Molecular Epidemiology of HTLV-II among United States Blood Donors and Intravenous Drug Users: An Age–Cohort Effect for HTLV-II RFLP Type a0

Edward L. Murphy,*2,1 Renaud Mahieux,2 Guy de The2 Fredj Tekaia,3 Dannie Ameti,§ Jacqueline Horton,§ and Antoine Gessain†

*Departments of Laboratory Medicine, Medicine and Epidemiology/Biostatistics, University of California San Francisco, San Francisco, California 94143; †Unité d'Epidémiologie des Virus Oncogènes, and ‡Unité de Génétique Moléculaire des Levures, Institut Pasteur, Paris, France; and §Westat Inc., Rockville, Maryland

Received August 19, 1997; returned to the author for revisions October 8, 1997; accepted December 19, 1997

Molecular subtyping was used to investigate the epidemiology of human T-lymphotropic virus type II (HTLV-II) in the United States. Nested polymerase chain reaction of the HTLV-II long terminal repeat region followed by restriction fragment length polymorphism (RFLP) analysis was performed on HTLV-II seropositive subjects including 97 U.S. blood donors without major risk factors for HTLV-II infection, 53 injection drug users (IDU), and 10 American Indian blood donors. Three new HTLV-II RFLP types were confirmed with DNA sequencing and phylogenetic analysis. HTLV-II RFLP type a0 (Switzer classification) was associated with older age [adjusted odds ratio (OR) 1.06 per year of age, 95% confidence interval (CI) 1.02–1.09] and with Black (OR 5.24, 95% CI 1.90–14.47) and White (OR 4.43, 95% CI 1.67–11.75) race/ethnicity. These data are consistent with an age–cohort effect for HTLV-II RFLP type a0 among older White and Black IDU and blood donors. This finding could be explained by an epidemic of non-a0 HTLV-II RFLP types among younger persons of Hispanic and other race/ethnicity, superimposed upon endemic HTLV-II RFLP type a0 among older Black and White persons. © 1998 Academic Press

INTRODUCTION

Human T-lymphotropic virus type II (HTLV-II), first isolated (Kalyanaraman et al., 1982) and characterized (Chen et al., 1983) in the early 1980’s, is molecularly related to other primate T-lymphotropic viruses including HTLV-I and related simian viruses (Vandamme et al., 1982; Cereseto et al., 1996). It is endemic in some but not all Amerindian tribes in North, Central, and South America, but has also become epidemic among injection drug users (IDU) in North and South America and, to a lesser extent, in Europe (Hall et al., 1996). Seroepidemiologic studies of HTLV-II among Amerindians reveal overall prevalence of 3 to 30% (Hjelle et al., 1994; Maloney et al., 1992; Vitek et al., 1995). Mother to child transmission is indicated by concordance of serostatus and viral sequence between mother and children (Vitek et al., 1995; Lal et al., 1993), and an increase in seroprevalence with age is consistent with sexual transmission (Vitek et al., 1995). Among IDU, there is evidence that the virus has been epidemic since at least the 1970’s in North America (Biggar et al., 1991; Khabbaz et al., 1991), with prevalences ranging from 1 to 18% in various North American cities (Khabbaz et al., 1992). Ongoing transmission of HTLV-II in the United States may be inferred from the strong association of HTLV-II seroprevalence with the duration of drug injection (Feigal et al., 1991), and from recent data revealing an HTLV-II sero-incidence of 0.7 per 100 person-years among IDU (Vlahov et al., 1995). Further complicating the epidemiologic picture is the sporadic demonstration of HTLV-II in Africans, including prostitutes in Ghana (Igarashi et al., 1993) and Cameroon (Mauclere et al., 1995), a family cluster in Gabon (Tuppin et al., 1996), and the isolated Central African Bambuti and Bakola pygmy tribes (Goubau et al., 1992; Gessain et al., 1995).

Thus, HTLV-II appears to be an endemic infection of ancient people in both Africa and the Americas. It has also been epidemic among IDU over a period of at least 20 years (Biggar et al., 1991), and may be considered an unrecognized precursor of the HIV epidemic. However, several important questions concerning the modern epidemiology of HTLV-II remain to be answered. When and from which endemic population did HTLV-II emerge to infect IDU? Are there populations in the United States with endemic HTLV-II other than Amerindians? And finally, to what extent has HTLV-II spread beyond Amerindians and IDU into the general North American population by means of sexual and maternal child transmission?

