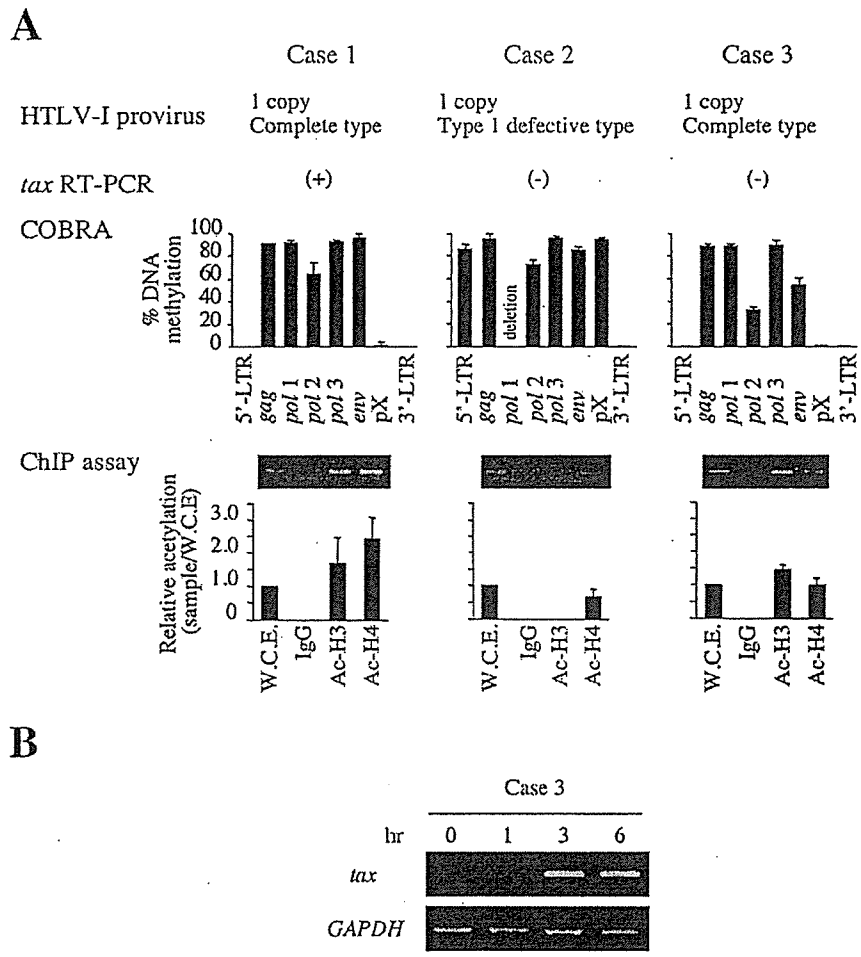

Figure 7

DNA methylation and histone modifications in fresh ATL cases. A. The relationships among DNA methylation, *tax* gene expression and histone modification in 5'-LTR were analyzed in three ATL cases. Cases 1 and 3 have one copy of the complete HTLV-I provirus, while Case 2 has a defective provirus that lacks part of the *pol* gene. DNA methylation was analyzed by COBRA. The *tax* gene transcripts could be detected in Case 1, but not in Cases 2 or 3, by RT-PCR. ChIP assays were also performed using primers for 5'-LTR to analyze acetylation of histone H3 (Ac-H3) and H4 (Ac-H4). W.C.E.: whole cell extract. B. Recovery of *tax* gene expression *ex vivo*. The PBMCs isolated from Case 3 were immediately cultured *ex vivo* for several hours and tested the transcription of *tax* mRNA by RT-PCR.

expressing cell line, ATL-48T, a difference was found in the acetylation of histone H3 in 5'-LTR (Fig. 6A and 6B). The histone H3 of 5'-LTR was hypoacetylated in ATL-43T compared with ATL-48T, whereas there were no differences in pX or 3'-LTR among these cell lines. Since the number of HTLV-I provirus in ATL-43T and -48T is one and two copies respectively, and acetylation of histone H3 in pX and 3'-LTR was similar in both cell lines, the number of provirus was thought to have no influence on the results of ChIP assay in 5'-LTR.

However, the *tax* gene transcription is silenced in about 20% of ATL cases despite no or partial methylation of 5'-LTR (Fig. 3B) [13], suggesting that there is aother mechanism(s) for suppressing viral gene transcription. To address this question, we studied the histone modification of 5'-LTR in fresh ATL cells with or without *tax* gene transcription. In a case with *tax* gene expression, 5'-LTR was not methylated and histone H3 was hyperacetylated (Fig. 7A, Case 1). On the other hand, in Case 2 with heavily methylated 5'-LTR, histone H3 was hypoacetylated in 5'-LTR, which was consistent with the lack of detection of *tax* gene transcription in this case. However, in Case 3, *tax* gene transcription could not be detected regardless of 5'-LTR hyperacetylation. After *in vitro* culture, such cells showed *tax* gene transcription within one hour (Fig. 7B). Although both Cases 1 and 3 exhibited hyperacetylation of 5'-LTR, *tax* gene transcription was silenced in Case 3.

