Frequent mutation in pX region of HTLV-1 is observed in HAM/TSP patients, but is not specifically associated with the central nervous system lesions

Mineki Saito1,2, Yoshitaka Furukawa1,2, Ryuji Kubota2, Kouichiro Usuku2, Shunro Sonoda3, Shuji Izumo2, Mitsuhiro Osame2 and Mitsuaki Yoshida1

1 Department of Cellular and Molecular Biology, The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108, Japan; 2 Third Department of Internal Medicine and 3 Department of Virology, Kagoshima University, 8-35-1 Sakuragaoka, Kagoshima 890, Japan

Human T-cell leukemia virus type 1 (HTLV-1) is an etiologic agent of adult T-cell leukemia (ATL) and of HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP). Recently it has been reported that defective HTLV-1 provirus was detected frequently in the central nervous system (CNS) lesions of HAM/TSP patients. Here we investigated sequence variations of the pX region of HTLV-1 in the CNS and peripheral blood lymphocytes (PBL) of the same patient. The results analyzing 9–13 clones isolated from each specimen indicated that the pX region is highly variable within a patient with HAM/TSP, and the mutations were found at almost random positions within the sequences analyzed. The frequency and pattern of those mutations did not appear to differ significantly between the CNS and PBL of the same patient, although they differed among patients. Similarly, frequent mutations were observed in an asymptomatic carrier of HTLV-1, although the variability was moderate, suggesting that the high variability of the pX sequence is not a specific event in HAM/TSP. However, one asymptomatic carrier showed much less frequent variations very similarly to an ATL patient; both of them harbored clonally expanded infected cells. Thus the apparent low variability was explained by clonal selection of a single species of the provirus by the clonal proliferation of infected cells. These results clearly indicate that mutations including defectives are not specifically associated with the CNS lesions in HAM/TSP patients, but suggest that the random mutations simply reflect the rate of viral replication in individuals and the variants were not inherited frequently.

Keywords: HAM/TSP; HTLV-1 instability; intrastrain variability; HTLV-1 in CNS

Introduction

A group of chronic progressive inflammatory diseases of the central nervous system (CNS) is established as HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP) (Gessain et al., 1983; Osame et al., 1986). The neuronal disease is associated with infection by human T-cell leukemia virus type 1 (HTLV-1) (Poiesz et al., 1980; Yoshida et al., 1982) which has been isolated as a causative agent of adult T-cell leukemia (ATL) (Yoshida et al., 1984). Many isolates of HTLV-1 from HAM/TSP patients have been analyzed and compared with those from ATL patients (Tsujimoto et al., 1988; Evangelista et al., 1990; Shirabe et al., 1990). HTLV-1 infection into cells of the CNS has not been established in vivo, however, infection into T cells infiltrated into the CNS has been described (Hara et al., 1994; Kubota et al., 1994). Analyzing the proviral sequences, microvariations of the sequence have been detected in almost all regions of the proviral genome; however, none of them was found to be specific to HAM/TSP (Daenke et al., 1990; Komurian et al., 1991; Kinoshita et al., 1991). Therefore, it has been thought that the same species of HTLV-1 is involved in both HAM/TSP and ATL (Yoshida et al., 1987). Although it was suspected that an immune mechanism affected by HTLV-1 infection is somehow involved in the development of HAM/TSP, little is known about the mechanism of HAM/TSP development.

Recently, Niewiesk et al. (1994) reported that the genomic sequence of HTLV-1 pX is more variable than previously expected and the sequence variabil-
ity within an infected individual is much greater than the variability between infected persons. Furthermore, they reported that the amino acid sequence of HTLV-1 Tax protein encoded by the pX region is more variable in asymptomatic carriers than in HAM/TSP patients, when the virus is isolated from peripheral blood lymphocytes (PBL), suggesting greater conservation of the Tax protein in HAM/TSP than in carriers. In contrast to these observations, when the proviruses in the CNS of HAM/TSP patients were analyzed, highly frequent (26%) detection of defective proviruses in the Tax protein has been reported, suggesting a specific contribution of the Tax-defective mutants in the development of HAM/TSP (Kira et al., 1994). However, in these two reports, HTLV-1 genomes were analyzed either in PBL or CNS of each patient; thus, direct comparison between CNS and PBL in the same individuals was not possible.

