

Preferential selection of human T-cell leukemia virus type I provirus integration sites in leukemic versus carrier states

Keitarou Doi, Xiaolin Wu, Yuko Taniguchi, Jun-ichirou Yasunaga, Yorifumi Satou, Akihiko Okayama, Kisato Nosaka, and Masao Matsuoka

Human T-cell leukemia virus type I (HTLV-I) is a causative agent of neoplastic disease, adult T-cell leukemia (ATL). Although the encoding viral proteins play an important role in oncogenesis, the role of the HTLV-I proviral integration site remains unsolved. We determined the integration sites of HTLV-I proviruses in ATL cells and HTLV-I-infected cells in asymptomatic carriers. In carrier and ATL cells, HTLV-I provirus was integrated into the transcriptional unit at frequencies of

26.8% (15/56) and 33.9% (20/59), respectively, which were equivalent to the frequency calculated based on random integration (33.2%). In addition, HTLV-I provirus was prone to integration near the transcriptional start sites in leukemic cells ($P = .006$), and the transcriptional direction of the provirus was in accordance with that of integrated cellular genes in 70% of cases. More importantly, the integration sites in the carrier cells favored the alphoid repetitive sequences

(11/56; 20%) whereas in leukemic cells they disfavored these sequences (2/59; 3.4%). Taken together, during natural course from carrier to onset of ATL, HTLV-I-infected cells with integration sites favorable for viral gene transcription are susceptible to malignant transformation due to increased viral gene expression. (Blood. 2005;106:1048-1053)

© 2005 by The American Society of Hematology

Introduction

After infection with retrovirus, reverse transcriptase synthesizes proviral DNA and then integrates the provirus into the host genome by the action of an integrase. In some retrovirus-associated neoplasms, provirus insertion enhances the transcription of oncogenes, such as the *myc* gene, resulting in transformation of infected cells.^{1,2} Although the integration sites of proviruses in the host genome have been considered random, recent studies regarding various retroviruses revealed that human immunodeficiency virus type 1 (HIV-1) prefers transcriptional units,³ whereas murine leukemia virus (MLV) tends to integrate near the transcriptional start sites.⁴ These findings suggest that the integration of proviruses depends on mechanisms unique to each retrovirus, which interact with host factors associated with nuclear transport, DNA repair, and chromatin structure.⁵

Human T-cell leukemia virus type I (HTLV-I) is the causative virus of adult T-cell leukemia (ATL) and inflammatory disease, HTLV-I-associated myelopathy (HAM)/tropical spastic paraparesis (TSP).⁶ HTLV-I infection induces ATL in a portion of infected individuals after a long latent period. The characteristic of HTLV-I is the presence of accessory genes, which are encoded by the pX region between *env* and the 3'-long terminal repeat (LTR).^{7,8} Among the accessory genes, *tax* is considered to play a central role in the proliferation of infected cells and leukemogenesis because of its pleiotropic actions. Tax activates transcriptional pathways such as nuclear factor κ B (NF- κ B), serum response factor (SRF), and cyclic AMP response element binding protein (CREB), leading to activated transcription of growth factor and its receptor genes, and inhibition of apoptosis. In addition, Tax can transrepress the

transcription of cellular genes and functionally inhibit p53,⁹ p16,¹⁰ and MAD1.¹¹ Such pleiotropic actions induce the proliferation of HTLV-I-infected cells, and inhibit apoptosis, resulting in clonal expansion in vivo.

Although it has been reported that the integration sites of HTLV-I provirus are random, preferential integration into the transcriptional units has been reported in ATL cells.^{12,13} In this study, we compared HTLV-I integration sites between carriers and leukemic cells, and found that the provirus was frequently integrated into alphoid repetitive sequences in the carrier state, but not the leukemic state.

Materials and methods

Patient samples

Peripheral blood mononuclear cells (PBMCs) were separated from heparinized peripheral blood by Ficoll-Hypaque density gradient centrifugation. They were then digested with proteinase K and treated with RNase A to eliminate RNA. Genomic DNAs were extracted from the PBMCs of 16 HTLV-I carriers and 59 patients with ATL. Approval for this study was obtained from the institutional review board of the Kyoto University. The informed consent was obtained from blood donors and patients according to the Declaration of Helsinki.