Restriction fragment length polymorphism (RFLP) has recently been applied to the molecular subtyping of HTLV-II (Switzer et al., 1995; Eiraku et al., 1995). The
literature clearly indicates that RFLP types a0 and a3 [classification of Switzer and Heneine (Switzer et al., 1995), phylogroup A-III], predominate among IDU but are rare among Amerindians, whereas RFLP types b5 is almost exclusively found among Amerindians (Switzer et al., 1995). In addition, overall HTLV-II seroprevalence is highest among older, Black and Hispanic IDU in eastern and western U.S. cities (Khabbaz et al., 1992; Feigal et al., 1991) groups with little apparent contact with American Indians. These observations led us to the hypothesis that Indians may not be the only source of the HTLV-II epidemic among IDU. We sought an endemic focus of HTLV-II among infected U.S. blood donors without major risk factors, and compared the prevalence of HTLV-II RFLP types among young and old IDU to search for evidence of an age–cohort effect for HTLV-II infection.

RESULTS

Study population

The study population consisted of 160 individuals, including 97 subjects from the Retrovirus Epidemiology Donor Study (REDS) cohort without major risk factors for HTLV-II, 53 young and old IDU from the San Francisco emergency room, and 10 samples from American Indian blood donors. The demographic characteristics of the subjects, by group, are given in Table 1. The median age of the subjects was 34 (range 18 to 78), and roughly two thirds of the subjects were female.

RFLP analysis

The HTLV-II RFLP types of the study population, as determined by RFLP, are given in Table 2, with representative photographs in Fig. 1. RFLP type a0 was most commonly observed, with 121 (76%) of all subjects classified in this group, followed by b4 (n=12, 8%), b5 (n=12, 8%), a3 (n=5, 3%). We observed three low risk subjects with RFLP type a5, all of whom were former blood donors from Los Angeles: a 22-year-old White male, a 38-year-old Hispanic female, and a 57-year-old American Indian male. We also found one San Francisco IDU with the a2 RFLP type and two American Indians with the a7 RFLP type.

Four specimens representing three potential new RFLP types could not be classified into the RFLP classification of Switzer and Heneine (Switzer et al., 1995) (see Fig. 1). These included two ‘bx’ specimens (samples 6–4 from a 32-year-old Hispanic female and 4–10 from a 31-year-old White female, both from San Francisco, IDU) which gave a b4 RFLP pattern except that they were cut with Sau I at a site similar to the ‘a’ RFLP types; one ‘by’ specimen (sample 15±21 from a 29-year-old low-risk Black female from San Francisco).
Francisco) with a b5 pattern except that it had a novel Eco47III restriction site; and one "bz" specimen (sample 9–17 from a 49-year-old low-risk White female from Detroit) with a b4 pattern except that it had three rather than two BanII sites. All new RFLP types were confirmed on subsequent proviral DNA sequencing.

### TABLE 2

HTLV-II Phylogenetic Groups and RFLP Types, Classified According to Switzer et al. (1995), by Subject Group

