Detection of a Major Gene Predisposing to Human T Lymphotropic Virus Type I Infection in Children among an Endemic Population of African Origin

S. Plancoulaine,1,2 A. Gessain,2,* M. Joubert,4 P. Tortevoye,2 I. Jeanne,3 A. Talarmin,3 G. de Thé,2,* and L. Abel1,3

Human T lymphotropic virus type I (HTLV-I), the first human oncoretrovirus to be discovered [1], causes a lymphoproliferative malignancy of CD4-activated cells called adult T cell leukemia/lymphoma (ATLL) and a chronic myelopathy called tropical spastic paraparesis/HTLV-I–associated myelopathy (TSP/HAM) [2]. HTLV-I infection has also been associated with clusters of high endemicity in certain geographic areas or ethnic groups. HTLV-I antibody prevalence rates, based on strict diagnostic criteria (Western blot and/or specific immunofluorescence confirmation), may range from 0.01% to 15% among adults in general population [5]. Areas with HTLV-I seroprevalence >2% in adults include the southwestern islands of Japan, the Caribbean basin, parts of South America, tropical Africa, parts of the Middle East (Iran), and Melanesia [5].

Three modes of transmission are recognized for HTLV-I. Mother-to-child transmission varies from 10% to 20% and is thought to occur after the decline of protective IgG maternal antibodies (around 6–9 months of age) through breast-feeding, by ingestion of maternal lymphocytes containing the HTLV-I provirus [6]. HTLV-I is also transmitted by sexual intercourse, mainly from male to female [7–9]. This difference may account for the higher age-specific HTLV-I seroprevalence in women than in men [5]. The intravenous route of infection, either by needle sharing or by blood transfusion, appears to be the most efficient mode of transmission, with a 15%–60% risk of infection for those receiving HTLV-I–infected cellular blood products [10]. Sexual and blood transfusion had been specifically linked to TSP/HAM, whereas ATLL appears to be more common after childhood infection [11, 12].

For an individual exposed to HTLV-I infection, the factors leading to the acquisition of the virus and/or to an anti HTLV-I immune humoral response (HTLV-I seropositive status) are unknown. They could be viral (e.g., infectious dose), environmental (e.g., age of infection), and/or genetic factors [5, 13–15].
Several arguments strongly suggest the existence of host genetic factors involved in HTLV-I infection, as well as in the development of HTLV-I-associated diseases among infected individuals. Familial aggregation of HTLV-I-seropositive individuals in high endemic areas (Japan, Caribbean and South America) have been reported in many studies [5, 9, 15-18], raising the problem of the distinction between shared familial exposure factors and genetic factors. Furthermore, mother-to-child transmission of HTLV-I occurs in only 10%-20% of the children born from infected mothers, despite a frequently similar exposure to HTLV-I infection [19, 20]. HTLV-I-associate pathologies also exhibit familial aggregation with familial clustering of ATLL [21, 22] and TSP/HAM [23, 24]. Furthermore, several association studies showed that persons who develop an ATLL or a TSP/HAM do not have the same HLA distribution as HTLV-I healthy carriers in Japanese [25-27] and black [28] populations. More recently, the class I allele HLA-A*02 was shown to be strongly protective against TSP/HAM in a Japanese population, whereas HLA-DRB1*0101 was found to increase the risk of TSP/HAM in the absence of HLA-A*02 [29]. Experimental studies in rats also showed difference in susceptibility to HTLV-I infection according to the genetic background of the animals [30]—Fisher 344 and Brown Norway inbred strains are more permissive to HTLV-I infection than are Lewis and BB strains [31, 32]. Furthermore, after infection by HTLV-I, some strains (WKA) seem to develop neurological diseases more frequently than do others [33, 34].

Since 1994, we have been conducting a large epidemiological study to determine the risk factors influencing HTLV-I infection in a general and highly endemic population of French Guiana [35]. In this population, the overall HTLV-I seroprevalence was found to be 8% (95% confidence interval 6.5%-9.5%). We have now completed familial data, and the goal of the present study was to investigate, by means of segregation analysis, the genetic factors involved in the control of HTLV-I infection. Results are consistent with the presence of a dominant major gene predisposing to HTLV-I infection in children.