Inverse long polymerase chain reaction

Inverse long polymerase chain reaction (IL PCR) was used to amplify the genomic DNA adjacent to the integration sites of the HTLV-I

From the Laboratory of Virus Immunology, Institute for Virus Research, Kyoto University, Kyoto, Japan; Laboratory of Molecular Technology, National Cancer Institute-Frederick, Frederick, MD; and the Department of Laboratory Medicine, Faculty of Medicine, University of Miyazaki, Miyazaki, Japan.

Submitted November 15, 2004; accepted April 3, 2005. Prepublished online as *Blood* First Edition Paper, April 19, 2005; DOI 10.1182/blood-2004-11-4350.

Supported by a grant-in-aid for Scientific Research from the Ministry of Education, Science, Sports, and Culture of Japan.

Reprints: Masao Matsuoka, Laboratory of Virus Immunology, Institute for Virus Research, Kyoto University, Kyoto 606-8507, Japan; e-mail: mmatsuok@virus.kyoto-u.ac.jp.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 U.S.C. section 1734.

© 2005 by The American Society of Hematology

provirus. First, genomic DNA (1.5 μ g) was digested with a restriction enzyme (*Hind*III, *Pst*I, or *Eco*RI), then ligated by T4 DNA ligase. When DNA was digested with *Eco*RI, it was digested with *Mlu*I after ligation so as not to amplify the HTLV-I provirus itself. The resulting DNA was used as a substrate for IL PCR, which was performed using TaKaRa LA PCR (Takara, Shiga, Japan). Briefly, primers (final concentration, 0.2 μ M), $MgCl_2$ (2.5 mM), and deoxynucleoside triphosphates (dNTPs, 0.4 mM) were mixed (total 20 μ L) then AmpliWax (Applied Biosystems, Norwalk, CT) was added to each tube. After wax layer formation by incubation at 80°C for 10 minutes and cooling at room temperature for 15 minutes, substrate DNA (0.5 μ g), 10 \times LA buffer (5 μ L), and LA Taq (0.4 μ L) were added (total 30 μ L). Cycles for long PCR were as follows: one cycle of 98°C for 2 minutes, 5 cycles of 98°C for 30 seconds and 64°C for 10 minutes, and 35 cycles of 94°C for 30 seconds, 64°C for 10 minutes, and 72°C for 15 minutes. The primers used in this experiment were as follows: primers 1 and 2 were used for *Pst*I-digested samples, primer 1: 5'-TAGCAGGAGTCTATAAAAGCGTGGAGACAG-3'; primer 2: 5'-TGGAATGTTGGGGGTTGTATGAGTGATTGG-3'. Primers 1 and 3 were used for *Hind*III-digested samples, primer 3: 5'-TGGGCAGGATTGCAGGGTTAGAGTGG-3'. Primers 4 and 5 were used for *Eco*RI-digested samples, primer 4: 5'-TGCCTGACCCTGCTTGCTCAACTCTACGTCTTTG-3', primer 5: 5'-AGTCTGGGCCCTGACCTTTTCAGACTTCTGTTTC-3'.

Cloning and sequencing

IL PCR products from ATL samples were used as a template for direct sequencing to determine the integration site. The primer used for sequencing was as follows: 5'-TCATTACAGACTGACTGCCGG-3'. To determine integration sites from carrier samples, IL PCR products were gel-isolated using a QIAquick Gel Extraction Kit (Qiagen, Valencia, CA), and ligated into a pPCR-Script Amp SK(+) cloning vector (Stratagene, La Jolla, CA) or pCR-TOPO-XL (Invitrogen, Carlsbad, CA). The plasmids were then used as a template for sequencing. Sequencing of IL PCR amplicon was performed using an ABI PRISM Genetic Analyzer 310 or 3100 (Applied Biosystems, Norwalk, CT) according to the manufacturer's instructions.

Mapping integration sites

The BLAST-like Alignment Tool program was used to map sequences to the human genome (University of California–Santa Cruz [UCSC] Human Genome Project Working Draft, July 2003 freeze).¹⁴ Sequence matches were judged to be authentic only if they (1) contained the LTR sequence, (2) showed 95% or greater identity to the genomic sequence over the high-quality sequence region, and (3) matched only one genomic locus with 95% or greater identity. Genomic features such as coding regions and repetitive sequences were investigated using the UCSC Genome Browser (<http://genome.ucsc.edu/cgi-bin/hgBlat?command=start>). Random integration sites (10 000 \times) in nongap regions of the human genome (UCSC Human Genome Project Working Draft, July 2003 freeze) were generated with a computer program using a uniform distribution algorithm, and used for comparison with the observed HTLV-I integration sites.