<table>
<thead>
<tr>
<th>HTLV-II phylogenetic group</th>
<th>HTLV-II RFLP type</th>
<th>Low risk n = 97</th>
<th>IDU n = 53</th>
<th>Amerindian n = 10</th>
<th>All subjects N = 160</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>a0</td>
<td>73 (75)</td>
<td>45 (85)</td>
<td>3 (30)</td>
<td>121 (76)</td>
</tr>
<tr>
<td></td>
<td>a2</td>
<td>0</td>
<td>1 (2)</td>
<td>0</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td>a3</td>
<td>3 (3)</td>
<td>2 (4)</td>
<td>0</td>
<td>5 (3)</td>
</tr>
<tr>
<td></td>
<td>a5</td>
<td>3 (3)</td>
<td>0</td>
<td>0</td>
<td>3 (2)</td>
</tr>
<tr>
<td></td>
<td>a7</td>
<td>0</td>
<td>0</td>
<td>2 (20)</td>
<td>2 (1)</td>
</tr>
<tr>
<td></td>
<td>All RFLP type a</td>
<td>79 (81)</td>
<td>48 (91)</td>
<td>5 (50)</td>
<td>132 (83)</td>
</tr>
<tr>
<td>IDU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b4</td>
<td>10 (10)</td>
<td>2 (4)</td>
<td>0</td>
<td>12 (8)</td>
</tr>
<tr>
<td></td>
<td>&quot;bz&quot;</td>
<td>1 (1)</td>
<td>0</td>
<td>0</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td>b5</td>
<td>6 (6)</td>
<td>1 (2)</td>
<td>5 (50)</td>
<td>12 (8)</td>
</tr>
<tr>
<td></td>
<td>&quot;by&quot;</td>
<td>1 (1)</td>
<td>0</td>
<td>0</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td>b-New</td>
<td>0</td>
<td>2 (4)</td>
<td>0</td>
<td>2 (1)</td>
</tr>
<tr>
<td></td>
<td>All RFLP type b</td>
<td>18 (19)</td>
<td>5 (9)</td>
<td>5 (50)</td>
<td>28 (18)</td>
</tr>
</tbody>
</table>

Note. Percentages may not add to 100 due to rounding.

- The two a7 RFLP subtypes and the PH230CAM sequences were related to both A-II and A-III in our phylogenetic analysis.
- The literature classifies RFLP subtype b4 in the B-II phylogroup; however, the one b4 sequenced in the current study clustered in the B-III group.

Francisco) with a b5 pattern except that it had a novel Eco47III restriction site; and one "bz" specimen (sample 9–17 from a 49-year-old low-risk White female from Detroit) with a b4 pattern except that it had three rather than two BanII sites. All new RFLP types were confirmed on subsequent proviral DNA sequencing.

**FIG. 1.** Restriction fragment length polymorphism (RFLP) patterns for representative specimens. Photos represent the products of restriction cleavage with enzymes AvaiI, BaniI, Eco47III, Saul, DraiI, and SacII electrophoresed on either 2% agarose or 1% agarose plus 2% low melting point agarose and visualized with ethidium bromide. For each panel, the lanes are as follows: molecular weight standard in 100 bp increments, negative control a0 specimen without enzyme, specimens representing RFLP types a0, a2, a3, a5, and a7, specimens representing RFLP types b4 and b5, and new b RFLP type samples "bx", "by", and "bz". The "bz" lanes from the AvaiI and Saul panels were added to the photograph from a separate gel. Restriction with XhoI is not shown since it was positive only in the one a2 specimen.
DNA sequence data

We chose to clone and sequence the 672-bp nested polymerase chain reaction (PCR) product from one specimen in each of the more common RFLP types that we observed (a0, a2, a3, b4, and b5) plus two a5, two a7, and all four specimens in the three novel RFLP types. All sequences were unique, although the two specimens in the new RFLP type bx RFLP type differed by only 2 bp. Phylogenetic analyses of the 625-bp long terminal repeat (LTR) segment using both the DNAPARS program and the neighbor-joining methods gave similar results; the DNAPARS-generated phylogenetic tree is shown in Fig. 2a.
complete separation of the HTLV-II-A and -B phylogroups was obtained (bootstrap value 100%).

Within the A group, the phylogroup A-I Navajo isolate NAVDS was well separated from all other A isolates (bootstrap 99%). Separate subgroups within the A group included the AII phylogroup (a5±a6 RFLP types) with Brazilian Kayapo and the Ghanian prostitute GHKT, the two a5 RFLP samples from the current study (samples 11±15 and 12±8) and the son of a Mexican prostitute (mexy17); the AIII phylogroup including the MO prototype as well as the a3 (sample 10±18), a2 (sample 6±2), and a0 (sample 5±5) RFLP types from the current study (bootstrap 82% for the A-II/A-III separation). In contrast to previous reports, our phylogenetic analysis classified the Cameroonian prostitute PH230CAM and our two a7 RFLP types (samples 15±5 and 15±8) into a small group related to both phylogroups A-III and A-II.

