U S A 88:5754–5758.

44. Kuroki M, Nakamura M, Itoyama Y, Tanaka Y, Shiraki H, Baba E, Esaki T, Tatsumoto T, Nagafuchi S NS. 1992. Identification of new epitopes recognized by human monoclonal antibodies with neutralizing and antibody-dependent cellular cytotoxicity activities specific for human T cell leukemia virus type 1 - PubMed. J Immunol 149:940–948.

45. Yamada K, Kuroda N, Washitani Y, Shiraki H MY. 1992. Enzyme immunoassay with synthetic peptides to detect anti-HTLV-I antibodies - PubMed. Clin Chem 38:699–703.

46. Londos-Gagliardi D, Dalibart R, Geoffre S, Dalbon P, Pouliquen JF, Georges-Courbot MC, Sainte-Foie S, Hajjar C, Georges AJ, Moreau JP, Guillemain B. 1996. Immunogenicity of variable regions of the surface envelope glycoprotein of HTLV type

656 I and identification of new major epitopes in the 239-261 region. AIDS Res Hum
657 Retroviruses 12:941–950.
658 47. Londos-Gagliardi D, Jauvin V, Armengaut MH, Astier-Gin T, Goetz M, Huet S,
659 Guillemain BJ. 1999. Influence of amino acid substitutions on antigenicity of
660 immunodominant regions of the HTLV type I envelope surface gylcoprotein: A study
661 using monoclonal antibodies raised against relevant peptides. AIDS Res Hum
662 Retroviruses 15:909–920.
663 48. Hernández M, Selles ME, Pozo Pea L, Gómez I, Melchor A. 2000. Antigenicity of
664 chimeric synthetic peptides based on HTLV-1 antigens and the impact of epitope
665 orientation. Biochem Biophys Res Commun 276:1085–1088.
666 49. Palker TJ, Tanner ME, Scearce RM, Streilein RD, Clark ME HB. 1989. Mapping of
667 immunogenic regions of human T cell leukemia virus type I (HTLV-I) gp46 and gp21
668 envelope glycoproteins with env-encoded synthetic peptides and a monoclonal
669 antibody to gp46. | The Journal of Immunology. J Immunol 142:971–978.
670 50. Lal RB, Rudolph DL, Lairmore MD, Khabbaz RF, Garfield M, Coligan JE FT. 1991.
671 Serologic Discrimination of Human T Cell Lymphotropic Virus Infection by Using a
672 Synthetic Peptide-Based Enzyme Immunoassay | The Journal of Infectious Diseases
673 | Oxford Academic. J Infect Dis 163:41–46.
674 51. Rudolph DL LR. 1993. Discrimination of human T-lymphotropic virus type-I and type-
675 II infections by synthetic peptides representing structural epitopes from the envelope
676 glycoproteins - PubMed. Clin Chem 39:288–292.
677 52. Hadlock KG, Goh CJ, Bradshaw PA, Perkins S, Lo J, Kaplan JE, Khabbaz R, Foung
678 SKH. 1995. Delineation of an immunodominant and human T-cell lymphotropic virus
679 (HTLV)-specific epitope within the HTLV-I transmembrane glycoprotein. Blood
680 86:1392–1399.

53. Loughran TP, Hadlock KG, Perzova R, Gentile TC, Yang Q, Foung SKH, Poiesz BJ. 1998. Epitope mapping of HTLV envelope seroreactivity in LGL leukaemia. Br J Haematol 101:318–324.

54. Lal RB, Rudolph DL, Griffis KP, Kitamura K, Honda M, Coligan JE, Folks TM. 1991. Characterization of immunodominant epitopes of gag and pol gene-encoded proteins of human T-cell lymphotropic virus type I. J Virol 65:1870–1876.

55. Marin MH, Rodríguez-Tanty C, Higginson-Clarke D, Bocalandro YM, Peña LP. 2005. Study of the peptide length and amino acid specific substitution in the antigenic activity of the chimeric synthetic peptides, containing the p19 core and gp46 envelope proteins of the HTLV-I virus. Biochem Biophys Res Commun 336:983–986.

56. Lal RB, Rudolph DL, Kaplan JE, Hjelle B, Levine PH, Coligan JE, Viscidii RP. 1992. Identification of immunodominant epitopes in envelope glycoprotein of human T Lymphotropic virus type II. Virology 186:274–279.

57. Hernández Marin M, Castellanos Pentón P, Márquez Bocalandro Y, Pozo Pea L, Díaz Navarro J, Javier González López L. 2001. Chimeric synthetic peptides from the envelope (gp46) and the transmembrane (gp21) glycoproteins for the detection of antibodies to human T-cell leukemia virus type II. Biochem Biophys Res Commun 289:7–12.

58. Marin MH, Bocalandro YM, Vallejo RV, Tanty CR, Clark DH, Peña LP, León CS. 2003. Use of a chimeric synthetic peptide from the core p19 protein and the envelope gp46 glycoprotein in the immunodiagnosis of HTLV-II virus infection. Prep Biochem Biotechnol 33:29–38.

59. Tanaka Y, Tanaka R, Hoshino H. 1994. Identification of a novel neutralization epitope on envelope GP46 antigen of human t-cell-leukemia virus-type-II (HTLV-II). Int J Cancer 59:655–660.

706 60. Swets JA. 1988. Measuring the accuracy of diagnostic systems. Sci Sci 240:1285–
707 1293.
708 61. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine
709 D, Reitsma JB, De Vet HCW, Bossuyt PMM. 2016. STARD 2015 guidelines for
710 reporting diagnostic accuracy studies: Explanation and elaboration. BMJ Open
711 6:e012799.
712 62. Khorasanizadeh S, Campos-Olivas R, Clark CA, Summers MF. 1999. Letter to the
713 editor: Sequence-specific 1H, 13C and 15N chemical shift assignment and secondary
714 structure of the HTLV-I capsid protein [8]. J Biomol NMR 14:199–200.
715 63. Center RJ, Kobe B, Wilson KA, Teh T, Howlett GJ, Kemp BE, Poumbourios P. 1998.
716 Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain
717 fragment as a chimera with maltose-binding protein. Protein Sci 7:1612–1619.
718 64. Kobe B, Center RL, Kemp BE, Poumbourios P. 1999. Crystal structure of human T
719 cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein
720 chimera reveals structural evolution of retroviral transmembrane proteins. Proc Natl
721 Acad Sci U S A 96:4319–4324.
722 65. Rosano GL, Ceccarelli EA. 2014. Recombinant protein expression in Escherichia coli:
723 Advances and challenges. Front Microbiol 5:172.
724 66. Calattini S, Betsem E, Bassot S, Chevalier SA, Mahieux R, Froment A, Gessain A.
725 2009. New strain of human T lymphotropic virus (HTLV) type 3 in a pygmy from
726 Cameroon with peculiar HTLV serologic results. J Infect Dis 199:561–564.
727 67. Zheng HQ, Wolfe ND, Sintasath DM, Tamoufe U, LeBreton M, Djoko CF, Diffo JLD,
728 Pike BL, Heneine W, Switzer WM. 2010. Emergence of a novel and highly divergent
729 HTLV-3 in a primate hunter in Cameroon. Virology 401:137–145.
730 68. Kline RL, Quinn TC, Brothers T, Halsey N, Boulos R, Lairmore MD. 1991. Evaluation