The pX region has three or four open reading frames (ORF) overlapping each other; thus, the sequence was found to be conserved (Kiyokawa et al., 1985; Tsujiimoto et al., 1988; Shirabe et al., 1990). One of the ORFs codes for the Tax protein, which is essential for efficient viral expression (Sodroski et al., 1984; Fujisawa et al., 1985) and also for activation of many cellular genes including cytokine genes and proto-oncogenes (Inoue et al., 1986; Cross et al., 1987; Fuji et al., 1988 and 1991; Miyatake et al., 1998). Furthermore, Tax is the dominant target antigen of the cytotoxic T-cell response to HTLV-1 in HAM/TSP patients. Therefore, involvement of the Tax protein in the virally induced pathogenesis has been proposed (Jacobson et al., 1990; Elovaa et al. 1993). We were thus interested in the viral sequence variability in CNS and PBL of the same infected individuals.

In this paper, we amplified by polymerase chain reaction (PCR) the nucleotide sequences of HTLV-1 pX region from freshly frozen CNS tissues and PBL obtained from four autopsied patients with HAM/TSP, and the sequences were compared within individual patients. The results indicated that the pX sequence is highly variable in HAM/TSP, but the variations were random and no difference was observed between CNS and PBL. In contrast to previous reports, it was suggested that the sequence variations of the pX region are correlated with viral replication, but not with specific diseases.

**Results**

**DNA sequences of HTLV-1 pX region in HAM/TSP patients**

To determine the nucleotide sequence of the pX region of HTLV-1 in CNS and PBL, high molecular weight DNA was isolated from both tissues from the same patients (case 1, 2, 3 and 5) with HAM/TSP, and the Tax-coding sequence in the pX region was amplified by PCR. The primers used for the amplification are indicated in Figure 1. The sequence amplified here covered most of the Tax coding region as illustrated in Figure 1 and included the region where most of the mutations were detected in the previous reports by Niewiesk et al. (1994) and Kira et al. (1994).

In all cases of DNA preparations from PBL, the amplified DNA products were detected after the first round of PCR; however, in CNS DNA preparations, the second round of PCR was required to obtain sufficient materials for subsequent analysis. The DNA amplified by the first round of PCR for PBL DNA and that by the second round of PCR for CNS DNA were subcloned into Bluescript. This difference in PCR did not affect the results as discussed later. Nine to 13 clones were randomly picked up from each specimen and sequenced.

Each sequence determined was compared with the sequence of AKT-1, a prototype sequence of HTLV-1, and only the nucleotides that differed from AKT-1 are summarized in Figure 2. Clones marked with gray shading are the dominant species in each specimen. In all cases, the dominant sequences were mostly close to AKT-1 and highly conserved the amino acid sequence of AKT-1 (Seiki et al., 1983) except for HAM3 (8322 C to T induces P to L). The sites of mutations were positions 15 (case HAM3) to 30 (case HAM5) in the sequence of 960 positions determined. It should be noted that two thirds of mutations induced alterations of amino acids in the Tax protein (Table 1). Generally the genetic codes consisted of three bases and one of the three bases is synonymous; thus, the fact that two thirds of mutations induced amino acid changes clearly indicates that the mutations were completely random in the Tax-coding sequence. More directly, when all mutations were plotted on the Tax-coding sequence (Figure 3), the mutations
Figure 2 Nucleotide changes in HTLV-1 clones isolated from PBL and CNS of HAM/TSP patients (a–d), of ATL patient (f) and of asymptomatic carriers (g and h). Only altered nucleotides are indicated. +C or -C at the position 7784 represents insertion or deletion of one C residue in the CCCCCCCC stretch. The numbers at the top of the figure represent the positions of the nucleotide in reference to the ATK-1 sequence (Seiki et al., 1983). At the lower part in each figure, amino acid changes (single letter abbreviation) or stop codon (*) by corresponding mutations are presented. ‘Minus’ indicates non-coding. Upper two lines are the amino acids of the wild type Tax (Seiki et al., 1983) or Rex protein and the lower lines are those of mutants.
distributed almost randomly and no clustering at particular regions was observed. Position 8405, where one A residue was inserted in all cases, is an exception and should represent the parental clone infected originally. Surprisingly, no identical clone was detected in each specimen except the dominant
clones. The latter finding strongly suggests that variants were accidentally produced during replication, but most variants were not replicative, and thus not inherited by the progenies.