Within the B group, the B-III phylogroup included several Amerindian isolates, including our b5 (14±17) and by (15±21) RFLP types, and, surprisingly, our b4 (12±7) isolate which should have been located in the B-II cluster with other published b4 samples. The bootstrap value for
the separation between the JG (phylogroup B-I) sample from New York and the B-II samples was not strong. Phylogroup B-II included a number of isolates from New York, Spanish, and Italian IDU as well as our new bz (9±17) RFLP type. Finally our two new bx (4±10, 6±4) samples were separate from the B-I, B-II, and B-III clusters (bootstrap = 53%) and the B-IV phylogroup (bootstrap = 41%), which included wyu1 from a Colombian Wayuu Indian and G12 from a Guaymi Indian.

The final phylogenetic analysis with the MEGA program yielded a tree (Fig. 2b) similar to that of the previous methods, but with an estimation of genetic distance given by the scale bar. In this tree, phylogroups B-II and B-III appear to be poorly separated. However, our new bx (4±10, 6±4) samples are still weakly separated from phylogroups B-II/B-III and B-IV with bootstrap values of 50 and 54%, respectively, similar to those of the DNAPARS analysis.

Epidemiologic analysis

Because 76% of subjects in the current study were the IDU-associated RFLP type a0 in the Switzer and Heneine classification (Switzer et al., 1995), we chose to analyze the demographic characteristics of this RFLP type in comparison to that of all others combined. The proportion of HTLV-II RFLP type a0 among all HTLV-II RFLP types is shown for different subgroups of subjects in Fig. 3. RFLP type a0 was relatively more common among older than younger subjects [adjusted odds ratio (OR) 1.05 per year of age, 95% confidence interval (CI) 1.02–1.09]; there was no RFLP type difference by gender (data not shown). Both Black subjects (OR 5.81, 95% CI 2.25–14.92) and White subjects (OR 3.80, 95% CI 1.53–9.45) had a higher proportion of RFLP type a0 compared to combined Hispanic, American Indian, and Asian subjects. The racial difference was even more marked when the analysis was limited to women; 84% of Black women (OR 7.20, 95% CI 1.99–26.09) and White women (OR 5.83, 95% CI 1.84–18.54), compared to 7/13 Hispanic, 0/2 Asian, or 2/6 American Indian women (43% combined), were RFLP type a0.

Within the IDU subject group, the prevalence of RFLP type a0 was 74% in those under 40 compared to 93% in those over 40 (P = 0.06, Fishers exact test). There was no difference in a0 prevalence between the 7 HIV seropositive and the 46 HIV seronegative IDUs. Within the "low risk" subject group (the only group with representation in all metropolitan areas) the prevalence of RFLP type a0 was highest in the Washington/Baltimore area (100%), intermediate in the Oklahoma City (83%), San Francisco (83%), and Detroit (75%) areas, and lowest in the Los Angeles (65%) area. The a0 prevalence for Los
Angeles was significantly lower than that of the other four metropolitan areas combined (OR 0.28, 95% CI 0.10–0.77). Because of the unexpectedly high RFLP type a0 prevalence among the low risk subjects, we also examined data on sexual behavior in this group. Although the mean number of lifetime male partners was over 10 for female low risk subjects, there was no difference in the number of partners between RFLP type a0 subjects versus non-a0 female subjects.

Logistic regression modeling was performed to control for potential confounding; however, the complexity of the models was limited by the moderate sample size of the study. In a model which included age, sex, race, and geographic location the variables significantly associated with HTLV-II RFLP type a0 were age (OR=1.06 per year, 95% CI 1.02–1.09), Black race (OR=5.23, 95% CI 1.90–14.43), and White race (OR=4.28, 95% CI 1.58–11.57) (both versus combined Hispanic/Indian/Asian race). The source of the subject (low risk vs. IDU vs. American Indian) was not associated with RFLP type a0 and was not retained in the final model. In a separate model for the low risk subjects with good geographic representation, subtype a0 was still associated with age (OR=1.06 per year, 95% CI 1.01–1.12) and Black race (OR=4.69, 95% CI 1.02–21.51), but only nonsignificantly with White race (OR=2.87, 95% CI 0.73–11.23) and inversely with Los Angeles (OR=0.42, 95% CI 0.13–1.39).

