MOLECULAR EPIDEMIOLOGY OF HUMAN T LYMPHOTROPIC VIRUS TYPE 1 TRANSMISSION IN OKINAWA, JAPAN

KYOJI KAKUDA, HIDEYUKI IKEMATSU, WEN LI YONG CHONG, JUN HAYASHI, AND SEIZABURO KASHIWAGI

Department of Environmental Medicine and Infectious Diseases, Internal Medicine and Surgery, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan; Department of General Medicine, Kyushu University Hospital, Fukuoka, Japan; Department of Clinical Research, Haro Doi Hospital, Fukuoka, Japan

Abstract. To clarify the route of human T lymphotropic virus type 1 (HTLV-1) transmission, we sequenced three proviral genome regions (gag, env, int) of HTLV-1 from 18 carriers in 7 families in Okinawa, Japan and compared the strains with isolates from other countries. The nucleotide substitution frequency among sequences derived from a single carrier was low; 0–0.24% in gag, 0–0.54% in env, and 0–0.34% in int. All sequences showed the closest identity to the Cosmopolitan strain, with differences of only 0–1.91%. All 8 mother/child pairs had identical nucleotide sequences. Of 3 pairs of spouses, 2 had identical sequences, with transmission probably from husband to wife. The mothers of both wives were HTLV-1-negative. The HTLV-1 sequence of the other wife showed three nucleotide differences from the sequence of her husband, but was identical to the sequence of her mother. These results support previous seroepidemiological studies that HTLV-1 transmission occurs from mother to children and also between spouses.

INTRODUCTION

Human T lymphotropic virus type-1 (HTLV-1), a human retrovirus, has been linked etiologically to adult T cell leu-kemia (ATL); a neoplasm of T cell origin1–3, HTLV-1-associated myelopathy/tropical spastic paraesthesia (HAM/TSP)4,5 and immunological disorders.6 A seroepidemiologic study of antibody to HTLV-1 (anti-HTLV-1) in an area endemic for ATL showed that there are many apparently healthy HTLV-1 carriers.7 We reported that the Yaeyama District of Okinawa Prefecture, a subtropical zone in the most southwestern part of Japan, is endemic for HTLV-1. Our seroepidemiological studies suggested that HTLV-1 is transmitted from mother to child by breast feeding and between husband and wife.8,9

Recent improvements in molecular biology techniques allow for an exact analysis of routes of transmission of microorganisms. A low degree of genomic variability within one or two generational transmissions has been found in Zaire,10 Papua New Guinea, the Solomon Islands,11 and in French patients given blood transfusions.12 Direct sequencing of polymerase chain reaction (PCR) products was used in these studies, hence the issue of intrachromosomal or the rate of genetic evolution could not be addressed. To determine the route of familial transmission and to compare Yaeyama strains with others, we sequenced three regions (gag, env, int) of the HTLV-1 proviral genome in 7 family clusters in Okinawa, Japan.

MATERIALS AND METHODS

Subjects. We investigated 7 families that included 18 healthy HTLV-I carriers (7 men, 11 women) (Table 1). All lived in the Yaeyama District of Okinawa, Japan, an area where HTLV-118,8,9 and hepatitis B virus13,14 infection are endemic. None of these HTLV-I carriers had a history of blood transfusion and all were healthy. All sera were positive for anti-HTLV-1 by particle agglutination (Serodia ATL kit, Fujirebio, Japan) and western blotting using HTLV-1 antigens prepared form MT2 cells. Family A had three children other than Carrier No.3, all of whom were negative for anti-HTLV-1. The fathers of Families B, C, D, E and F were all negative for anti-HTLV-1. Two young children of Family G refused blood testing. These 7 families had no blood relation-ship with each other. Informed written consent to participate in the present study was obtained from 18 HTLV-1 carriers and from their parents if they were under 20 years old. The presence of anti-HTLV-1 antibody was detected as described above. Anti-p40tax antibody was assayed by ELISA using recombinant p40tax protein expressed in Escherichia coli with a full length HTLV-1 tax gene.15

Genomic DNA preparation and PCR amplification. Mononuclear cells were separated from peripheral blood on LSM, as described by Hirata and others.16 Genomic DNA was extracted from 5 million peripheral blood cells using IsoQuick (Orca Research, WA) following the manufacturer’s instructions. The first PCR mixture contained 300 ng of genomic DNA, 50pM each of the outer primer pair, 0.625U of Taq polymerase (GeneTaq, Wako, Japan), the buffer (50mM KCl, 10mM Tris-HCl, 2.5 mM MgCl2), and 0.2 mM of each of four dNTP. The second PCR was done with the inner primer pair under the same conditions as described above. The strategy for PCR amplification of the HTLV-1 gag, env, and int regions is depicted schematically in Figure 1. Primer sequences are as follows.