Statistical analyses

The χ^2 or one-sided Fisher exact test was used to determine statistical significance.

Results

HTLV-I provirus integration sites

We determined the genomic sequences adjacent to the HTLV-I provirus integration sites in ATL cells (59 cases) and HTLV-I–infected cells (56 sites) in the 16 carriers by inverse PCR, and then analyzed (1) chromosomal locations, (2) the genes containing integration sites or neighboring genes, (3) the relation of transcriptional direction between provirus and cellular genes when the

HTLV-I provirus was integrated within the gene, (4) distance from the transcriptional start site, and (5) repetitive sequences. The results are summarized in Table 1.

As reported previously,^{15,16} the integration sites of HTLV-I provirus were random in both leukemic cells and carriers (Figure 1) and were not associated with the guanine and cytosine (GC) content of surrounding genomes (data not shown). In addition, there was no specific chromosome for HTLV-I integration. To investigate whether HTLV-I integration sites prefer transcription unit in the human genome, we used the UCSC Genome Browser RefSeq Genes track, which represents annotated genes based on National Center for Biotechnology Information mRNA reference sequences. An integration site is considered in a transcription unit if it locates within the transcription start site and stop site of a RefSeq gene. Among the 56 HTLV-I integration sites in the carrier cells, 15 (26.8%) were identified within the transcriptional units (RefSeq). On the other hand, 20 integration sites (33.9%) existed within the transcriptional units in the 59 ATL cases (Table 2). As a control, we simulated random HTLV-I integration by placing 10 000 integration sites randomly into the same human genome, of which 33.2% were found within the transcription units (Table 2), a percentage identical to the estimated transcribed human genome.^{17,18} Although the frequency of integration into transcriptional units was low in carriers, the difference between the carriers and ATL cells was not statistically significant.

Since annotation for the human genome has been changing rapidly in the last few years, the previous reports on HTLV-I integration sites should be reexamined.^{12,13} The frequency of integration into the transcriptional units has reduced to 46.6% from 56.4%. When these data are combined with those presented here, the provirus landed in 48 of 119 cases (40.3%), which is not statistically significant compared with random integration (33.2%; $P = .17$ by a χ^2 test). The frequency of integration into the transcriptional units did not differ between ATL and carriers ($P = .08$ by a χ^2 test).

When HTLV-I provirus was integrated into transcriptional units, the transcriptional direction of the provirus was in accordance with that of integrated cellular genes in 70% of cases among the leukemic cells. Since the deletion of 5'-LTR is frequently observed in the provirus of ATL cells, which is a designated type 2 defective provirus,¹⁹ the cellular promoter might act as a promoter for viral genes in such proviruses as reported previously.²⁰ Therefore, we investigated the relationship of transcriptional direction between cellular genes and type 2 defective provirus. In 2 of 4 cases with type 2 defective provirus, the provirus was integrated into transcriptional units, which were the mutated in colorectal cancers (*MCC*) and protocadherin 9 genes (Table 1). HTLV-I provirus was inserted in both genes in a sense direction, suggesting that the promoter of cellular genes might transcribe the viral genes in such cases.

HTLV-I provirus tends to be integrated near the transcriptional start sites in leukemic cells

In MLV, the provirus tended to be integrated near the transcriptional start sites.⁴ On the other hand, HIV-1 has no such tendency in spite of its preference for transcriptional units. Since HTLV-I is a human retrovirus, it is of interest to determine whether HTLV-I has a similar tendency to MLV or HIV-1. In HTLV-I, 9 of the 59 proviruses were integrated near the transcriptional start sites (± 5 kb) in ATL cells (9/59; 15.3%) as shown in Table 3 and Figure 2. In contrast, using the same genome assembly, only 5.6% of the random integration landed near transcriptional start sites. Hence, the frequency of integration near the transcriptional start sites was statistically significant ($P = .006$), which is similar to MLV. Since the number of integration within plus or minus 5 kb from transcription start sites is only 3, there is no preference of integration near the transcriptional start sites.