More importantly, the frequency of sequence variations was not significantly different between CNS and PBL in all four cases, including their nucleotide substitutions and dominant clones (Figure 2). Furthermore, the frequency of mutants coding for defective Tax proteins was not particularly high and almost the same between CNS and PBL; the incidence of defective mutants of the Tax protein was in three out of 41 clones of PBL and seven out of 56 clones of CNS (Table 1). These observations clearly indicated that a nucleotide substitution in the pX region including those for defective Tax protein is not associated with the CNS, where the pathological lesion is located in.

Sequence variations in ATL patients and carriers
To examine whether the results obtained from patients with HAM/TSP are specific to HAM/TSP or a general finding in infected individuals, similar analysis was carried out on a CNS sample from an ATL patient and also on samples from PBL of asymptomatic carriers (Figure 2). Tax-coding sequence in the CNS of an ATL patient showed much less variability and 10 out of 12 clones had the same sequence. ATL patients generally harbor clonally expanded leukemic T-cells which are infected with HTLV-1, and infiltration of leukemic T cells into the spinal cord was observed in this ATL patient (data not shown). Therefore, the finding in this ATL patient is explained by the presence of clonal leukemic cells in the CNS specimen. The low variability in CNS of this case provided a good evidence that two rounds of PCR on CNS DNA did not give any higher incidence of mutation comparing with that of PBL.

Analyzing asymptomatic carriers, two cases gave rather different results (Figure 2). Case AC1 showed almost identical results to that of ATL patients; that is, the sequence was stably conserved; however, case AC2 gave rather similar results to those of HAM/TSP patients, although the variations were moderate when compared with those in HAM/TSP; four clones were dominant and mutations were observed at eight positions of nucleotides. The reason for such striking difference between two asymptomatic carriers is further investigated in the following section.

---

**Table 1** Summary of mutations observed in Tax-coding region of HAM/TSP patients

<table>
<thead>
<tr>
<th>Patient</th>
<th>ATL</th>
<th>HAM-PBL</th>
<th>HAM-CNS</th>
<th>HAM-total</th>
<th>AC-PBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of clones analyzed</td>
<td>10</td>
<td>41</td>
<td>56</td>
<td>97</td>
<td>22</td>
</tr>
<tr>
<td>No. of defective mutations</td>
<td>Tax Rex</td>
<td>Tax Rex</td>
<td>Tax Rex</td>
<td>Tax Rex</td>
<td>Tax Rex</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of mutations</td>
<td>Tax Rex</td>
<td>Tax Rex</td>
<td>Tax Rex</td>
<td>Tax Rex</td>
<td>Tax Rex</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>32</td>
<td>9</td>
<td>61</td>
<td>21</td>
</tr>
<tr>
<td>93</td>
<td>30</td>
<td>11</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of amino acid alterations</td>
<td>1</td>
<td>2</td>
<td>20</td>
<td>8</td>
<td>41</td>
</tr>
<tr>
<td>61</td>
<td>26</td>
<td>10</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Figure 3** Summary of mutations in HAM/TSP patients. The positions of mutations detected in all samples from HAM/TSP patients are cumulatively plotted by vertical bars. Open box and horizontal thick line represent Tax-coding sequence and non-coding sequence, respectively. The region amplified and sequenced is also indicated. A bar with two arrowheads indicates the region of an epitope recognized by cytotoxic T cells from a HAM/TSP patient with B14 HLA haplotype (Elovaa et al, 1993).
Southern blot analysis of HTLV-1 genome