Gag13, 5'-CTTAAACCATCGCCCATGGCAA-3' (nt 1020 to 1140); Gag32, 5'GAGTTGCTGTATTCTGCTT-3' (nt 1239 to 1218); Gag31, 5'-TTTTCGATCTGAGTAGCTAA-3' (nt 1540 to 1519); PE1, 5'-TGAATGGATAACACTGCA-3' (nt 4973 to 4895); PE2, 5'-TGCATGCCCAAGACGTCGGA-3' (nt 5121 to 5142); E1, 5'-TAATAGCTCCCGCCATTCGGT-3' (nt 5478 to 5457); Int51, 5'-CTTCTTTATTCTATTGGCAGGA-3' (nt 6539 to 6559); Int31, 5'-GAGGAAGAGTGTTAACACAGCGCA-3' (nt 7069 to 7089); Int32, 5'-AACGCATCGTGATCGGCAGG-3' (nt 8062 to 8067).

Nucleotide numbers are referenced to ATK-1.3

Cloning and sequencing. PCR products were ligated into pT7 Bluescript (Novagen, CA). Plasmid DNA was purified from selected bacterial clones using the QIAprep spin plasmid kit (Qiagen, Hilden, Germany) and sequenced using the 7-Deaza Thermosequenase kit (Amersham Pharmacia Biotech, Uppsala, Sweden) according to the manufacturer’s instructions. The obtained sequences were compared with the published HTLV-I sequences. Nucleotide numbering was referenced to the ATK-1 sequence.3
INTRA- AND INTERFAMILIAL GENETIC ANALYSIS HTLV-1

Genetic analysis. Sequences of HTLV-1 carriers were compared with each other and with the prototype HTLV-1 sequence, ATK-1, derived from a Japanese ATL patient. The portion of the HTLV-1 proviral DNA tested was 1079 bp (gag 279 bp, env 314 bp, and int 486 bp). We also compared the consensus sequences of our HTLV-1 carriers with the ATK-1, MT2,17,18 CH,19 HS35,20 EL,19 and Mel521 sequences.

In order to obtain a reference sequence for the gag, env, and int region, a grand consensus was generated by comparing sequences from 18 HTLV-1 carriers.

Nucleotide sequences were aligned and compared with reported sequences of HTLV-1 strains from various geographic regions. Sequence alignments were facilitated by using the FASTA computer program in the DNA Data Bank of Japan (DDBJ).22 Phylogenetic trees were constructed by use of the CLUSTALW computer program of the DDBJ.23

RESULTS

HTLV-I proviral DNA nucleotide mutations in the three sequenced regions (gag, env, and int) derived form each of 18 HTLV-1 carriers form 7 families are summarized in Table 2. Nucleotide substitutions were found in the env or int regions of 16 carriers, but in the gag region of only 8 carriers studied. Analysis of HTLV-1 clones from amplified DNA samples revealed genetic variations in all three sequenced regions. The nucleotide substitution frequency ranges were 0–0.24% in gag, 0–0.54% in env, and 0–0.34% in int. The nucleotide substitutions were not associated with age, sex, or anti-p40 tax antibody status.

HTLV-1 nucleotide differences from the consensus of the 18 analyzed sequences were divided into 2 groups depending on the env and int regions. The nucleotide sequences of Families D, E, F, and G were identical to the consensus sequences in all three regions studied.

TABLE 2

HTLV-I proviral DNA nucleotide mutations in three regions (gag, env, int) in 18 HTLV-1 carriers from 7 families

<table>
<thead>
<tr>
<th>HTLV-I carrier</th>
<th>gag</th>
<th>env</th>
<th>int</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of clones</td>
<td>Number of clones</td>
<td>Number of clones</td>
<td></td>
</tr>
<tr>
<td>0.07%</td>
<td>0.21%</td>
<td>0.11%</td>
<td></td>
</tr>
</tbody>
</table>

nt. = nucleotide.
Family A had 12 different nucleotides from consensus (Figure 2). There was one substitution in the \textit{gag} region, four in the \textit{env}, and seven in the \textit{int} region. Family B had three substitutions in the \textit{env} region and five in the \textit{int} region. Family C had three substitutions in the \textit{env} region and four in the \textit{int} region.

Carrier No.8 had nucleotide differences from Carrier No.7 (his wife) and from Carrier No.6 (his wife’s mother) in the \textit{env} and \textit{int} regions, but not in the \textit{gag} region. The wife and her mother had identical sequence in all three regions studied.

Substitutions at 1432 (C to T) in the \textit{gag} region, 5295 (C to T) in the \textit{env} region, 6680 (C to G) and 6726 (C to A) in the \textit{int} region were seen only in family A. Substitutions at 6657 (T to C) in the \textit{int} region were found in families A and B, but not in family C. Families A, B, and C were apparently distinct from each other and also from families D, E, F, and G.

The ratio of substitutions to total nucleotides in the three regions were 0–0.36% in the \textit{gag} region, 0–1.6% in the \textit{env} region, and 0–1.4% in the \textit{int} region. The proviral DNA of the \textit{gag} region had the fewest substitutions among the three regions in the 18 sequences studied ($P < 0.05$).