731 of enzyme immunoassays for antibody to human T-lymphotropic viruses type I/II.
732 Lancet 337:30–33.
733 69. Campos KR, Gonçalves MG, Caterino-de-Araujo A. 2017. Short Communication:
734 Failures in Detecting HTLV-1 and HTLV-2 in Patients Infected with HIV-1. AIDS Res
735 Hum Retroviruses 33:382–385.
736 70. Tosswill JHC, Taylor GP. 2018. Interpretation of low reactivity in the Abbott Architect
737 rHTLV I/II assay. Transfus Med 28:326–330.
738 71. Campos KR, Caterino-de-Araujo A. 2020. Provirus Mutations of Human T-
739 Lymphotropic Virus 1 and 2 (HTLV-1 and HTLV-2) in HIV-1-Coinfected Individuals.
740 mSphere 5:e00923-20.
741
742

Figure Legends

Figure 1. Representation of the genetic constructs and SDS-PAGE of purified antigens. (A) Scheme of HTLV1/2 genome structure. (B and E) Representation of the constructs used to express HTLV-1 and HTLV-1 antigens, respectively. (C and D) Image of an SDS-PAGE showing purified HTLV-1 and HLTV-2 antigens, respectively.

Figure 2. Evaluation of recombinant HTLV-1 antigen performance using ELISA. (A) Graphs showing the reactivity index for each antigen tested against a panel of 94 HTLV-1-positive and 94 HTLV-1-negative samples. The cut-off value is 1.0 and shadowed the area represents the grey-zone. Horizontal lines and numbers for each group of results represent the geometric means (± 95%CI). (B) Antigen performance parameters obtained from the assays shown in A. AUC: area under the ROC curve; Sens.: sensibility; Spec.: specificity; Acc.: accuracy; kappa: Cohen's kappa index; CI: confidence interval.

Figure 3. Evaluation of HTLV-1 antigen combinations by ELISA. (A) Reactivity index for each antigen combination tested against a panel of 94 HTLV-1-positive and 94 HTLV-1-negative samples. The cut-off value is 1.0 and shadowed area represents the grey-zone. Horizontal lines and numbers for each group of results represent the geometric means (± 95%CI). (B) Performance parameters obtained from the assays shown in A. AUC: area under the ROC curve; Sens.: sensibility; Spec.: specificity; Acc.: accuracy; kappa: Cohen's kappa index; CI: confidence interval.

Figure 4. Evaluation of HTLV-2 antigen combinations by ELISA. (A) Reactivity index for each antigen combination tested against a panel of 94 HTLV-1-positive and 94 HTLV-1-

Journal of Clinical Microbiology

768 negative samples. The cut-off value is 1.0 and shadowed area represents the grey-zone.
769 Horizontal lines and numbers for each group of results represent the geometric means (±
770 95%CI). (B) Performance parameters obtained from the assays shown in A. AUC: area
771 under the ROC curve; Sens.: sensibility; Spec.: specificity; Acc.: accuracy; kappa: Cohen's
772 kappa index; CI: confidence interval.
773
774
775 **Figure 5.** Cross-reaction of recombinant HTLV-1/HTLV-2 antigens with HIV-1-positive
776 samples assayed by ELISA. Data are shown as the reactivity index for each antigen tested
777 against 15 HIV-1 positive samples. The cut-off value is 1.0 and shadowed area represents
778 the grey-zone. RI: reactivity index. CI 9%: confidence interval of 95%. CR (%): number and
779 percentage of cross-reacting samples. GZ (%): number and percentage of samples falling
780 in the gray zone.
781
782 **Figure 6.** Conservation and diversity of the immunogenic regions in structural proteins of
783 HTLV-1 to HTLV-4. (A) Multiple sequence alignment of p19, p24 and envelope protein.
784 Symbols ***** and **:** or **.** underneath the alignments indicate conserved residues in all
785 sequences and conserved amino acid substitutions, respectively. p24 is highly conserved in
786 all HTLV types. The immunogenic regions previously identified for p19 and Env of HTLV-1
787 and HTLV-2 are marked with colored boxes according to the references cited in Table 1.
788 The regions corresponding to the gp46 and gp21 segments present in our constructs are
789 shown in bold. The relatively divergent immunogenic regions of the gp46 segment are
790 indicated by turquois boxes. The green boxes indicate the conserved immunogenic regions
791 of the gp21 segment. (B) Summary of protein conservation between the HTLV types.
792 HTLV-3 strain Lobak18 and HTLV-3 isolate Cam2013OK shown highly divergent p19 and

793 Env sequences. Therefore, they were analyzed separately. p24 is more conserved, sharing
794 from 85% to 95% amino acid identity. p19 and the envelope protein show lower amino acid
795 identity, in the range of 57% to 82% and from 66% to 84%, respectively. Accession
796 numbers of the sequences used in this alignment and comparisons are given in the
797 materials and methods section. HTLV-3-L indicates HTLV-3 strain Lobak18 and HTLV-3-C
798 indicates HTLV-3 isolate Cam2013OK.
799

800 **Table 1.** Immunogenic regions of HTLV1/2 structural proteins reported in the literature.