Discussion

DNA methylation is regarded as a host defense mechanism for inactivating transportable elements such as retroviruses to inhibit viral transcription and the generation of new viruses. On the other hand, it also renders the provirus into a latent state, resulting in the establishment of latent infection. However, it remained unclear how and when the provirus was methylated, and whether DNA methylation changed *in vivo*.

Tax has the remarkable potency to promote the proliferation of infected cells [3], however, it is also a major target of CTL *in vivo* [8]. Therefore, HTLV-I controls *tax* gene expression by own viral proteins, Rex [19], p30 [20,21] and HBZ [22]. In the leukemic cells, several mechanisms have been identified to suppress or abolish Tax expression, including genetic changes of *tax* gene, deletion of 5'-LTR, and DNA methylation of 5'-LTR. In this study, DNA methylation was shown to occur in internal provirus sequences, such as the *gag*, *pol* and *env* regions, and then extend to 5' (5'-LTR) and 3' (pX) regions. Since DNA methylation of 5'-LTR is associated with *tax* gene transcription, the finding that 5'-LTR was more highly methylated in ATL cells than in carriers, among cases with methylated 5'-LTR, suggests that such HTLV-I-infected cells and ATL cells with the methylated provirus, which produce lower amounts of viral proteins, are selected *in vivo* by the host immune system. In this regard, HTLV-I is quite different from another human retrovirus, HIV-1. HIV-1 vectors were resistant to gene silencing *in vivo* [23,24]. It is noteworthy that the number of CpG sites in the U3 region of HIV-1 LTR (9 sites in LTR of NL43) is much fewer than that of HTLV-I (47 sites in LTR of ATK). This is consistent to the previous report that transcriptional suppression was not associated with DNA methylation of HIV-1 provirus [25]. In addition, HIV-1 provirus is frequently integrated within transcriptional units, which encode the genes that are transcribed in T-cells [15,26]. In such regions, it is possible that HIV-1 tends to escape from transcriptional silencing that is observed in the heterochromatin region such as alphoid repetitive sequences [18]. These data suggest that HIV-1 is more resistant to gene silencing than HTLV-I. Alternatively, it is possible that HTLV-I takes advantage of susceptibility to DNA methylation to escape from the host immune system.

This study shows that 3'-LTR is unmethylated in carriers and ATL cells while 5'-LTR is methylated in about half of cases. In HTLV-I, *HTLV-I bZIP* (*HBZ*) *gene* is encoded by minus strand of provirus [22,27]. We observed that *HBZ* gene was transcribed in all ATL cells, suggesting that *HBZ* gene play a critical role in growth of HTLV-I infected cells and ATL cells (submitted for publication). The finding that 3'-LTR is unmethylated in all ATL cases and carriers suggests that *HBZ* gene transcription is important for proliferation of ATL and HTLV-I infected cells.

Why does DNA methylation occur from the internal sequences of the HTLV-I provirus? Since CpG island is recognized as DNA region that is susceptible to DNA methylation, we analyzed HTLV-I provirus by the criterion by Takai and Jones [28]. CpG islands are present throughout the provirus in 5'-LTR-*gag* (1–1360), *pol* (3876–4509), *env* (5648–6166), *env*-pX (6446–7561), and pX-3'-LTR (8212–9045) regions. Therefore, the presence of CpG island could not explain why DNA methylation occurred in the internal region of HTLV-I provirus. Among tumor-suppressor genes, which are transcriptionally silenced by DNA methylation, the exon regions are first methylated, and then DNA methylation progresses to the promoter region [29]. When the promoter region is heavily methylated, the transcription of the corresponding gene is silenced. Since 5'-LTR is the promoter/enhancer for viral gene transcription, there might be a similar scenario between the exon/promoter and DNA methylation in both virus and tumor-suppressor genes. Thus, it is possible that gene coding regions are first methylated and DNA methylation spreads to the promoter region of provirus, 5'-LTR.