The sequence variations observed in each specimen might be affected by the population of HTLV-1-infected cells such as virus dose in vivo and clonality of infected cells. To collect such information, blot analysis of the viral DNA was carried out on DNA from PBL. The DNA was digested with PstI and the viral sequences were detected by hybridization either with the total viral probe carrying the whole HTLV-1 sequence or with the LTR probe. When sufficient cells are randomly infected, the total viral probe will give three bands and the LTR probe will not detect any band. On the other hand, for infected cells expanded clonally, the total viral probe will give four or five bands and the LTR probe will detect one or two bands of the viral sequences flanked by cellular sequences if one proviral copy is integrated. Blotting results of three HAM/TSP patients and two asymptomatic carriers are shown in Figure 4. PBL DNA of HAM2, HAM3 and HAM5 gave three bands with the total viral probe, although they were rather weak, but no significant band with the LTR probe. These results indicate that a significantly large population of cells is randomly infected in these patients, confirming the previous findings that HAM/TSP patients have a greater population of HTLV-1-infected cells than asymptomatic carriers (Yoshida et al. 1989; Furukawa et al. 1992). On the other hand, PBL DNA of one asymptomatic carrier, AC1, gave four bands with the total probe and three bands with the LTR probe as shown in Figure 4, indicating the limits carrier had two or three expanded infected cells. If one or two copies of the proviruses are integrated, one should expect two or four bands with the LTR probe, but only three bands were detected in this case. The probable explanation is that either two bands were overlapped or one fragment was very small and thus ran off the gel. The other asymptomatic carrier, AC2, did not show any significant band even with the total probe, suggesting that AC2 had a much smaller population of HTLV-1-infected cells than HAM/TSP patients, and thus the proviruses were not detected by the standard Southern blot analysis.

Discussion

The present study investigated the genetic variation of HTLV-1 in CNS and PBL from the same patients with HAM/TSP and found that the Tax-coding sequence is highly variable in HAM/TSP patients as described previously (Niewiesk et al., 1994; Kira et al., 1994); however, the incidence of defective mutation was not significantly different between CNS and PBL. Furthermore, the pattern and frequency of the mutations were very similar in the CNS and PBL. Thus, we could not confirm the previous report by Kira et al. (1994) and our observations strongly suggest that the highly frequent variations in the Tax-coding region are not associated with the pathogenicity of HAM/TSP.

Various mutations which induce defective Tax protein or amino acid alterations in the Tax protein were in fact detected in higher frequency than previously expected from the stability of the viral genome (Watanabe et al., 1984). Thus, one may speculate that some of these mutants are pathogenic in the development of HAM/TSP. However, this would be highly unlikely because these mutations in the CNS differ from one clone to another and no particular mutant had propagated preferentially. Therefore, we concluded that the variants just evolved randomly during viral replication, and present as a very rare population, and thus would not be able to explain the massive degeneration of CNS tissues.
unpublished observation), but we confirmed that some patients with HAM/TSP analyzed in this paper showed a random population of infected cells.