Region*	Amino acid sequence (one-letter code)	Ref.
	HTLV-1 Env gp46	
89-110	TKKPNRNGGGYYSASYSDPCSL	(43)
162-209	LLVDAPGYDPIWFLNTEPSQLPPTAPPLLPHSNLDHILEPSIPWKSK	(38)
175-199	FLNTEPSQLPPTAPPLLPHSNLDHI	(44–47)
176-199	LNTEPSQLPPTAPPLLPHSNLDHI	(43)
190-207	LLPHSNLDHILEPSIPW	(48)
190-209	LLPHSNLDHILEPSIPWKS	(49)
190-210	LLPHSNLDHILEPSIPWKSK	(43)
190-212	LLPHSNLDHILEPSIPWKSKLLT	(43)
191-215	LPHSNLDHILEPSIPWKSKLLTLV	(50, 51)
209-231	SKLLTLVQLTLQSTNYTCIVCID	(43)
224-244	YTCIVCIDRASLSTWHVLYSP	(43)
240-262	VLYSPNVSVPSSSSTPLLYPSLA	(43, 46, 47)
242-257	SPNVSVPSSSSTPLLY	(50, 51)
274-314	WTHCFDPQIQAIVSSPCHNSLILPPFSLSPVPTLGSRSRRA	(43)
296-312	ILPPFSLSPVPTLGSRSR	(49)
	HTLV-1 Env gp21	
346-405	SLLHEVDKDISQLTQAIVKNHKNLLKIAQYAAQNRRGLDLLFWEQGGLCKALQEQCRFPN	(50)
361-404	IVKNHKNLLKIAQYAAQNRRGLDLLFWEQGGLCKACQEQCRFPN	(52, 53)
374-392	AAQNRRGLDLLFWEQGGL	(49)
397-430	QEQCRFPNITNSHVSILQERPPLENRVLTGWGLN	(40, 52, 53)
417-425	PPLENRVLTG	(40)
417-430	PPLENRVLTGWGLN	(51)
467-489	QLRHLPSRVRYPHYSLIKPESSL	(43)
	HTLV-1 gag p19	
102-117	PPSSPTHDPPDSDPQI	(54)
105-124	SPTHDPPDSDPQIPPPYVEP	(48, 55)
100-130	PPPSSPTHDPPDSDPQIPPPYVEPTAPQVL	(44, 45)
	HTLV-2 Env gp46	
85-106	IKKPNRQGLGYYSPSYNDPCSL	(43)
85-102	KKPNRQGLGYYSPSYNDP	(51, 56)
172-208	ITSEPTQPPPTSPPLVHDSDLEHVLTPSTSWTTKILK	(43)
162-205	DAPGYDPLWFITSEPTQPPPTSPPLVHDSDLEHVLTPSTSWTTK	(38)
187-210	VHDSDLEHVLTPSTSWTTKILKFI	(50, 51)
173-209	TSEPTQPPPTSPPLVHDSDLEHVLTPSTSWITKILKF	(56)
178-205	QPPPTSPPLVHDSDLEHVLTPSTSWTTK	(57)
178-200	QPPPTSPPLVHDSDLEHVLTPST	(58)
182-199	TSPPLVHDSDLEHVLTPS	(59)
219-256	YSCMVCVDRSSLSSWHVLYTPNISIPQQTSSRTILFPS	(56)
	HTLV-2 Env gp21	
377-400	GLDLLPWEQGGLCKAIQEQCCFLN	(43)
370-396	YAAQNRRGLDLLFWEQGGLCKAIQEQC	(57)
	HTLV-2 gag p19	
116-134	PPSPEAHVPPPYVEPTTTQ	(58)

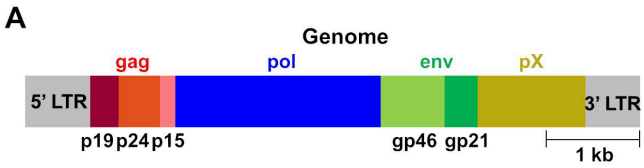
801 * regions are according to the cited reference.

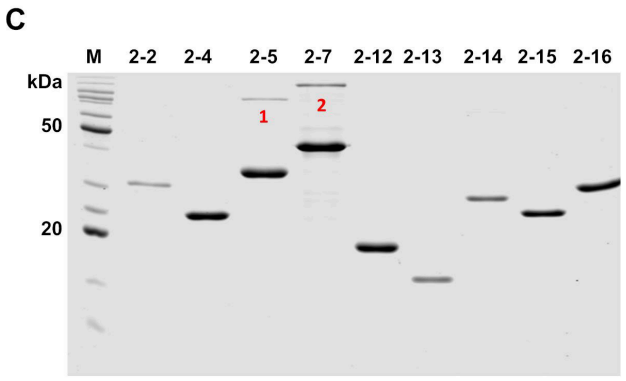
802

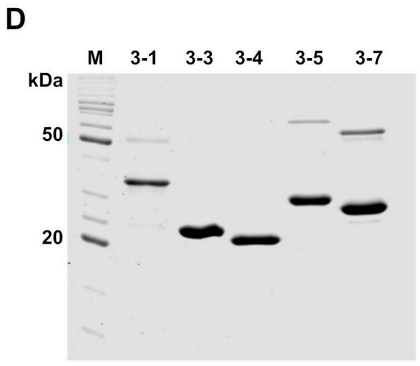
803 **Table 2**. Genetic constructs for expression of HTLV1/2 antigenic protein regions in *E. coli*.