DISCUSSION

Our demonstration of HTLV-II RFLP types a2, a5, and a7 in North American IDU, low risk blood donors and Indians, respectively, adds to the reported number of these rare RFLP types in the United States. RFLP type a2 has previously been reported in five Norwegian IDU and two blood donors from Atlanta, Georgia (Switzer et al., 1995). Given the high prevalence of HTLV-II among U.S. IDU (Khabbaz et al., 1992), it is more likely that RFLP type a2 was introduced from U.S. into Norwegian IDU, rather than vice versa. The opposite is probably true for RFLP type a5, which is most prevalent among the Kayapo, Kraho, and other Brazilians. This RFLP type has previously been reported in a prostitute from Ghana and the child of a Mexican prostitute (Switzer et al., 1996), and our finding of this RFLP type among three low risk U.S. blood donors in Los Angeles implies the penetration of the RFLP type into North America. Finally, our data on HTLV-II RFLP type a7 among two American Indians from Oklahoma is difficult to reconcile with the single previous isolation of this RFLP type from a prostitute in Cameroon (Mauclere et al., 1995).

We have also identified three new HTLV-II RFLP types among our low risk blood donors and IDU. Our new by genotype differed from the b5 RFLP type only in having an extra Eco47III restriction site and was located near b5 in the phylogenetic analysis. The one bz and the two bx samples differed from the b4 RFLP type only in having extra BanII and SauI restriction sites, respectively. Sample b2 was grouped in the genotype B-II with other published RFLP type b4 sequences, whereas the two bx samples formed a new genotype intermediate between the Wayuu/Guaymi Indian B-IV genotype and the other B genotypes. The case of our sample 12±7 which had an RFLP b4 phenotype, but was located with other published b5 rather than b4 sequences in the phylogenetic analysis, was the single exception in our study to the correspondence between RFLP and sequence analysis. These data indicate that RFLP is useful in screening for new RFLP types. However, novel RFLP types based upon single restriction enzyme differences could be caused by PCR errors and should be interpreted with caution. Phylogenetic analysis based upon proviral DNA sequence remains the gold standard for defining new HTLV-II RFLP types.

The distribution of RFLP types in both the low risk and IDU groups was similar to that of samples from multiple U.S. sources (Switzer et al., 1995), or New York blood donors (Eiraku et al., 1995) in the two other major studies using RFLP of the LTR region to classify HTLV-II RFLP types. Thus, although all low risk subjects denied IDU, sex with an IDU and blood transfusion, their infections probably originated indirectly from the large pool of IDU-related HTLV-II in the United States, probably by sexual transmission. This inference is supported by the preponderance of females in the low risk group compared to the male excess in the IDU group and by the relatively high number of lifetime sexual partners in these females. Previously published data from the REDS cohort indicate that the number of lifetime sexual partners is a risk factor for HTLV-II infection independent of IDU or sex with an IDU (Schreiber et al., 1997).

We found that the distribution of HTLV-II RFLP type a0 was not uniform among the subjects, being significantly higher in older vs. younger, and in Black and White vs. other race subjects. There was also a statistically insignificant lower prevalence of RFLP type a0 in subjects from Los Angeles. Within the IDU group, a0 prevalence was higher in subjects over 40 years old compared to subjects less than 40 years old, a finding that persisted after controlling for sex and race. These data are consistent with an age-cohort effect for HTLV-II RFLP type a0 among older White and Black IDU and blood donors. Such an age-cohort effect could be produced by either an historic epidemic of HTLV-II RFLP type a0 or a more recent epidemic of non-a0 RFLP types. Another study has reported age, race, and geographic variation in HTLV-II prevalence among IDU, although HTLV-II subtype was not measured (Briggs et al., 1995).