Table 1. The genes that hosted integration event

Sample	Locus	RefSeq gene	Entrez Gene ID	Exon or intron	Direction
A 1	2q14.2	Protein tyrosine phosphatase, nonreceptor type 4	5775	intron 10	sense
A 2	2q21.1	POTE14	404785	intron 8	antisense
A 3	3p14.3	Calcium channel, voltage-dependent, alpha 2/delta 3 subunit	55799	intron 9	sense
A 4	3p26.3	Contactin 4	152330	intron 2	sense
A 5	4p15.32	Hypothetical protein FLJ90013	202018	intron 1	sense
A 6	4q28.2	PHD finger protein 17	79960	exon 9	sense
A 7	5q22.2	Mutated in colorectal cancers	4163	intron 1	sense
A 8	6q21	REV3-like, catalytic subunit of DNA polymerase zeta	5980	exon 16	sense
A 9	7q11.23	Williams-Beuren syndrome chromosome region 5	7462	intron 1	sense
A 10	8q21.13	CGI-62 protein	51101	intron 1	antisense
A 11	9p21.3	Cyclin-dependent kinase inhibitor 2A (p16)	1029	intron 1	antisense
A 12	10q22.1	Cadherin-like 23	64072	intron 58	antisense
A 13	10q24.32	Chromosome 10 open reading frame 76	79591	intron 3	antisense
A 14	12q24.11	Carnitine deficiency-associated gene expressed in ventricle 1	28981	intron 10	antisense
A 15	13q21.32	Protocadherin 9	5101	intron 2	sense
A 16	14q11.2	Heterogeneous nuclear ribonucleoprotein C (C1/C2)	3183	intron 1	sense
A 17	14q32.11	Echinoderm microtubule associated protein-like 5	161436	intron 2	sense
A 18	19q13.11	CCAAT/enhancer binding protein (C/EBP), alpha	1050	exon 1	sense
A 19	20p12.3	Hypothetical protein DJ971N18.2	56255	intron 1	sense
A 20	20q13.12	Zinc finger, SWIM domain containing 3	140831	intron 1	sense
C 1	1q41	UNQ6077	375056	intron 4	antisense
C 2	1q43	Ryanodine receptor 2 (cardiac)	6262	intron 2	antisense
C 3	2q37.3	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 10, 42 kDa	4705	exon 3	antisense
C 4	4q23	Alcohol dehydrogenase 5 (class III), chi polypeptide	128	exon 3	sense
C 5	5q21.3	F-box and leucine-rich repeat protein 17	64839	intron 6	sense
C 6	6p21.2	Ring finger protein (C3HC4 type) 8	9025	intron 2	antisense
C 7	6q15	RNA guanylyltransferase and 5'-phosphatase	8732	intron 13	sense
C 8	8q24.22	Thyroglobulin	7038	intron 41	antisense
C 9	9q21.11	LOC220869: dopamine responsive protein	220869	intron 2	sense
C 10	10p15.1	GDP dissociation inhibitor 2	2665	intron 9	sense
C 11	11q13.1	EGF-containing fibulin-like extracellular matrix	30008	exon 12	sense
C 12	14q24.1	RAD51-like 1	5890	intron 7	antisense
C 13	15q21.3	Hypothetical protein FLJ38736	256764	intron 1	sense
C 14	20q13.33	GTP binding protein 5 (putative)	26164	intron 5	antisense
C 15	21q21.1	Ubiquitin specific protease 25	29761	intron 19	antisense

Genes were identified by UCSC Genome Browser, <http://genome.ucsc.edu/cgi-bin/hgBlat?command=start>. Entrez gene ID was assigned by NCBI Entrez Gene, <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>.

A and C indicate ATL and carrier samples, respectively.