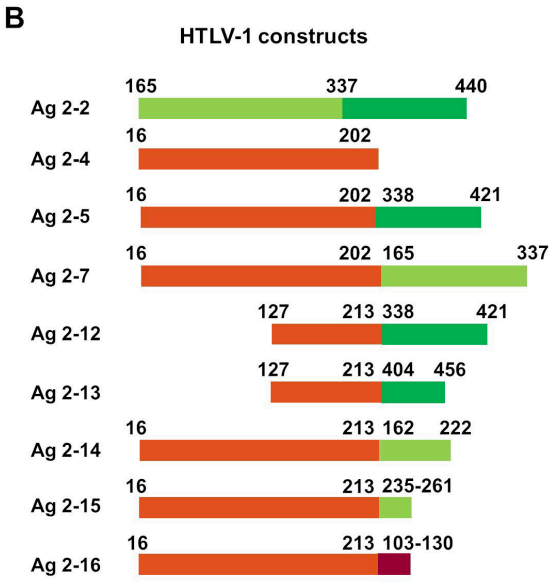
Name	Protein/Protein fusion	AA range	NCBI accession/ PDB reference	Histidine tag/ fusion linker	Expression	Approximate yield (mg/L of culture)
			HTLV-1			
Ag 2-1	gag p19	1 - 130	P03362.3	N-terminal/-	no expression	-
Ag 2-2	Env gp46-gp21	165 - 440	NP_057865.1	N-terminal/-	Insoluble	1.0
Ag 2-4	gag p24	16 – 202	X91888.1	N-terminal/-	Soluble	12.0
Ag 2-5	gag p24-Env gp21 fusion	16 – 202 338 – 421	X91888.1 NP_057865.1	N-terminal /GGSSCAAAAN	Insoluble	10.0
Ag 2-7	gag p24-Env gp46 fusion	16 – 202 165 – 337	X91888.1 NP_057865.1	N-terminal /GGSSCAAAAN	Insoluble	8.0
Ag 2-11[#]	gag p24-Env p21 fusion	16 - 213 360 - 421	X91888.1 NP_057865.1	C-terminal/GGS	Insoluble	-
Ag 2-12	gag p24-Env gp21 fusion	127– 213 338 – 421	1QRJ_A BAD95662.1	C-terminal /GGS	Soluble	6.0
Ag 2-13	gag p24-Env gp21 fusion	127– 213 404 – 456	1QRJ_A BAD95662.1	C-terminal /GGS	Soluble	2.0
Ag 2-14	gag p24-Env gp46 fusion	17 – 215 162 – 222	1QRJ_A AAF66016.1	C-terminal /GGS	Insoluble	1.5
Ag 2-15	gag p24-Env gp46 fusion	16 – 213 235 – 261	1QRJ_A AAF66016.1	C-terminal /GGS	Soluble	1.5
Ag 2-16	gag p24-gag p19 fusion	16 – 213 103 – 130	1QRJ_A P03362.3	C-terminal /GGS	Soluble	12.0
			HTLV2			
Ag 3-1	Env gp46-gp21	161 – 436	CAA61545	N-terminal/-	Insoluble	4.0
Ag 3-3	gag p24	17 – 204	NP_954567.1	N-terminal/-	Soluble	20.0
Ag 3-4	gag p19	2 – 137	1JVR	*N-terminal/-	Soluble	10.0
Ag 3-5	gag p24-Env gp21 fusion	17 – 204 334 – 436	NP_954567.1 CAA61545	C-terminal/GGS	Insoluble	5.0
Ag 3-6[$]	gag p24-env gp21 fusion	17 – 204 366 – 436	NP_954567.1 CAA61545	C-terminal /GGS	Insoluble	-
Ag 3-7	gag p24-Env gp46 fusion	17 – 204 162 – 208	NP_954567.1 CAA61545	C-terminal /GGS	Insoluble	12.0

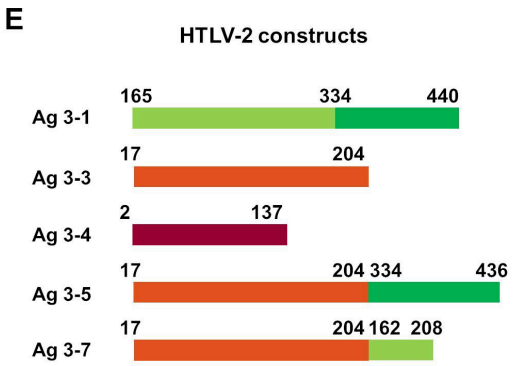
804 *cysteine residues 107 and 136 replaced by serine; [#]antigen 2-11 was not purified, its sequences are found in antigens 2-4
805 and 2-12. [$]antigen 3-6 was not purified, its sequence is comprised in antigens 3-3 and 3-5.