Subjects and Methods

Study population. A large epidemiological study was conducted from November 1994 to November 1998 in 2 isolated villages of French Guiana, Maripasoula and Papalitchon (2000 and 700 inhabitants, respectively), which are located in the Amazonian rain forest of northeastern South America between Brazil and Suriname. Practical details have been provided elsewhere [35] and will be briefly summarized. Most subjects of these villages (80%) belong to an ethnic group referred to as Noir-Marron, who are descendants of Africans who escaped from plantations in Suriname in the 18th century. Our first epidemiological study, conducted to determine the risk factors influencing HTLV-I infection in a general population and limited to the village of Maripasoula, showed high HTLV-I seroprevalence in this ethnic group [35]. For the present analysis, all consenting Noir-Marrons ≥2 years of age (1638 subjects) living in the 2 villages were included. Demographic and medical data were collected by interview and/or from medical files. Information on familial relationships was obtained on the basis of several crossed interviews (i.e., the same information was obtained by different interviewers), and the validity of genealogical data was checked with the local medical team and the population.

Seroanalytical assays. A 10-mL blood sample was taken from 1238 of the 1638 subjects included for HTLV-I determination, and drops of blood were put on filter paper for the 400 remaining individuals, mainly children. Plasma samples were screened by ELISA (Cobas Core, anti-HTLV-I/II EIA; Roche, Basel, Switzerland). All samples were also screened by immunofluorescence (IFA) on HTLV-I-producing MT2 cells. All positive or borderline ELISA or IFA samples were tested by a confirmatory Western blot (Western blot HTLV2.3; Diagnostic Biotech, Singapore) with stringent positive criteria [35].

Segregation analysis. The phenotype of interest was a binary trait (i.e., HTLV-I seropositive/seronegative status denoted as HTLV-I+/HTLV-I−). Segregation analysis was performed using the regressive logistic model [36], which specified a regression relationship between the probability of a person to be infected (i.e., to be HTLV-I+) and a set of explanatory variables including major genotype, phenotype of preceding individuals (or antecedents), and other covariates (or risk factors). The use of the regressive model allows us to analyze large pedigrees as a whole, to estimate simultaneously the genetic and risk factors (age, sex) effects, and to consider a pattern of familial correlations on the HTLV-I status, including dependence between spouses.

The parameters of the major gene are (1) q, the frequency of allele A predisposing to be HTLV-I infected, and (2) aAA, aAa, and aaa, the 3 baseline risks of being HTLV-I+ on the logit scale corresponding to the 3 genotypes, AA, Aa, and aa. To ensure that the parent-offspring transmission of the major gene is Mendelian, 3 additional parameters are defined tAA, tAa, and taa, which denote the probabilities of transmitting A for individuals AA, Aa, and aa, respectively [37]. These parameters are used to test the hypothesis of Mendelian transmission (tAA = 1, tAa = 0, and taa = 0) against alternative hypotheses such as general transmission (free τ’s) and no parent-offspring transmission (equal τ’s).

The dependence on relative phenotypes is parameterized through regression coefficients accounting for the relationship between the phenotype of an individual and the phenotype of his preceding relatives. In the class D pattern of familial dependence [36] used in the present analysis, 4 types of phenotypic familial dependences are considered: father-mother (spouse-spouse), father-offspring, mother-offspring, and sib-sib, with corresponding regression coefficients denoted as GFM, GFO, GMO, and GSS, respectively. To account for unknown phenotypes, each of the G parameters is a vector of 2 coefficients [38]: GFM = (γF, γM), GFO = (γF, γO), GMO = (γMO, γMO), and GSS = (γSS, γSS). For example, the logit of being HTLV-I+ for an individual is modified by γF or γO remains unchanged when his mother is HTLV-I−, HTLV-I+, or unobserved, respectively.

The effects of covariates are specified through regression coefficients that may be genotype dependent. The covariates of which effects on the phenotype were evaluated in the present analysis were sex (coded 0 for male and 1 for female) and age in years. Different functions of age were assessed (polynomial in age and...
Figure 1. Age- and sex-dependent human T lymphotropic virus type I seroprevalence rates in 1638 Noir-Marrons (878 women and 760 men) from the villages of Maripasoula and Papaïchon, French Guiana. Error bars, 95% confidence intervals of the seroprevalence rates.

logarithm of age), and the best-fitting function, based on Akaike’s (1974) information criterion (AIC), was a regression on log(age). When sex and age were taken into account, other potential risk factors measured in this population (hospitalization history, transfusion, and gravidity and parity in women) had no significant influence on HTLV-I infection, as shown elsewhere [35]. Furthermore, interactions between age or sex and genotype were never found to be significant, and the results of the present analysis are presented with two parameters, \( b_{sex} \) and \( b_{age} \), which are the regression coefficients associated with sex and log(age), respectively.