Integration sites in repetitive sequences

Next, we studied the relationship between repetitive sequences and integration sites. Percentages of repetitive sequences in human genome were based on the previous report by Venter et al.¹⁸ Among the carriers, 11 sites (11/56; 20%) resided in the aliphoid repetitive

sequences, which are a component of centromeric heterochromatin and have a monomeric repeating unit of 171 bp,²¹ whereas only 2 integration sites were identified within aliphoid sequences in ATL (2/59; 3.4%). Since aliphoid sequences were estimated to compose 3% to 5% of the human genome, we used 5% for the statistical

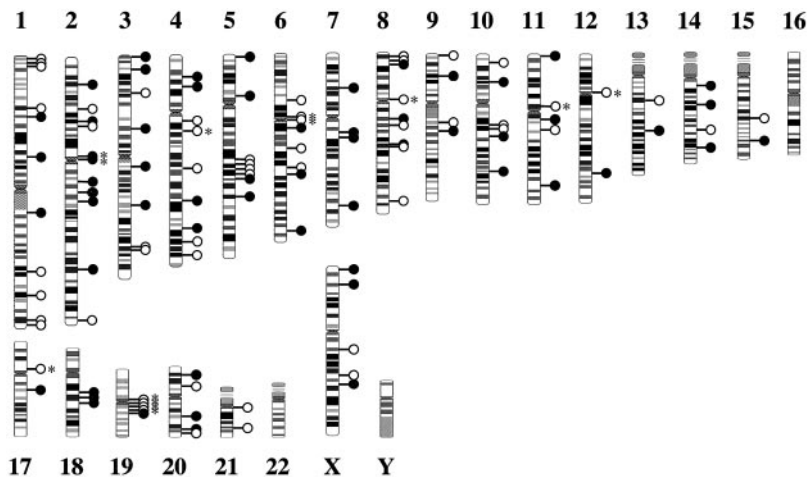


Figure 1. Distribution of HTLV-I provirus integration sites in human chromosomes. The integration sites of HTLV-I provirus in ATL patients (●) and HTLV-I carriers (○) are demonstrated. Integration within aliphoid repetitive sequence is marked by an asterisk.

Table 2. Correlation between RefSeq genes and integration sites

	Carrier	ATL	Random
	No. (%)	No. (%)	No. (%)
Total	56 (100.0)	59 (100.0)	10 000 (100.0)
In RefSeq genes	15 (26.8)	20 (33.9)	3 324 (33.2)
Sense*	7 (46.7)	14 (70.0)	ND (ND)

ND indicates not done.

*The percentage of sense-direction has been calculated among sites, in which HTLV-I provirus was integrated in transcriptional units.

analyses (Table 4). HTLV-I-infected cells that have the provirus integrated in aliphoid sequences are enriched in HTLV-I-infected cells during the carrier state compared with ATL ($P = .0153$; one-sided Fisher exact test; Table 4). The difference between ATL and carrier cells was statistically significant ($P = .0059$; one-sided Fisher exact test), indicating that integration within aliphoid sequences is disfavored in leukemic cells. There were no preferences of HTLV-I integration with other repetitive sequences by statistical analyses.

Discussion

Recent studies on the integration sites of proviruses have provided new insights into the mechanism of integration and pathogenesis of retroviral infections.^{3,4,22} In MLV,⁴ provirus integration tends to occur near transcriptional start sites although there was no preference toward transcriptional units (34.2%) as observed in HTLV-I. On the other hand, HIV-1 tends to be integrated within transcriptional units (57.8%⁴ and 69%³) in vitro. In vivo data on the integration of HIV-1 provirus demonstrated that most (91%) was integrated within transcriptional units, the genes of which were transcribed in T lymphocytes,²² indicating that HIV-1 integration targets transcriptional active regions more than expected from in vitro data. On the other hand, this report showed that integration of the HTLV-I provirus into transcriptional units was not frequent compared with random integration, which is not consistent with the previous studies.^{12,13} This is because of the changing database. We analyzed the integration sites and genes by the UCSC Genome Browser (July 2003). In HTLV-I carriers, integration into the transcriptional units was rather less frequent than random integration although the difference was not statistically significant. Combined with the finding that HTLV-I tends to be integrated near the transcriptional start sites, the pattern of HTLV-I integration is similar to that of MLV rather than that of another human retrovirus, HIV-1.