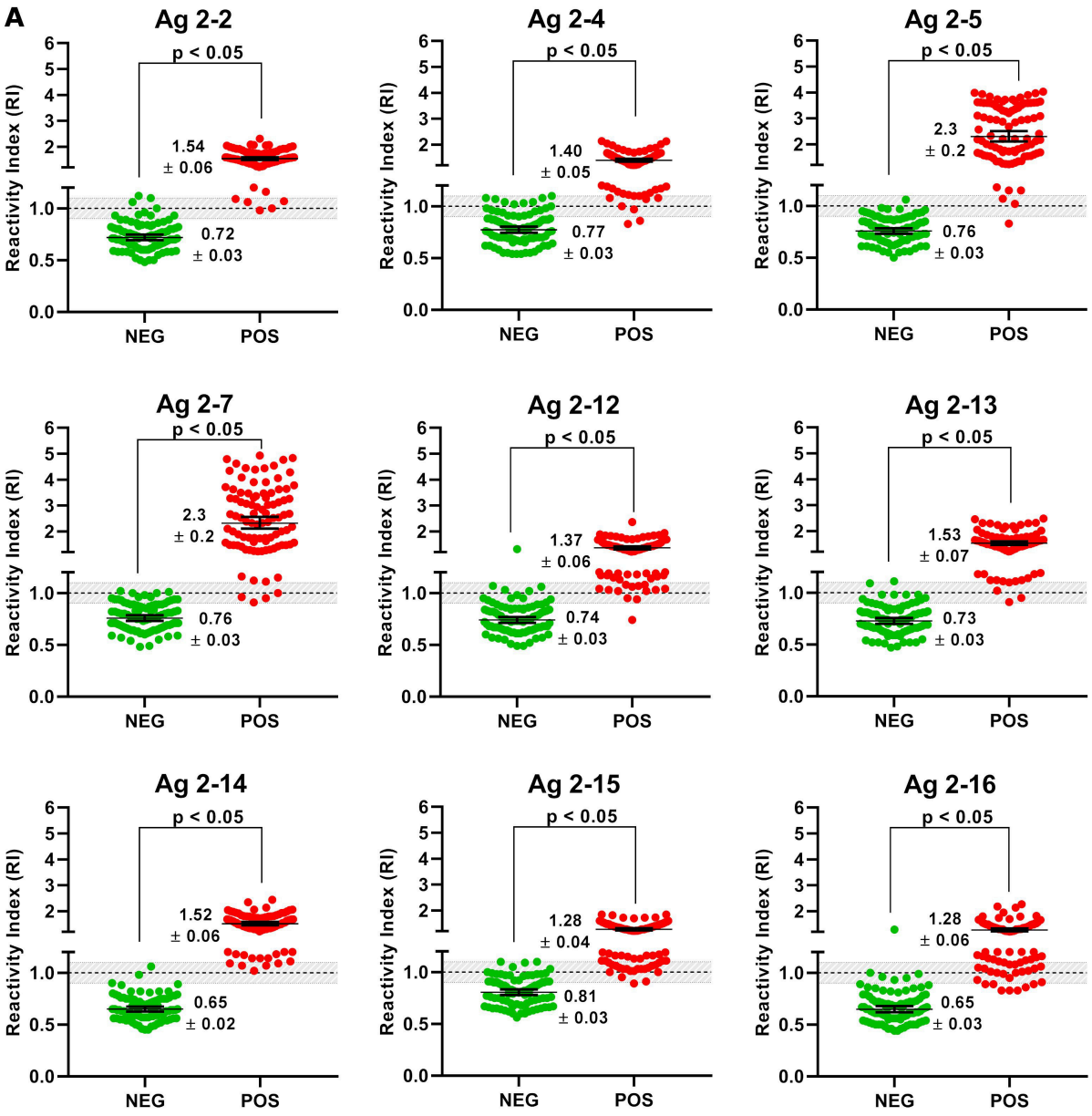








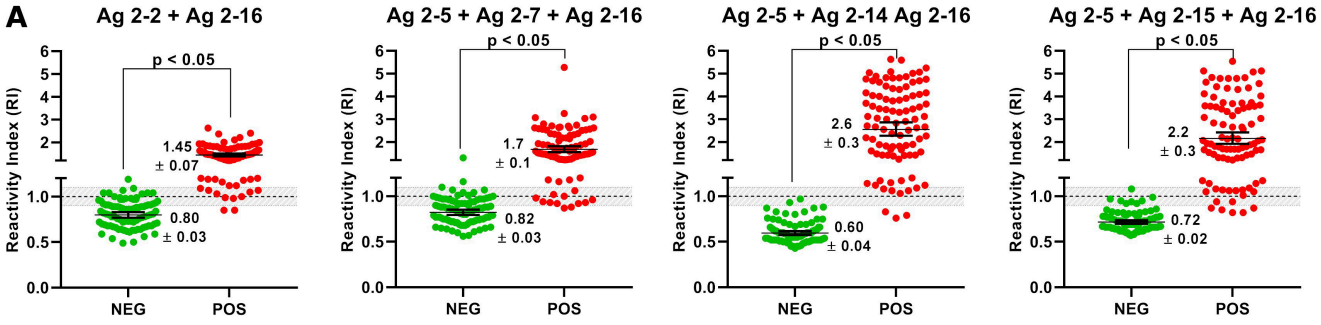
A



B

Antigen	AUC (CI 95%)	Sen. (CI 95%)	Spe. (CI 95%)	Acc. (CI 95%)	Kappa
Ag 2-2	99.8 (99.6-100)	98.9 (94.22–99.81)	95.7 (89.56–98.33)	97.3 (94.0-98.9)	0.95 (0.90-0.99)
Ag 2-4	98.9 (97.8-100)	96.8 (91.03–98.91)	90.4 (82.80–94.88)	93.6 (89.2-96.3)	0.87 (0.80-0.94)
Ag 2-5	99.6 (98.9-100)	98.9 (94.22–99.81)	98.9 (94.22–99.81)	99.9 (96.2-99.7)	0.98 (0.95-1.00)
Ag 2-7	99.6 (99.2-100)	96.8 (91.03–98.91)	95.7 (89.56–98.33)	96.3 (92.5-98.2)	0.93 (0.87-0.98)
Ag 2-12	98.6 (97.2-100)	96.8 (91.03–98.91)	94.7 (88.15–97.71)	95.7 (91.8-97.8)	0.92 (0.86-0.97)
Ag 2-13	99.8 (99.4-100)	97.9 (92.57–99.41)	97.9 (92.57–99.41)	97.9 (94.7-99.2)	0.96 (0.92-0.99)
Ag 2-14	99.9 (99.9-100)	100 (96.07–100)	98.9 (94.22–99.81)	99.5 (97.0-99.9)	0.99 (0.97-1.00)
Ag 2-15	98.4 (97.2-99.7)	96.8 (91.03–98.91)	93.6 (86.77–97.04)	95.2 (91.2-97.5)	0.90 (0.84-0.96)
Ag 2-16	98.6 (97.3-99.9)	88.3 (80.25–93.34)	97.9 (92.57–99.41)	93.1 (88.5-95.9)	0.86 (0.79-0.93)