Preliminary analysis of risk factors was conducted with Statistical Analysis System software (SAS Institute, Cary, NC). Segregation analysis was performed using the computer program REGRESS [39], which incorporates the regressive approach into the LINKAGE package [40]. No ascertainment correction was needed for likelihood computation, because all families of the village were included in the analysis. Nested models were compared by means of the likelihood-ratio test. As usual, all tests were performed using a type I error of 0.05. Models were also compared by use of the AIC [41], which gives the best-fitting model as the one with minimum AIC, defined as \( \text{AIC} = -2 \log(\text{likelihood}) + 2 \times \text{no. estimated parameters} \).

Results

Description of the raw data. The entire sample included 1638 Noir-Marrons (878 women and 760 men) ranging in age from 2 to 91 years old and representing ~80% of the total Noir-Marron population of the 2 villages. The overall HTLV-I seroprevalence was 10.1% (165 HTLV-I+ /1638). HTLV-I seroprevalence was significantly higher among women than among men (13.2% vs. 6.4%; \( P < .001 \)). Although HTLV-I seropositive persons were found in every age group, there was a sharp increase of HTLV-I seroprevalence with age (as shown in figure 1), ranging from 1.9% in children <10 years of age to 47.7% for subjects >49 years of age (\( \chi^2 \) trend, \( P = .001 \)).

The 1638 Noir-Marrons belonged to 81 families distributed among 21 nuclear families (i.e., including 2 parents and their offspring), 33 families with <30 individuals, 22 pedigrees between 30 and 100 individuals, and 5 large pedigrees with >100 individuals. Overall, these families include 224 fathers (21% HTLV-I+), 404 mothers (26% HTLV-I+), 743 sons (6% HTLV-I+), and 828 daughters (13% HTLV-I+). The total number is >1638 because some individuals can be both parent and offspring, owing to the complex structure of pedigrees. Figure 2 displays an example of pedigree that includes a total of 27 subjects with known HTLV-I status (providing 5 fathers, 7 mothers, 4 sons, and 17 daughters).

Segregation analysis results. Results are shown in table 1. Familial dependences were first studied one by one. There was strong evidence for a father-mother (FM; model I vs. IIa, \( \chi^2 \) with 2 df [\( \chi^2 = 17.72; P < 2 \times 10^{-4} \)]) and a mother-offspring (MO) dependence (I vs. IIc, \( \chi^2 = 26.91; P < 10^{-4} \)). When there is only 1 estimated familial dependence, the corresponding \( \Gamma \) parameter is interpretable in terms of conditional odds ratio [42]; that is, the FM odds ratio, OR_{FM}, is equal to
Table 1. Segregation analysis of human T lymphocyte virus type I (HTLV-I) serological status in 81 Noir-Marron families.