In the carrier state, HTLV-I provirus tends to be integrated into aliphoid repetitive sequences, whereas the frequency of integration into aliphoid sequences was significantly decreased in leukemic cells. When HIV-1 provirus is integrated into aliphoid sequences, it establishes a latent infection by influencing the surrounding heterochromatin in vitro.²³ Heterochromatin decreases the basal

Table 3. Correlation between transcription start sites and integration sites

	Carrier		ATL		Random, no.
	No.	P	No.	P	
Total	56	NA	59	NA	10 000
Within 5 kb of transcription start sites	3	.610	9	.006	565

P compared with random integration with the use of the Fisher exact test. NA indicates not applicable.

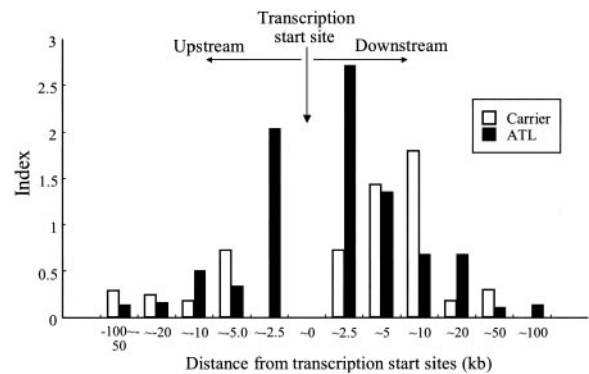


Figure 2. Relationship between HTLV-I provirus integration sites and transcriptional start sites. The distances from the transcriptional start sites are shown. The index was defined by the following calculation: (observed number of integrations) \times 100 / (total number) \times (window size).

transcription of viral genes, resulting in a latent state. Taken together, it is possible that infected cells in which HTLV-I provirus is integrated into aliphoid sequences are enriched in the carrier state. Such cells are considered to produce lesser amounts of viral proteins, which facilitate cells to escape from the host immune surveillance system.²⁴ However, integration into aliphoid sequences was not frequent in the leukemic cells, indicating that those with higher amounts of viral proteins are more likely to transform into malignant cells among surviving HTLV-I-infected cells. Since ATL occurs among HTLV-I carriers after a long latent period, these findings indicate the scenario as follows: after transmission of HTLV-I, HTLV-I provirus is randomly integrated into the host genome. The host immune system, including cytotoxic T lymphocytes (CTLs), excludes the HTLV-I-infected cells, in which Tax protein is the major target of CTLs.²⁵ In such circumstance, HTLV-I-infected cells expressing lesser amount of viral proteins are selected in vivo. However, among such infected cells, HTLV-I-infected cells expressing viral protein, such as Tax, tend to proliferate in vivo. Such cells should have greater chance to transform into malignant cells. A higher provirus load has been reported to be a risk factor for development of ATL, which is consistent with this hypothesis. Therefore, integration into aliphoid repetitive sequences is less frequent among leukemic cells since viral transcription of such cells tends to be silenced.

Although MLV can infect only dividing cells,²⁶ lentiviruses such as HIV-1 can infect nondividing cells by transfer of a preintegration complex through the nuclear pore.²⁷ Since transcriptional active sites are associated with nuclear transport machinery,²⁸ preintegration complexes that pass through the nuclear pore might be integrated into transcriptional active sites due to the open structure of chromatin. This might be the reason for the high frequency at which HIV-1 provirus is integrated within transcriptional active genes. The data in this study reveal that the characteristics of HTLV-I integration in ATL cells resemble those of MLV, suggesting that HTLV-I cannot infect nondividing cells, although this requires clarification. Preference to transcriptional start sites was observed only in leukemic cells, indicating that such integration sites confer the advantage in leukemogenesis. Since the integration sites of HTLV-I provirus concentrate within 5 kb from transcriptional start sites, it is possible that such sites are suitable for transcription of viral genes.

Type 2 defective HTLV-I provirus lacks 5'-LTR and internal viral sequences such as *gag* and *pol*.¹⁹ It is possible that this provirus traps the cellular promoter, thus ensuring transcription. In this study, the provirus of 2 of the 4 cases with type 2 defective

Table 4. Correlation between chromosomal features and integration sites

	Human genome, %	Carrier		ATL		
		No. (%)	P_1^*	No. (%)	P_1^*	P_2^\dagger
Total		56	NA	59	NA	NA
SINES	12.8	3 (5.3)	NS	4 (6.8)	NS	NS
LINES	20	10 (17.9)	NS	9 (15.3)	NS	NS
DNA elements	2.8	1 (1.8)	NS	2 (3.4)	NS	NS
LTR elements	8.3	6 (10.7)	NS	4 (6.8)	NS	NS
Satellite						
Alpha satellite	5	11 (19.6)	.0153	2 (3.4)	NS	.0059
Beta satellite	UN	0 (0.0)	UN	0 (0.0)	UN	NS

Percentages of repetitive sequences in the human genome were based on Venter et al.¹⁸

NS indicates not significant; UN, unknown; NA, not applicable.