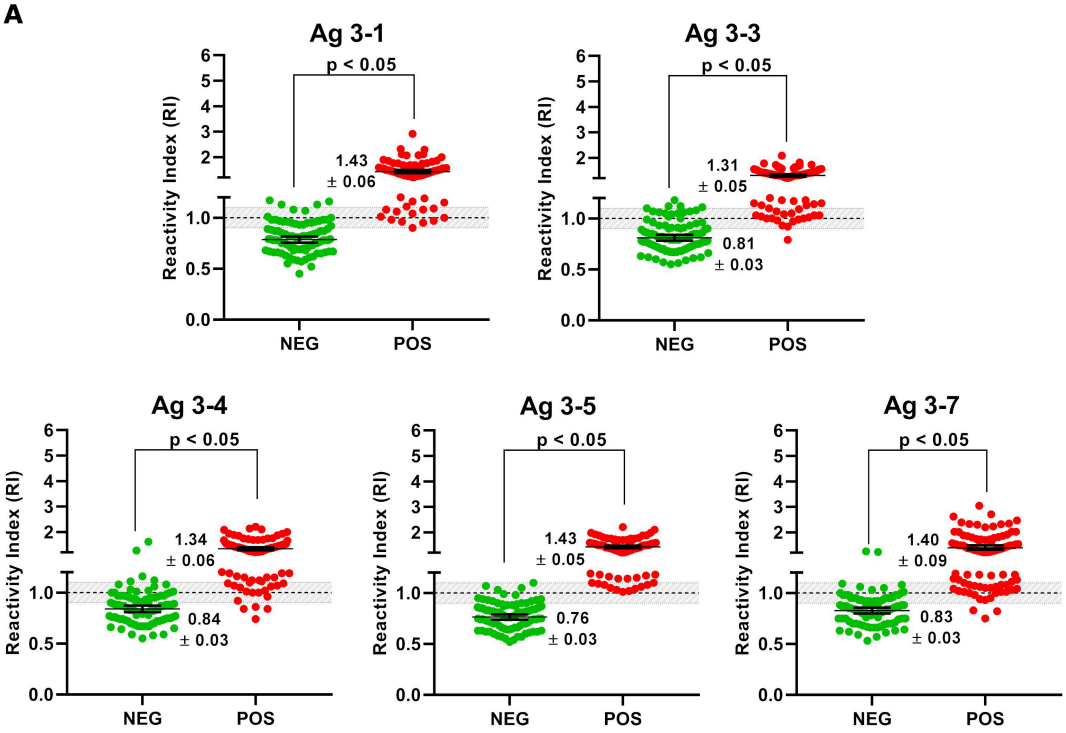
A



B

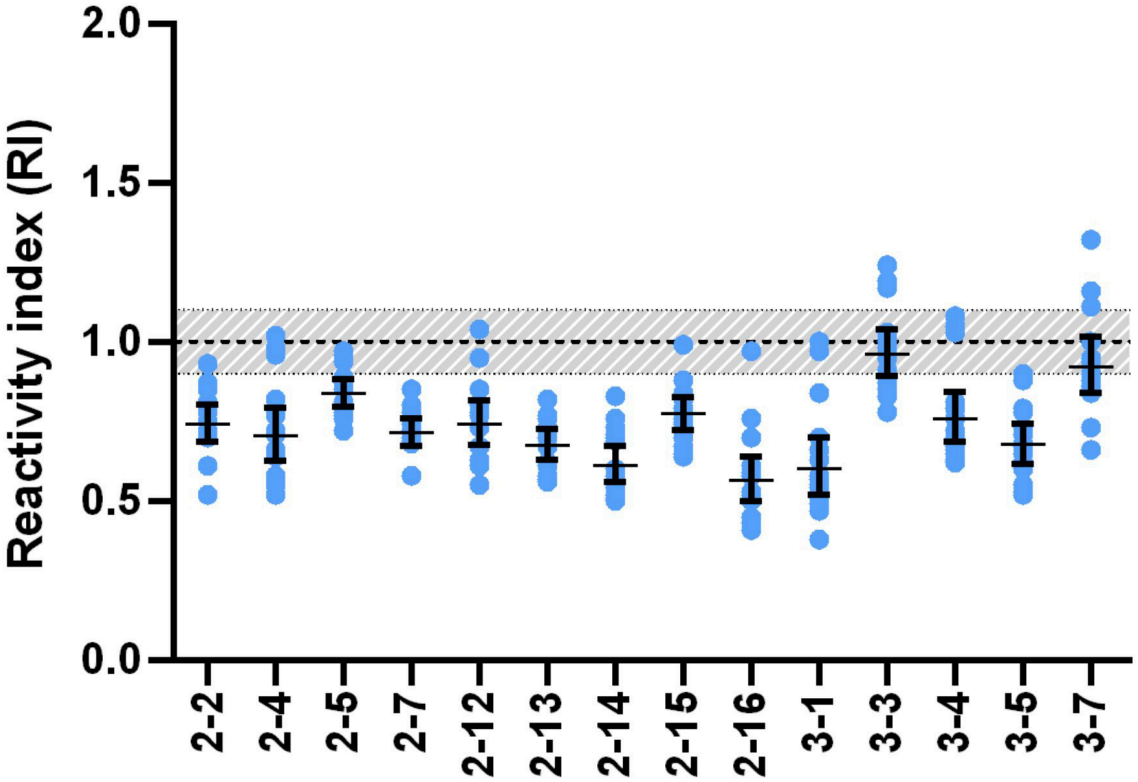
Antigen	AUC (CI 95%)	Sen. (CI 95%)	Spe. (CI 95%)	Acc. (CI 95%)	Kappa
2-2, 2-16	98.4 (96.9-99.9)	98.4 (96.93-99.92)	90.4 (82.80-94.88)	93.1 (88.5-95.9)	0.86 (0.79-0.93)
2-5, 2-7, 2-16	96.9 (94.9-98.9)	90.4 (82.80-94.88)	88.3 (80.25-93.34)	89.4 (84.1-93.0)	0.79 (0.70-0.87)
2-5, 2-14, 2-16	99.7 (99.4-100)	96.8 (91.03-98.91)	100 (96.07-100)	98.4 (95.4-99.5)	0.97 (0.93-1.00)
2-5, 2-15, 2-16	99.2 (98.4-99.9)	92.6 (85.42-96-35)	98.9 (94.22-99.81)	95.7 (91.8-97.8)	0.92 (0.86-0.97)

A



B

Antigen	AUC (CI 95%)	Sen. (CI 95%)	Spe. (CI 95%)	Acc. (CI 95%)	Kappa
Ag 3-1	98.5 (97.4-99.7)	94.7 (88.15-97.71)	91.4 (84.10–95.62)	93.1 (88.5-95.9)	0.86 (0.79-0.93)
Ag 3-3	96.6 (94.5-98.7)	93.6 (86.77-97.04)	84.0 (75.33–90.08)	88.8 (83.5-92.6)	0.78 (0.69-0.87)
Ag 3-4	95.0 (91.8-98.1)	93.6 (86.77-97.04)	84.0 (75.33–90.08)	88.8 (83.5-92.6)	0.78 (0.69-0.87)
Ag 3-5	99.7 (99.4-100)	100 (96.00-100)	95.7 (89.56–98.33)	97.9 (94.7-99.2)	0.96 (0.92-0.99)
Ag 3-7	95.7 (93.0-98.5)	92.6 (85.42-96.35)	89.4 (81.51–94.12)	91.0 (86.00-94.28)	0.82 (0.74-0.90)



Antigen	IR (95%CI)	CR (%)	GZ (%)
Ag 2-2	0.74 ± 0.06	0	1 (6.67%)
Ag 2-4	0.70 ± 0.08	1 (6.67%)	2 (13.33%)
Ag 2-5	0.84 ±0.04	0	3 (20.00%)
Ag 2-7	0.72 ±0.01	0	0
Ag 2-12	0.74 ± 0.07	1 (6.67%)	1 (6.67%)
Ag 2-13	0.68 ± 0.05	0	0
Ag 2-14	0.61 ± 0.06	0	0
Ag 2-15	0.77 ±0.05	0	1 (6.67%)
Ag 2-16	0.57 ± 0.07	0	1 (6.67%)
Ag 3-1	0.60 ± 0.09	1 (6.67%)	1 (6.67%)
Ag 3-3	0.96 ± 0.07	6 (40.00%)	4 (26.67%)
Ag 3-4	0.76 ± 0.08	3 (20.00%)	0
Ag 3-5	0.68 ± 0.06	0	1 (6.67%)
Ag 3-7	0.92 ± 0.09	4 (26.67%)	5 (33.33%)