<table>
<thead>
<tr>
<th>Model and hypothesis</th>
<th>$q^2$</th>
<th>$a_{aa}^b$</th>
<th>$a_{AA}$</th>
<th>$g_{FM1}$</th>
<th>$g_{FO1}$</th>
<th>$g_{MO1}$</th>
<th>$g_{SS1}$</th>
<th>$g_{FM2}$</th>
<th>$g_{FO2}$</th>
<th>$g_{MO2}$</th>
<th>$g_{SS2}$</th>
<th>$\beta_{aa}d$</th>
<th>$\beta_{AA}^e$</th>
<th>$\tau_{AAA}$</th>
<th>$\tau_{AA}$</th>
<th>$\tau_{A}$</th>
<th>$-2\ln L + cf$</th>
<th>AIC$g^f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Sporadic</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>0.64</td>
<td>1.91</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>82.39</td>
<td>63.84</td>
</tr>
<tr>
<td>II. Familial dependence</td>
<td>a. FM</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>0.66</td>
<td>1.90</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>64.67</td>
<td>50.67</td>
</tr>
<tr>
<td></td>
<td>b. FO</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>0.64</td>
<td>1.84</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>79.53</td>
<td>65.53</td>
</tr>
<tr>
<td></td>
<td>c. MO</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>0.68</td>
<td>1.80</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>55.48</td>
<td>41.48</td>
</tr>
<tr>
<td></td>
<td>d. SS</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>0.65</td>
<td>1.83</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>76.42</td>
<td>62.42</td>
</tr>
<tr>
<td></td>
<td>e. FM</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>0.69</td>
<td>1.81</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>37.38</td>
<td>27.38</td>
</tr>
<tr>
<td></td>
<td>f. FM</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>[0]</td>
<td>0.67</td>
<td>1.74</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>31.29</td>
<td>25.29</td>
</tr>
<tr>
<td>III. Mendelian major gene</td>
<td>a. Codominant</td>
<td>1. Residual FM + MO</td>
<td>0.0076</td>
<td>−11.51</td>
<td>−3.84</td>
<td>−0.69</td>
<td>−0.65</td>
<td>[0]</td>
<td>0</td>
<td>−0.38</td>
<td>[0]</td>
<td>[0]</td>
<td>0.52</td>
<td>[0]</td>
<td>0.82</td>
<td>2.40</td>
<td>[1]</td>
<td>[5]</td>
</tr>
<tr>
<td></td>
<td>b. Dominant</td>
<td>1. Residual FM + MO</td>
<td>0.0089</td>
<td>−12.59</td>
<td>−3.77</td>
<td>−0.88</td>
<td>−0.85</td>
<td>[0]</td>
<td>0</td>
<td>−0.43</td>
<td>0.17</td>
<td>0.98</td>
<td>[0]</td>
<td>0.49</td>
<td>0.17</td>
<td>0.96</td>
<td>2.62</td>
<td>[1]</td>
</tr>
<tr>
<td></td>
<td>c. Recessive with FM + MO</td>
<td>0.018</td>
<td>−10.69</td>
<td>−3.84</td>
<td>−0.65</td>
<td>[0]</td>
<td>−0.38</td>
<td>[0]</td>
<td>1.09</td>
<td>[0]</td>
<td>0.52</td>
<td>[0]</td>
<td>0.82</td>
<td>2.40</td>
<td>[1]</td>
<td>[5]</td>
<td>[0]</td>
<td>9.25</td>
</tr>
<tr>
<td>IV. Nontransmitted dominant effect, FM + MO</td>
<td>0.001</td>
<td>−10.00</td>
<td>−9.26</td>
<td>−0.51</td>
<td>[0]</td>
<td>−0.91</td>
<td>[0]</td>
<td>0.38</td>
<td>[0]</td>
<td>0.68</td>
<td>1.98</td>
<td>[92]</td>
<td>[92]</td>
<td>992</td>
<td>992</td>
<td>992</td>
<td>992</td>
<td>28.62</td>
</tr>
<tr>
<td>V. General transmitted dominant effect, FM + MO</td>
<td>0.001</td>
<td>−10.00</td>
<td>−4.19</td>
<td>−0.57</td>
<td>[0]</td>
<td>−0.72</td>
<td>[0]</td>
<td>1.16</td>
<td>[0]</td>
<td>0.39</td>
<td>[0]</td>
<td>0.62</td>
<td>2.10</td>
<td>0.84</td>
<td>[0]</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

NOTE. All parameters in brackets are fixed at the value shown. FM, father-mother; FO, father-offspring; MO, mother-offspring; SS, sib-sib.

a Frequency of allele A predisposing to be HTLV-I infected.
b Baseline risks of being HTLV-I on logit scale corresponding to 3 genotypes: AA ($g_{AA}$), Aa ($g_{Aa}$), and aa ($g_{aa}$).
c Regression coefficients associated with the following familial dependences: father-mother ($g_{FM1}$, $g_{FM2}$), father-offspring ($g_{FO1}$, $g_{FO2}$), mother-offspring ($g_{MO1}$, $g_{MO2}$), and sib-sib ($g_{SS1}$, $g_{SS2}$).
d Regression coefficients associated with sex ($b_{sex}$) and age ($b_{age}$).
e Probabilities of transmitting A for individuals AA ($t_{AAA}$), Aa ($t_{Aa}$), and aa ($t_{aa}$).
f C represents twice the log likelihood ($2\ln L$) of the general model V ($\approx 928.16$).
g Akaike's information criterion (AIC) with baseline taken as AIC value of general model V ($\approx 952.16$).
There was no evidence for a father-offspring (FO) correlation (I vs. IIb, \( \chi^2 = 2.86; P > .2 \)) and the \( \Gamma_{1O} \) parameter was fixed at (0, 0) for further analyses. The sib-sib (SS) dependence was borderline significant either when tested against a sporadic model (I vs. IIId, \( \chi^2 = 5.97; P < .05 \)) or in the presence of FM and MO dependences (IIe vs. IIf, \( \chi^2 = 6.09; P = .05 \)). Therefore, FM, MO, and SS dependences were conserved for further analyses.