The demonstration of a cohort effect for RFLP type a0 has implications for understanding the origin of HTLV-II infection in the United States. There is little evidence in the literature for RFLP type a0 infection among American
Indians (Switzer et al., 1995; Eiraku et al., 1995). Since the duration of IDU is correlated with age among IDU (Feigal et al., 1991), data from the current study suggest that HTLV-II RFLP type a0 was endemic among both Black and White IDU at least 20 to 30 years ago, perhaps from an unrecognized endemic focus in these racial/ethnic groups. Younger subjects, and those of Hispanic, Asian, and American Indian race/ethnicity, had a higher proportion of non-a0 RFLP types, suggesting the subsequent admixture of these RFLP types with the predominant a0 RFLP type. In such a scenario, the introduction of HTLV-II from American Indians or another source into IDU is probably a more recent phenomenon than inferred by other authors (Hall et al., 1996; Heneine et al., 1996). Further RFLP subtyping of older, non-IDU HTLV-II seropositive individuals could confirm our hypothesis of an unrecognized endemic focus of HTLV-II RFLP type a0 in the United States.

An alternate explanation for the HTLV-II epidemic among IDU would be a founder effect, initiated by the contact between a rare Indian with RFLP type a0 and an IDU more than 30 years ago. This RFLP type then expanded, only to be admixed with other HTLV-II RFLP types among younger IDU. Although there is little data on the relative viral load or transmission efficiency of different HTLV-II RFLP types, another possibility is greater epidemiologic fitness of RFLP subtype a0. One study has reported a similar distribution of RFLP types among HTLV-II infected men who did or did not transmit their infection to their long-term sexual partner (Kaplan et al., 1996), the correlation between viral load and HTLV-II RFLP type was not examined. Additional studies of the transmission and pathogenesis of HTLV-II infection should include information on RFLP types.

In conclusion, we report novel and unusual HTLV-II RFLP types among both low risk blood donors and IDU in the United States, suggesting that the epidemiology of this virus is complex and not determined solely by geography or race/ethnicity. We have also found evidence for an age-cohort effect of HTLV-II RFLP type a0 among older HTLV-II infected subjects, and an association of this RFLP type with Black and White race/ethnicity. This finding could be explained by an epidemic of non-a0 HTLV-II RFLP types among younger persons of Hispanic and other race/ethnicity, superimposed upon low-level endemic HTLV-II RFLP type a0 among older Black and White persons.

METHODS

Sample selection

Three sets of samples from individuals with confirmed HTLV-II seropositivity were studied. First, in a search for a non-Indian endemic focus of HTLV-II in the United States, we selected only those subjects with negative questionnaire responses to questions about IDU, sex with an IDU, and receipt of a blood transfusion (Schreiber et al., 1997) from among 404 HTLV-II positive, HIV seronegative blood donors enrolled in the REDS cohort from 1990 to 1992 (Murphy et al., 1997). Second, to study temporal trends of HTLV-II among IDU, we studied HTLV-II positive samples from young and old IDU enrolled in another study of San Francisco emergency room patients in 1993-1994, selected without regard to HIV status. Finally, we also studied a small number of specimens from American Indian blood donors collected by the REDS study in 1991-1994. Risk factor information was not available from the last group because they were not enrolled into the prospective cohort. Either Ficoll-separated peripheral blood mononuclear cells or white blood cell pellets were obtained from frozen repositories, and demographic and risk factor data were obtained from the source studies. The study protocol was approved by the human subjects committee of the University of California San Francisco.

PCR

Cellular samples equivalent to 0.5 ml of whole blood for the WBC pellet specimens or 15 million cells for the PBMC specimens were treated with a saponin lysis sample preparation kit (Roche Molecular Systems, Nutley, NJ). Nested PCR of the HTLV-II LTR region was then performed on lysed sample according to the method of Switzer et al. (1995). Initial primers BSQF6 and BSDR3 amplified a 712-bp LTR fragment, followed by the nested primers BSQF2 and BSDR4 to amplify a 672-bp product. The initial PCR mixture consisted of 0.2 mM dNTP (Pharmacia, Uppsala, Sweden), 10× buffer (Perkin-Elmer-Cetus, Gaithersburg, MD), each primer at 10 μM, 2.5 units of Taq polymerase (Perkin-Elmer-Cetus), and water in a total volume of 100 μl. No magnesium was added because of preexisting magnesium in the lysis buffer. The nested PCR mixture consisted of the same reagents plus MgCl₂ (Perkin-Elmer-Cetus) at 1.25 mM final concentration. Following denaturation at 94°C for 5 min, the reaction mixture were cycled 35 times at 94°C for 1 min, 55°C for 1 min and 72°C for 1 min with an extension time of 2 s per cycle, followed by a final extension at 72°C for 7 min, using a Perkin-Elmer 4800 thermocycler. Physical separation was maintained between the sample preparation and post-PCR manipulations. The presence of a PCR product of the expected size was assessed by electrophoresis on a 2% agarose (Gibco Brl., Gaithersburg, MD) gel with ethidium bromide.