* P_1 shows comparison with genome frequency with the use of Fisher exact test.

† P_2 shows comparison with carrier integration with the use of Fisher exact test.

provirus was integrated in the transcriptional unit in a sense orientation, indicating that the cellular promoter might transcribe the viral gene. In addition, HTLV-I provirus contains an internal promoter sequence in the *pol* region,²⁹ which is considered to transcribe the viral gene. This is thought to have occurred especially in the remaining 2 cases, which showed integration of the provirus outside the transcriptional units.

In HTLV-1–induced oncogenesis, viral proteins such as Tax promote the proliferation of HTLV-I–infected cells and induce ATL in about 2% to 6% of carriers after a long latent period.^{7,8} Since HTLV-I provirus integration is random,¹⁵ integration itself does not directly influence leukemogenesis. Viral products such as Tax promote the proliferation of HTLV-I–infected cells and induce transformation of infected T lymphocytes.^{30,31} However, expression of Tax protein is impaired by several mechanisms in ATL cells, including deletion of 5′-LTR,¹⁹ DNA methylation of 5′-LTR,³² and genetic changes (deletion, insertion and nonsense mutations) of the *tax* gene itself.^{33,34} In the carrier state, presence of the Tax protein is advantageous for the proliferation and survival of infected cells; however, since Tax is the major target of CTLs in vivo,²⁵ the

growth of Tax-expressing cells is suppressed by CTLs.³⁵ When HTLV-I provirus is integrated into transcriptional active sites, viral gene transcription is thought to be active. During the carrier state, these producer cells are possibly eliminated by CTLs. Therefore, cells infected with HTLV-I provirus integrated into aliphoid sequences have been enriched. On the other hand, higher production of viral proteins in HTLV-I–infected cells is thought to take advantage during malignant transformation. Therefore, the frequency of integration into aliphoid sequences was low in leukemic cells compared with carrier states.

In this study, analyses of HTLV-I integration sites in both leukemic and HTLV-I–infected cells of carriers have emphasized the structural significance of the host genome, which influences viral gene transcription, during leukemogenesis.

Acknowledgment

We thank Suzuko Ohsako for technical help.