The fact observed in this paper that high variability of Tax-coding sequence is observed in random population of infected cells. Therefore, it is suggested that high variability of the Tax-coding sequence is associated with higher replication of HTLV-1 in vivo. HTLV-1 replicates very efficiently in HAM/TSP patients, thus providing more chances for mutations. If asymptomatic carriers have a high virus dose like HAM/TSP patients, it would be predicted that they will show high sequence variability of the viral sequence similarly to that in HAM/TSP patients. Consistent with this conclusion, AC2 who showed moderate variation of the Tax-coding sequence had much fewer copies of the provirus, and thus had much less chance of viral replication. Infection of HTLV-1 to central nervous system is not established and infiltration of infected T cells has been described (Hara et al., 1994; Kubota et al., 1994). The similarity of HTLV-1 variations between CNS and PBL strongly suggest that the major infected cells in the CNS were T cells and represent the random replication of HTLV-1 in PBL.

The pX region can code for two or three proteins using different but overlapping open reading frames. Thus, the region was mostly conserved (Tsujimoto et al., 1988; Shirabe et al., 1990). The reason why the pX region is so variable during the viral replication is not well understood. It is possible that the Tax protein may be the dominant target antigen of cytotoxic T-cells (Jacobson et al., 1992; Kannagi et al., 1992; Elowa et al., 1993; Parker et al., 1994), thus, sequence variation in the Tax protein may allow the infected cells to escape from the host immune response. If this explanation is correct, we should see selective variations at some specific sequences that are the targets for cytotoxic T cells. However, this was not the case (Figure 3). Almost all variants were different from each other, suggesting that these variations are not replicative. This was expected since Tax is essential for efficient replication (Sodroski et al., 1984; Fujisawa et al., 1985). About 0.4% of the mutations observed in the HAM/TSP patients induced amino acid alterations in the Tax protein and even one amino acid substitution destroys Tax activity (Seiki et al., 1986) or alters its transcriptional activity (Smith and Greene, 1996; Semmes and Jeans, 1992). Therefore, it is reasonable to expect that most of Tax mutants were inactive in transcriptional activation. Random mutations without any restriction indicate that mutations should not result in any particular activity of the Tax protein. Our results indicate that random mutations took place in the k-HTLV coding sequence at the half to three quarters proviruses. If mutations occur at such frequency in all region of HTLV-1 provirus (9kb long), it is easily expected that the mutants would not be replicative. Thus the results suggest that only the parental or highly conserved variants can replicate efficiently; thus, the variations are not inherited. This possibility also explains why the dominant clones are always the parental strain. Therefore, the high variability of the HTLV-1 genome observed in individuals is not contradictory to the previous conclusion that the viral genome is stably conserved among individuals, races and geographic regions. Based on these discussions, we are now analyzing the sequence variations in other regions than Tax-gene.

### Materials and methods

#### Patients

The clinical diagnosis of five patients with HAM/TSP was based on the criteria proposed by Osame et al. (Osame et al., 1987). Four patients (HAM 1, 3-5) were derived from Kagoshima and the one (HAM 2) was from Miyazaki, Japan. Both are the endemic area of HTLV-1 infection in Japan. All patients except HAM 3 had never received blood transfusion, and the clinical and pathological char-

<table>
<thead>
<tr>
<th>Patient</th>
<th>Sex</th>
<th>Age</th>
<th>Duration of illness</th>
<th>Anti-HTLV-1 antibody</th>
<th>Clonality of infected cells</th>
<th>No. of variants per 10 clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAM1</td>
<td>M</td>
<td>77</td>
<td>2yr 6m</td>
<td>× 2048 × 16</td>
<td>(NE)</td>
<td>4.4</td>
</tr>
<tr>
<td>HAM2</td>
<td>F</td>
<td>71</td>
<td>4yr 6m</td>
<td>× 16384 × 1024</td>
<td>Random</td>
<td>6.0</td>
</tr>
<tr>
<td>HAM3</td>
<td>F</td>
<td>52</td>
<td>8y</td>
<td>× 8192 × 128</td>
<td>Random</td>
<td>4.4</td>
</tr>
<tr>
<td>HAM4</td>
<td>F</td>
<td>68</td>
<td>9y</td>
<td>× 2048 × 516</td>
<td>(NE)</td>
<td>6.0</td>
</tr>
<tr>
<td>HAM5</td>
<td>M</td>
<td>73</td>
<td>10y</td>
<td>× 32763 × 1024</td>
<td>Random</td>
<td>6.9</td>
</tr>
<tr>
<td>ATL</td>
<td>M</td>
<td>83</td>
<td>1y</td>
<td>× 2048 × 32</td>
<td>Clonal</td>
<td>2.0</td>
</tr>
<tr>
<td>AC1</td>
<td>M</td>
<td>78</td>
<td>—</td>
<td>× 8192 (NE)</td>
<td>Clonal</td>
<td>1.7</td>
</tr>
<tr>
<td>AC2</td>
<td>M</td>
<td>35</td>
<td>—</td>
<td>× 256 (NE)</td>
<td>TLTD</td>
<td>6.0</td>
</tr>
</tbody>
</table>