RFLP

RFLP analysis was performed on all specimens with positive PCR results using the enzymes DraI, XhoI, AvalI, SauI (also called AocI) (Boehringer, Mannheim, Germany), and BanI, Eco47III, and SacII (Promega, Madi-
son, WI). The only difference from the RFLP method of Switzer and Heneine (Switzer et al., 1995) was that BglII was not utilized because this enzyme did not add to the discrimination provided by the above enzymes. The nested PCR product was not purified prior to RFLP. Reaction conditions included 2 to 4 μl of nested PCR product, 5±10 units of enzyme plus 10× enzyme-specific buffer provided by manufacturer, at a final concentration of 1×. Incubation was for 2 h at 37°C, and reaction products were visualized either on 2% agarose or 1% agarose plus 2% NuSieve low melting point agarose (FMC Bio Products, Rockland, ME) with etidium bromide.

Cloning and sequencing

Twenty-five microliters of nested HTLV-II LTR PCR product from selected samples was subjected to 2% agarose gel electrophoresis and the resulting 672 bp band was excised and purified using a Gene-clean kit (Bio 101, La Jolla, CA). The HTLV-II DNA was bluntly ligated into ampicillin-resistant pCR Script amp SK+ plasmid (Stratagene, La Jolla CA) and then transfected into HB101 Escherichia coli (Gibco BrL) using a heat shock methodology and plated onto ampicillin media. Positive colonies were selected by filter hybridization to HTLV-II LTR probes 27-LTRII and 481RM (Maucere et al., 1995) labeled with 32P (Amersham, Little Chalfont, Buckinghamshire, UK). Plasmid DNA was purified using Qiagen midiprep kits (Qiagen, Milden, Germany). DNA sequencing was done on one clone from each sample using the chain termination method with the Sequenase reaction (USB, Beverly, MA), universal primers T3 and T7 (Joseph Felsenstein, joe@genetics.washington.edu). The SEQBOOT program was used to generate 200 data sets that are randomly resampled versions of the previously aligned sequences. For both the maximum parsimony and neighbor-joining methods, a consensus tree was constructed using the CONSENSE program with the majority rule` criteria. Bootstrap values given on the trees were percentage values based upon 200 bootstrap iterations. In these trees, the tree branch lengths have no particular meaning. To remedy this, we performed a final analysis using the molecular evolutionary genetic analysis (MEGA Version 101) program (S. Kumar, K. Tamura, and M. Nei, Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, PA) to construct a consensus tree obtained by the neighbor-joining method (using Kimura two-parameter distance and 1000 bootstrap iterations), with estimated branch lengths.

Epidemiologic and RFLP type data were processed using SAS software (Statistical Analysis System, Cary, NC). Univariate analyses were performed using 2 × 2 tables, and associations were tested using the Fishers exact test and odds ratios (OR) with exact 95% confidence intervals. Multivariate models were analyzed using the SAS Proc Logist procedure.

ACKNOWLEDGMENTS

We are grateful to W. Switzer and W. Heneine for technical advice and to M. Shah, N. Murgai, K. Young and the staff of the Retrovirus Epidemiology Donor Study (REDS) study who collected the specimens and data used in this study. Dr. Murphy was a Fulbright Scholar and an NIH Fogarty Senior International Fellow during the period of this research. This work was also partially supported by a CANAM (France) grant to R. Mahieux, a grant from the Agence Nationale de Recherche sur le SIDA to A. Gessain, and the U.S. National Heart Lung and Blood Institute Research Contracts N01-HB-97077 (superseded by -47114), -97078, -97079, -97080, -97081, and -97082.

REFERENCES