In the presence of FM, MO, and SS dependences, there was strong evidence for a codominant major gene (IIf vs. IIIa2, \( \chi^2 = 23.78; P < 10^{-4} \)). In presence of this codominant major gene, there was no evidence for any residual SS dependence (IIIa1 vs. IIIa2, \( \chi^2 = 1.73; P > .4 \); \( \chi^2_{SS2} = \chi^2_{SS1} \approx 0 \)), and \( \Gamma_{ss} \) was further fixed at (0, 0). With respect to the degree of dominance, a recessive model was rejected (IIic vs. IIIa1, 7.23; \( P < .01 \)), and the best fit was obtained with a dominant major gene (model IIIb1). In presence of this dominant major gene, the FM correlation was still highly significant (IIIb2 vs. IIIb1, 16.09; \( P < 10^{-4} \)), whereas the MO dependence was much less significant (IIIb3 vs. IIIb1, 8.92; \( P < .02 \)) than without the major gene (IIa vs. IIe, \( \chi^2 = 27.29; P < 10^{-4} \)). No gene-by-sex or gene-by-age interaction was found (data not shown). In the presence of FM and MO correlations, the Mendelian transmission of the dominant major effect was rejected (IIIb1 vs. V, \( \chi^2 = 9.25; P < .03 \)), and the hypothesis of no transmission of the dominant major effect was also rejected (IV vs. V, \( \chi^2 = 28.62; P < 10^{-4} \)).

To further investigate the observed rejection of the Mendelian transmission hypothesis, we conducted a simulation study. This study was motivated by (1) the observation that the Mendelian transmission hypothesis was not strongly rejected with a frequency of allele A tending toward 0 under the general transmission hypothesis and (2) the characteristics of the Mendelian model IIIb1 detailed in the next paragraph and predicting that a very high proportion of adult HTLV-I+ cases is sporadic. The aim of the simulation study was to test whether analysis of familial data generated under such a genetic model can lead to rejection of the Mendelian hypothesis more often than expected by a given type I error level. More specifically, we generated within our family sample structure and under Mendelian model IIIb1, shown in table 1, 50 replicates of phenotypic data (keeping unknown the phenotypes of subjects who were actually missing) that provided evidence for a dominant major gene. We observed that, of the 50 replicates, 36 led to the false rejection of the Mendelian transmission hypothesis using a 0.05 type I error (the probability that these 36 rejections are due only to type I error is \( \approx 10^{-9} \)). Furthermore, when we used the 50 \( \chi^2 \) values observed in the simulation study to compute an empirical significance level for our actual value of 9.25, we obtained a Monte Carlo \( P \) value >.60. These results indicate that, under this specific genetic model, the asymptotic distribution could not be used for testing the Mendelian hypothesis against the general transmission hypothesis. When an appropriate empirical distribution for this test is used, our data are consistent with the presence of a dominant major gene, with FM and MO residual familial dependences accounted for the
familial distribution of HTLV-I+/HTLV-I− status among the Noir-Marron population.

Under the dominant major gene model IIIb1, the frequency of the allele A was estimated as .0076, indicating that ~1.5% of the individuals are predisposed to HTLV-I infection. The penetrance—that is, the probability of being HTLV-I+—in children (0–15 years of age) according to age, sex, and genotype is shown in figure 3A. The results are presented considering an HTLV-I+ mother, because the major route of HTLV-I transmission in children is through breast-feeding. The penetrance is almost complete by age 10 years for Aa (or AA) children, whereas it remains close to 0 in aa children by the same age. With respect to these results, the variation of penetrance with age in subjects older than age 15 years is shown only for the aa genotype, according to sex and spouse’s HTLV-I status (figure 3B). By age 40, the probability to be HTLV-I+ is estimated at .09 (.17) for a man (woman) with an HTLV-I− spouse. Overall, these data clearly indicate that almost all HTLV-I-seropositive children <10 years old (who have acquired infection through breast-feeding) are genetic cases, whereas most of HTLV-I seropositive adults are sporadic cases.