A

```
                                                P24
HTLV-1    MKDLQAIKQEVSQAAPGSPQFMQTIRLAVQQFDPTAKDLQDLLQYLCSSLVASLHHQQLDSLISEAETRGITGYNPLAGPLRVQANNPQQQGLRREYQQL  100
HTLV-2    MKDLQAIKQEVSSSALGSPQFMQTLRLAVQQFDPTAKDLQDLLQYLCSSLVVSLHHQQLNTLITEAETRGMTGYNPMAGPLRMQANNPAQQGLRREYQNL  100
HTLV-3-L  MKDLQAIKQEVSSSAPGSPQFMQTVRLAVQQFDPTAKDLHDLLQYLCSSLVASLHHQQLETLIAQAETQGITGYNPLAGPLRVQANNPNQQGLRREYQNL  100
HTLV-3_C  MKDLQAIKQEVTSSAPGSPQFMQTVRLAVQQFDPTAKDLHDLLQYLCSSLVASLHHQQLETLIAQAETQGITGYNPLAGPLRVQANNPTQQGLRREYQNL  100
HTLV-4    MKDLQAIKQEISTSAPGSPQFMQTIRLAVQQFDPTAKDLHDLLQYLCSSLIVSLHHQQLQALIAEAETRGLTGYNPMAGPLRVQANNPAQQGLRREYQSL  100
          **********:: :* ********:**************:**********:.*******::**:**:***:*:*****:*****:***** *********.*

HTLV-1    WLAAFAALPGSAKDPSWASILQGLEEPYHAFVERLNIALDNGLPEGTPKDPILRSLAYSNANKECQKLLQARGHTNSPLGDMLRACQTWTPKDKTKAL  198
HTLV-2    WLAAFSTLPGNTRDPSWAAILQGLEEPYCAFVERLNVALDNGLPEGTPKEPILRSLAYSNANKECQKILQARGHTNSPLGEMLRTCQAWTPKDKTKVL  198
HTLV-3-L  WLSAFSALPGNTKDPTWAAILQGPEELFCSFVERLNVALDNGLPEGTPKDPILRSLAYSNANKECQKLLQARGQTNSSLGEMLRACQTWTPRDKNKIL  198
HTLV-3-C  WLSAFSALPGNTKDPTWAAILQGPEEPFCTFVERLNVALDNGLPEGTPKEPILRSLAYSNANKECQKLLQARGQTNGPLGDMLRACQAWTPRDKNKVL  198
HTLV-4    WLAAFAALPGNTRDPSWAAILQGLEEPYCAFVERLNVALDNGLPEGTPKEPILRSLAYSNANKECQKLLQARGHTNSPLGEMLRACQAWTPKDKTKVL  198
          **:**::***.::**:**:**** ** : :******:***********:****************:*****:**. **:***:**:***:**.* *

                                                P19
HTLV-1    MGQIFSRSASPIPRPPRGLAAHHWLNFLQAAYRLEPGPSSYDFHQLKKFLKIALETPARICPINYSLLASLLPKGYPGRVNEILHILIQTQAQIPSRPA-  100
HTLV-2    MGQIHGLSPTPIPKAPRGLSTHHWLNFLQAAYRLQPGPSDFDFQQLRRFLKLALKTPIWLNPIDYSLLASLIPKGYPGRVVEIINILVKNQVSPSAPAAP  100
HTLV-3-L  MGKTYSSPVNPIPKAPKGLAIHHWLNFLQAAYRLQPGPSEFDFHQLRKFLKLAIKTPVWLNPINYSVLAGLIPKNYPGRVHEIVAILIQETPAREA----  100
HTLV-3-C  MGNSYSRAANPIPKAPKGLAIHHWLNFLQAAYRLQPGPSEFDFHQLRNFLKLAIKTPVWLNPINYSVLAELVPKNYPGRIQEIIAILIQETSTQEV----  100
HTLV-4    MGQTHAS--SPVPKAPRGLSTHHWLNFLQAAYRLQPGPSEFDFHQLRRFLKLALQTPVWLNPIDYSLLASLIPKGYPGRVAEIVNILLRAHPPPSAPAIS   98
          **: ..   .*:*: *:**: *************:****.:**:**:.***:*::**  : **:**:** *:**.****: **: **::

HTLV-1    -P----PPPSSPTHDPPDSDPQIPPPYVEPTA--PQVL---  130
HTLV-2    VPTPISPTTTPPPPPPPSPEAHVPPPYVEPTT--TQSF---  136
HTLV-3-L  ----------PPSAPPADDPQKPPPYPEHAQVEPQCLPVL  126
HTLV-3-C  ----------PPSAPPASEPQNPPPYPEPGQAIPQCL---  123
HTLV-4    MPTAT-----GPAPAPQPQEAHTPPPYAEPAA--LQCL---  129
                    *   *   : : **** *      * :

                                                ENV
HTLV-1  MGKFLATLILFFQFCPLILGDYSPSCCTLTVGVSSYHSKPCNPAQPVCSWTLDLLALSADQALQPPCPNLVSYSSYHATYSLYLFPHWIKKPNRNGGGYY  100
HTLV-2  MGN--VFFLLL--FSLTHFPPVQQSRCTLTVGISSYHSSPCSPTQPVCTWNLDLNSLTTDQRLHPPCPNLITYSGFHKTYSLYLFPHWIKKPNRQGLGYY   96
HTLV-3L MGKSGLYFSLICFYTLFPSSFGNPSRCTLFIGASSYHSDPCGSNHPRCTWRLDLFSLTKDQSLSPPCPGLVTYSQYHKPYSLYVFPHWIAKPDRRGLGYY  100
HTLV-3C MGKSSLFICLFCSYMASLFVPGDPSRCTLFIGASSYHSSPCGSNYPQCTWTLDLVSLTRDQSLNPPCPDLVTYSQYHRPYSLYLFPHWITKPNRQGLGYY  100
HTLV-4  MGN--VFFLIL--LATLGAPVLRASRCTLTVGISSYHSSPCSPAQPLCTWALDLVSITKDQLLYPPCQNLITYSNYHKTYSLYLFPHWVQKPLRRGLGYY   96