References

- Hayward WS, Neel BG, Astrin SM. Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukemia. *Nature*. 1981;290:475-480.
- Nusse R, Varmus HE. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell*. 1982;31:99-109.
- Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*. 2002;110:521-529.
- Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. *Science*. 2003;300:1749-1751.
- Mitchell RS, Beitzel BF, Schroder AR, et al. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol*. 2004;2:1127-1137.
- Matsuoka M. Human T-cell leukemia virus type I and adult T-cell leukemia. *Oncogene*. 2003;22:5131-5140.
- Yoshida M, Suzuki T. HTLV type 1 Tax oncoprotein binds to DNA topoisomerase I and inhibits its catalytic activity. *AIDS Res Hum Retroviruses*. 2000;16:1639-1645.
- Franchini G, Fukumoto R, Fullen JR. T-cell control by human T-cell leukemia/lymphoma virus type 1. *Int J Hematol*. 2003;78:280-296.
- Suzuki T, Uchida-Toita M, Yoshida M. Tax protein of HTLV-1 inhibits CBP/p300-mediated transcription by interfering with recruitment of CBP/p300 onto DNA element of E-box or p53 binding site. *Oncogene*. 1999;18:4137-4143.
- Suzuki T, Kitao S, Matsushime H, Yoshida M. HTLV-1 Tax protein interacts with cyclin-dependent kinase inhibitor p16INK4A and counteracts its inhibitory activity towards CDK4. *EMBO J*. 1996;15:1607-1614.
- Jin DY, Spencer F, Jeang KT. Human T cell leukemia virus type 1 oncoprotein Tax targets the human mitotic checkpoint protein MAD1. *Cell*. 1998;93:81-91.
- Ozawa T, Itoyama T, Sadamori N, et al. Rapid isolation of viral integration site reveals frequent integration of HTLV-1 into expressed loci. *J Hum Genet*. 2004;49:154-165.
- Hanai S, Nitta T, Shoda M, et al. Integration of human T-cell leukemia virus type 1 in genes of leukemia cells of patients with adult T-cell leukemia. *Cancer Sci*. 2004;95:306-310.
- Kent WJ. BLAT—the BLAST-like alignment tool. *Genome Res*. 2002;12:656-664.
- Seiki M, Eddy R, Shows TB, Yoshida M. Nonspecific integration of the HTLV provirus genome into adult T-cell leukaemia cells. *Nature*. 1984;309:640-642.
- Ohshima K, Ohgami A, Matsuoka M, et al. Random integration of HTLV-1 provirus: increasing chromosomal instability. *Cancer Lett*. 1998;132:203-212.
- Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409:860-921.
- Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science*. 2001;291:1304-1351.
- Tamiya S, Matsuoka M, Etoh K, et al. Two types of defective human T-lymphotropic virus type I provirus in adult T-cell leukemia. *Blood*. 1996;88:3065-3073.
- Kubota S, Furuta RA, Siomi H, Maki M, Hatanaka M. Analysis of a novel defective HTLV-I provirus and detection of a new HTLV-I-induced cellular transcript. *FEBS Lett*. 1995;375:31-36.
- Choo KH, Vissel B, Nagy A, Earle E, Kalitsis P. A survey of the genomic distribution of alpha satellite DNA on all the human chromosomes, and derivation of a new consensus sequence. *Nucleic Acids Res*. 1991;19:1179-1182.
- Han Y, Lassen K, Monie D, et al. Resting CD4⁺ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. *J Virol*. 2004;78:6122-6133.
- Jordan A, Biggrove D, Verdin E. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *EMBO J*. 2003;22:1868-1877.
- Bangham CR. Human T-lymphotropic virus type 1 (HTLV-1): persistence and immune control. *Int J Hematol*. 2003;78:297-303.

25. Kannagi M, Harada S, Maruyama I, et al. Pre-dominant recognition of human T cell leukemia virus type I (HTLV-I) pX gene products by human CD8⁺ cytotoxic T cells directed against HTLV-I-infected cells. *Int Immunol*. 1991;3:761-767.
26. Mulligan RC. The basic science of gene therapy. *Science*. 1993;260:926-932.
27. Naldini L, Blomer U, Gally P, et al. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*. 1996;272:263-267.
28. Casolari JM, Brown CR, Komili S, West J, Hieronymus H, Silver PA. Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell*. 2004;117:427-439.
29. Ariumi Y, Shimotohno K, Noda M, Hatanaka M. Characterization of the internal promoter of human T-cell leukemia virus type I. *FEBS Lett*. 1998;423:25-30.
30. Grassmann R, Dengler C, Muller-Fleckenstein I, et al. Transformation to continuous growth of primary human T lymphocytes by human T-cell leukemia virus type I X-region genes transduced by a Herpesvirus saimiri vector. *Proc Natl Acad Sci U S A*. 1989;86:3351-3355.
31. Akagi T, Ono H, Shimotohno K. Characterization of T cells immortalized by Tax1 of human T-cell leukemia virus type 1. *Blood*. 1995;86:4243-4249.
32. Koiwa T, Hamano-Usami A, Ishida T, et al. 5'-long terminal repeat-selective CpG methylation of latent human T-cell leukemia virus type 1 provirus in vitro and in vivo. *J Virol*. 2002;76:9389-9397.
33. Furukawa Y, Kubota R, Tara M, Izumo S, Osame M. Existence of escape mutant in HTLV-I tax during the development of adult T-cell leukemia. *Blood*. 2001;97:987-993.
34. Takeda S, Maeda M, Morikawa S, et al. Genetic and epigenetic inactivation of tax gene in adult T-cell leukemia cells. *Int J Cancer*. 2004;109:559-567.
35. Kannagi M, Ohashi T, Harashima N, Hanabuchi S, Hasegawa A. Immunological risks of adult T-cell leukemia at primary HTLV-I infection. *Trends Microbiol*. 2004;12:346-352.