Our consensus sequence was compared with 6 reported sequences (Table 3). It showed the closest identity to the ATK-1 and MT2 strains, within 2%, and was the same as the CH strain in the \textit{gag} region. It was also similar to the MT2 strain in the \textit{env} region and from the MT2 and CH strains in the \textit{int} region within 1%. The HS35 (1.23–1.43%), EL (1.79–6.37%), and Mel5 (5.73–13.4%) strains were very different from our consensus in all three regions studied.

The phylogenetic trees of 7 strains, including our consensus, are shown in Figure 3. By analysis of the \textit{gag} (Figure 3A) and \textit{env} regions (Figure 3B), the strains were classified into two groups: our consensus, CH, MT2, ATK-1, HS35; and others. In the \textit{int} region (Figure 3C), only the Mel5 strain was isolated from the other strains, including our consensus.

\section*{DISCUSSION}

HTLV-1 is the blood borne virus, as are human immunodeficiency virus, hepatitis B virus, and hepatitis C virus, all of which have viral mutations within individuals. The substitution frequency of the HTLV-1 carriers in our study was 0–0.54% in the three regions studied, and indicates that HTLV-1 is highly conserved in individual carriers.

Our previous study demonstrated that the levels of proviral DNA and double spliced mRNA were significantly higher in

\begin{table}
\centering
\begin{tabular}{|l|c|c|c|}
\hline
\textbf{Isolates} & \textbf{\textit{gag}} & \textbf{\textit{env}} & \textbf{\textit{int}} \\
\hline
\textbf{Japanese strains} & & & \\
ATK-1 & 1 (0.36) & 6 (1.91) & 9 (1.85) \\
MT2 & 2 (0.72) & 3 (0.96) & 4 (0.82)* \\
\hline
\textbf{Caribbean strains} & & & \\
CH & 0 (0) & 6 (1.91) & 4 (0.82) \\
HS35 & 4 (1.43) & 4 (1.27) & 6 (1.23) \\
\hline
\textbf{African strain} & & & \\
EL & 5 (1.79) & 7 (2.22) & 31 (6.37)$\ddagger$ \\
\hline
\textbf{Melanesian strain} & & & \\
Mel5 & 16 (5.73)$\S$ & 20 (6.36) & 65 (13.4) \\
\hline
\end{tabular}
\caption{Nucleotide differences between geographical isolates and the consensus of carriers of Okinawa, Japan}
\end{table}

* Underestimated for partial analysis. The MT2 isolate has not been completely sequenced.

$\ddagger$ Mel5 has an additional nucleotide between 1261 and 1262, that causes a frameshift. Amino acid changes happened more often than nucleic changes.

\section*{FIGURE 3.} Phylogenetic tree based on geographical strains and consensus sequences in the \textit{gag} (A), \textit{env} (B), and the \textit{int} regions (C). The MT2 sequence was not analyzed in the \textit{int} region because only part of this region has been published.
anti-p40\textsuperscript{tax}-positive carriers than in negative individuals. We previously concluded that the presence of anti-p40\textsuperscript{tax} may serve as a marker for a higher virus load and virus replication in asymptomatic HTLV-1 carriers. However, the results of this study showed that neither frequency nor position of mutation within individuals correlated with age, sex, or anti-p40\textsuperscript{tax} antibody status. Nucleotide substitutions may therefore not be affected by replication circumstances.

Until recently, there was no method to accurately determine the route of HTLV-1 transmission. Our study found nucleotide sequence analysis to be useful; however, of the three regions studied, the gag region was not useful because viral mutations were infrequent. Previous studies revealed three main HTLV-1 transmission routes: mother-to-child transmission, most likely by the resulting from breast feeding, sexual transmission (mainly husband to wife) and blood transfusion. No HTLV-1 carrier in the present study had a history of blood transfusion. Our results that 8 mother/child pairs had identical nucleotide sequences confirmed our previous findings of mother-to-child transmission in a seroepidemiologic survey of this same area. Even in Families A, B, and C, where members who had relatively variable mutations, the sequences of mothers were identical to those from their children.

Because this study concerned too few HTLV-1 carriers, the exact route of HTLV-1 infection between spouses was not confirmed. Of 3 pairs of spouses, 2 (Carriers No.1 and No.2; Carriers No.16 and No.17) had identical sequences. The mother of Carriers No.2 and No.17 were negative for anti-HTLV-1, indicating that transmission was probably from their husbands. The remaining spousal pair had different sequences, but Carrier No.7 and her mother (Carrier No.6) had identical sequences, indicating probable transmission from the mother.

We extended the analysis to determine the geographical origin of the HTLV-1 strains of our carriers. The Japanese (ATK-1 and MT2) and Caribbean (CH and HS35) are the so-called “Cosmopolitan” strains. Our finding that most HTLV-1 carriers in Okinawa were infected with the Cosmopolitan strain supports the notion that there is a low rate of genetic drift among the various geographical isolates. Molecular analysis of HTLV-1 strains revealed the geographic distribution. We were unable to confirm the geographical origins of our species because we had sequences covering only 2 generations.

Because the HTLV-1 genome is well conserved, nucleotide analysis is useful for confirming the route of HTLV-1 transmission.

**REFERENCES**