        **:    : ::             * *** :* *****.**.   * *:* *** ::: ** * *** .*::** :*  ****:****: ** *.* ***

HTLV-1  SASYSDPCSLKCPYLGCQSWTCPYTGAVSSPYWKFQQDVNFTQEVSHLNINLHFSKCGFSFSLLVDAPGYDPIWFLNTEPSQLPPTAPPLLSHSNLDHIL  200
HTLV-2  SPSYNDPCSLQCPYLGCQSWTCPYTGPVSSPSWKFHSDVNFTQEVSQVSFRLHFSKCGSSMTLLVDAPGYDPLWFITSEPTQPPPTPPPLVHDSDLEHVL  196
HTLV-3L SASYSDPCAIQCPYLGCQSWTCPYTGPVSSPHWKYTSDLNFTQEVSSISLHLHFSKCGSSFSFLLDAPGYDPVWLLSSQATQIPPTPAPLIRDSDLQHIL  200
HTLV-3C SASYSDPCAIKCPYLGCQSWTCPYTGPMSSPYWKYTSDLNFTQKVSSVTLHLHFSKCGSSFSLLLDAPGYDPVWFLSSQTTQAPPTPAPLTQDSDFQHIL  200
HTLV-4  SASYSDPCSLQCPYLGSQSWTCPYTGPVSSPTWRFFTDVNFTQEVSRVSLKLHFSKCGSSLTLLIDAPGYDPLWYLTSEPTQEPPTPPPLVSDSDLDHVL  196
        * **.***:::*****.********* :*** *::  *:****:** :.:.******* *:::*:*******:* :.:: :* ***  **  .*:::*:*

HTLV-1  EPSIPWKSKLLTLVQLTLQSTNYTCIVCIDRASLSTWHVLYSPNVSVPS--PSSTPLLYPSLALPAPHLTLPFNWTHCFDPQIQAIVSSPCHNSLILPPF  298
HTLV-2  TPSTSWTTKMLKFIQLTLQSTNYSCMVCVDRSSLSSWHVLYTPNISIPQ-QTSSRTILFPSLALPAPP-FQPFPWTHCYQPRLQAITTDNCNNSIILPPF  294
HTLV-3L EPSIPWSSKILNLILLALKSTNYSCMVCVDRSSLSSWHVLYDPLKAPSSPDPQAQSILRPSLAIPASNITPPFPWTHCYRPPLQAISSENCNNSVILPPF  300
HTLV-3C EPSVPWSSKILNLILLTLKSTNYSCMVCVDRSSLSSWHVLYDPLKVPKQHEPRARALLRPSLAIPITNTTPPFPWSHCYCPLLQAVISNNCNNSVILPPF  300
HTLV-4  TPSASWASKMLTLIHLTLQSTNYSCMVCVDRASLSSWHVLYTPNISSN---APSKPIVRPSLALSAPR-PQPFPWTHCYQPQVQAITTAKCNNSIILPPF  292
         **  * :*:*.:: *:*:****:*:**:**:***:***** *         :  :: ****:       ** *:**: * :**: :  *:**:*****

HTLV-1  SLSPVPTL-GSRSRRAVPVAVWLVSALAMGAGVAGRITGSMSLASGKSLLHEVDKDISQLTQAIVKNHKNLLKIAQYAAQNRRGLDLLFWEQGGLCKALQ  397
HTLV-2  SLAPVPPP-ATRRRRAVPIAVWFVSALAAGTGIAGGVTGSLSLASSKSLLFEVDKDISHLTQAIVKNHQNILRVAQYAAQNRRGLDLLFWEQGGLCKAIQ  393
HTLV-3L SLSPIPDVSRPRKRRAVPIAIWLVSALAAGTGIAGGVTGSLSLASSKSLLREVDQDIDHLTRAIVKNHDNILRVAQYAAQNRRGLDLLFWEQGGLCKAIQ  400
HTLV-3C SLSPVPDLSKPRQRRAVPIAVWLVSALAVGTGIAGGTTGSLSLASSRSLLHEVDQDISHLTQAIVKNHNNILRVAQYAAQNRRGLDLLFWEQGGLCKAIR  400
HTLV-4  SLSPLPGASLTRRRRAVPVAVWLVSALAAGTGIAGGVTGSLSLASSRSLLSEVDKDISHLTQAIVKNHQNILRVAQYAAQNRRGLDLLFWEQGGLCKAIQ  392
        **:*:*     * *****:*:*:***** *:*:**  ***:****.:*** ***:**.:**:******.*:*::**********************::

HTLV-1  EQCCFLNITNSHVSILQERPPLENRVLTGWGLNWDLGLSQWAREALQTGITLVALLLLVILAGPCILRQLRHLPSRVR--YPHYSLINPESSL  488
HTLV-2  EQCCFLNISNTHVSVLQERPPLEKRVITGWGLNWDLGLSQWAREALQTGITILALLLLVILFGPCILRQIQALPQRLQNRHSQYALINPETML  486
HTLV-3L EQCCFLNISNTHVSVLQERPPLEKRVITGWGLNWDLGLSQWAREALQTGITLLALFLLLIVVGPCVIRQLQTLPSRLQHRSQPYSLLNYETNL  493
HTLV-3C EQCCFLNISNTHVSVLQERPPLEKRVITGWGLNWDLGLSQWAREALQTGITLLALFLLLILVGPCVLRQLQALLFRLQHRSHPYSLLNRETNL  493
HTLV-4  EQCCFLNISNTHISVLQERPPLETRVTTGWGLNWDLGLSQWAREALQTGITLLALLLLIIILGPCIIRQLQALPQRLQQRPDQYPLLNPETPL  485
        ********:*:*:*:********.** ************************::**:**:*: ***:**:: *  *::     * *:* *: *
```

B

Amino acid identity (%)												
	P24				P19				ENV			
	HTLV-2	HTLV-3-L	HTLV-3-C	HTLV-4	HTLV-2	HTLV-3-L	HTLV-3-C	HTLV-4	HTLV-2	HTLV-3-L	HTLV-3-C	HTLV-4
HTLV-1	86	85	85	88	59	67	67	58	72	66	67	69
HTLV-2		85	86	92		71	57	77		73	74	81
HTLV-3-L			95	86			83	64			84	73
HTLV-3-C				87				72				71

Journal of Clinical Microbiology