AC = asymptomatic carrier; (NE) = not examined; Clonal = monoclonal integration; TLTD = HTLV-1 DNA was not detected by blot analysis; Random = random integration

Anti-HTLV-1 antibodies were titrated by the particle agglutination method.
acteristics of these patients were previously described (Umehara et al, 1993). Information for the present study is summarized in Table 2. Spinal cords from these five patients with HAM/TSP (HAM1-5) were obtained at autopsy and these specimens were stored in liquid nitrogen until use for DNA extraction. Peripheral blood lymphocytes (PBL) had been collected from the same patients 1–2 years prior to death and these samples had been frozen soon after their isolation. The spinal cord isolated from a patient with ATL, who showed infiltration of leukemic cells into the CNS, was also analyzed to compare with those of HAM/TSP patients.

**DNA extraction**

PBLs were obtained from heparinized venous blood by Ficoll–Hypaque gradient centrifugation. Frozen specimens of the spinal cord were directly used for DNA extraction. To avoid cross contamination, microtome blades were renewed for each sample. Genomic DNA was extracted from PBL and CNS tissues by proteinase K treatment followed by phenol extraction (Yoshida et al, 1984).

**PCR**

One microgram of DNA was amplified by 35 cycles of PCR using reaction buffer containing 50mM KCl, 10mM Tris (pH8.3), 1.5mM MgCl₂, 0.2mM deoxynucleotide triphosphate, 2.5U of Taq polymerase (Takara, Japan), and 1μM primers. Each PCR cycle consisted of denaturation at 95°C for 60 s, annealing at 56°C for 75 s, extension at 72°C for 90 s and extension of the final cycle at 72°C for 10 min. After the first PCR reaction, 5 μl aliquots of the amplified products were analyzed by electrophoresis in 2% Nusieve (FMC Bioproducts, ME) agarose gel followed by staining with ethidium bromide. When the product was not significantly detected on the gel, aliquots of the PCR reaction mixture were subjected to a further 20 cycles of second PCR.

**Cloning and sequencing**

Amplified DNA products with the expected size were excised from the agarose gel and the DNA was precipitated with ethanol after phenol extractions. The isolated DNA was subcloned into pBluescript KS+ and sequenced using the Cycle Sequencing Kit (Applied Biosystems, CA) and an automatic sequencer (373A DNA Sequencer, Applied Biosystems).

**Southern blot analysis**

Ten μg of genomic DNA (roughly equivalent to 10⁶ cells) from PBL samples was digested with PstI and subjected to Southern blot hybridization as described previously (Yoshida et al, 1989; Furukawa et al, 1992). The DNA was transferred to nylon filters and hybridized with 32p-labeled HTLV-1 DNA containing the whole viral sequence (total probe) or LTR sequence (LTR probe).

**Acknowledgements**

This work was supported in part by a grant for Special Project Research, Cancer Bioscience, from the Ministry of Education, Culture and Science of Japan.

**References**


p40' encoded by human T-cell leukemia virus type 1. EMBO J 5: 2883-2888.