Discussion

The investigation of host genes controlling susceptibility/resistance to viral agents in humans is a recent research domain. So far, genetic studies in HTLV-I infection have focused on testing the association between HTLV-I related diseases and HLA genes and found that persons who develop an ATLL or a TSP/HAM do not have the same HLA distribution as HTLV-I healthy carriers [25–29]. In particular, the more recent study [29] showed that the class I allele HLA-A*02 was strongly protective against TSP/HAM, whereas HLA-DRB1*0101 was
found to increase the risk of TSP/HAM in the absence of HLA-A*02; however, no studies have been performed to identify genetic factors controlling HTLV-I infection per se, as measured by the HTLV-I+HTLV-I− status assessing the presence of persistent infection with HTLV-I provirus in the peripheral blood lymphocytes. It should be noted that, by use of such a serological phenotype, major genetic findings have been obtained in another retrovirus infection, with the demonstration that human immunodeficiency virus type 1 (HIV-1)−seropositive/seronegative status was associated with the gene encoding the CC-chemokine receptor 5 (CCR5)—that is, individuals homozygous for a defective CCR5 allele were highly protected against HIV-1 infection [43–45]. The present study, which indicates that a dominant gene is involved in the control of HTLV-I infection in a highly endemic population from African origin, is a major step in the understanding of the mechanisms leading to chronic HTLV-I infection.

Our segregation analysis results are consistent with the presence of a dominant major gene predisposing to HTLV-I infection with residual familial dependences. It was interesting to note that a large part of the MO dependence, as well as the weak SS correlation observed in nongenetic models, appeared to be taken into account by the major gene effect. In contrast, the FM dependence remained stable regardless of whether the major gene was included in the model and is well explained by the transmission route of HTLV-I from men to women through sexual intercourse. The test of the Mendelian transmission hypothesis for the dominant major gene detected in the present study raised some questions. Although the hypothesis of Mendelian transmission was first rejected using asymptotic distribution (asymptotic $P < .03$), simulation studies clearly indicated that this asymptotic distribution could not be used for testing this hypothesis and that our familial data were consistent with Mendelian transmission of a dominant major gene when using an appropriate empirical distribution (Monte Carlo $P > .60$). This latter result is likely to be explained by the specificities of the present dominant major gene model predicting that most of HTLV-I seropositive adults are sporadic cases. In this situation, the estimation of $q$, the frequency of allele A, became very difficult under the general transmission model, because HTLV-I seropositive founders are very likely to be sporadic cases (because of free transmission probabilities, founder genotypes can be relatively independent of those of their descending relatives). Therefore, it is not surprising to obtain an estimation of $q$ tending toward 0 and a distortion of transmission probabilities to restore the presence of allele A in the families with seropositive children. Under these particular conditions, the simulation study clearly showed that an empirical distribution should be used to properly assess the significance level of the Mendelian transmission test.

A major characteristic of the gene detected by the present segregation analysis is that almost all HTLV-I−seropositive children (<10 years old) are predicted to be genetic cases. Therefore, this gene appears to account for most infections occurring in children through breast-feeding and can explain, at least in part, the reason why mother-to-child transmission of HTLV-I only occurs in a certain proportion of children fed by infected mothers. Because ATLL has been shown to be highly associated with childhood infection [11, 13, 46] and to exhibit familial aggregation [5], the investigation of the role of this dominant gene in ATLL development will be of major interest. Linkage studies with genetic markers are now ongoing to locate the major gene detected by the present analysis.

Acknowledgments

We thank the medical team of Maripasoula and Papêchéton; the nurses and interpreters; and Dr. B. Cottrelle and P. Chaud, from the Direction Départementale des Affaires Sanitaires et Sociales de French Guiana. We are also grateful to J. F. Pouliquen and M. Van Beveren, for the serological assays; R.P. Buigues, for his epidemiological field work; Prof. J. P. Moreau and Dr. J. L. Sarthou, from the Institut Pasteur of French Guiana, for their continuous active support; and F. Demenais, for providing us the REGRESS software.

References