Transcriptional silencing of *tax* gene in spite of hyperacetylated histone H3 is recognized as another mechanism to suppress the viral gene transcription in addition to DNA methylation. The prompt recovery of *tax* gene expression after *in vitro* culture suggests the presence of an inhibitory factor(s) that binds to 5'-LTR, and suppresses the viral gene transcription *in vivo*. It is noteworthy that this phenotype is very similar to that of a mouse T-cell line transfected with an HTLV-I LTR-derived reporter plasmid [30]. In that study, a green fluorescent protein-fused Tax (*Gax*) gene was transfected into a mouse T-cell line, EL-4, and the transduced cells were then injected into Tax-immunized and non-immunized mice. Although Tax-induced cytotoxic T-cells suppressed the expression of the *Gax* gene *in vivo*, its expression was shown to recover within three hours when the transduced cells were transferred to *in vitro* culture. This phenotype resembles that observed in Case 3 in Fig. 7. Considering that Tax is the major target of CTL *in vivo*, and at the same time, confers growth advantages on the infected cells, such reversible suppression of *tax* gene expression is thought to be suitable for the survival of HTLV-I infected cells, and ATL cells. In this regard, potentiation of anti-Tax immunity might protect against the development of ATL when combined with possible therapeutics to induce Tax expression [31]. For this purpose, the mechanism for silencing viral transcription regardless of histone H3 hyperacetylation should be studied.

In general, gene silencing is associated with several different mechanisms. DNA methylation in the promoter region silences the gene transcription, whereas gene silencing is often not associated with DNA methylation [32,33]. In such situations, methylation of H3K9 is linked with loss of transcriptions [34]. It is possible that silencing of viral gene transcription renders proviral DNA vulnerable to methylation. Once proviral DNA is methylated, such silencing would be fixed unless such cells are treated with demethylating agents such as 5-aza-deoxy-cytidine.

DNA methylation of the HTLV-I provirus did not accumulate in a cell line that was cultured *in vitro* for more than 9 years. The finding that the *p16* gene was heavily methylated in this cell line excluded the possibility that hypermethylation did not occur in this cell line due to aberrant methylation machinery. Among the seroconverters, the provirus was heavily methylated in internal regions such as *gag*, *pol* and *env*. Taken together, DNA methylation in the provirus is considered to reflect the selection *in vivo*. Since the growth of *in vitro* HTLV-I-transformed cell lines depends on Tax expression, cells with suppressed expression of the *tax* gene do not have the growth advantage *in vitro*. However, the immune system exerts selection of the infected cells with suppressed *tax* gene expression *in vivo*.

Recently, both 5'- and 3'-LTR have been reported to be transcriptionally active, and transcriptional factors and Tax bind equally to both [35]. 3'-LTR may activate the transcription of cellular genes, which are located in the downstream of integration sites. In addition, unmethylated 3'-LTR is critical for transcription of the *HBZ* gene. Since 5'-LTR is a promoter/enhancer for viral gene transcription, selective methylation of 5'-LTR is considered to silence the transcription of viral genes.

Conclusion

We have demonstrated how DNA methylation of HTLV-I provirus occurred, and how it suppressed viral gene transcription. When 5'-LTR was heavily methylated, viral transcription was silenced, which is thought to reflect the immune system selection *in vivo*. In addition, mechanisms other than DNA methylation suppresses viral gene transcription regardless of histone H3 hyperacetylation. The mechanism of such suppression requires further investigation.

Methods

Cells

HTLV-I-associated cell lines (MT-1, MT-2, MT-4, ATL-2, TL-Oml and Sez627) were cultured in RPMI1640 medium supplemented with 10% fetal bovine serum and penicillin/streptomycin. For interleukin-2-dependent cell lines (ATL-43T, 48T and 55T), 100 U/ml of recombinant interleukin-2 (Shionogi, Osaka) was added to the medium. Peripheral blood mononuclear cells (PBMC) or lymph node cells were isolated from HTLV-I carriers and ATL patients after informed consent was obtained. The polyclonal integration of HTLV-I provirus in carriers has been shown by inverse PCR [36], and provirus load was determined by real-time PCR as reported previously [37].

Sodium bisulfite treatment of genomic DNA

Sodium bisulfite treatment was performed as described previously [29]. Briefly, 1–3 μg of genomic DNA was denatured in 0.3 N NaOH at 37°C for 15 min, and 1 μg of salmon sperm DNA was added to each sample as a carrier. Sodium bisulfite (pH 5.0) and hydroquinone were added to each sample to final concentrations of 3 M and 0.05 mM, respectively. The reaction was performed at 55°C for 16 h and the samples were then desalted using the Wizard DNA Clean-Up System (Promega, Madison, WI). Finally, samples were desulfonated in 0.3 N NaOH at 37°C for 15 min.

Sequencing of sodium bisulfite-treated genomic DNA

The sodium bisulfite-treated DNA (200–500 ng) was used as a template for PCR amplification of eight HTLV-I provirus regions. The PCR reactions were performed using FastStart Taq DNA Polymerase (Roche, Mannheim, Germany). The PCR primer pairs and annealing temperatures

Table 1: Primer sets for COBRA and ChIP assay

	Site in HTLV-I[a]		Forward primer	Reverse primer	Anneal (°C)	Enzyme for COBRA
COBRA	620	1st	5'-TTTGGAGTTTATTTAGATTTAG-3'	5'-CCAATAATAAACRACCAACCC-3'	45	*Taq*I
	(5'-LTR)	2nd	5'-GTTTTGTTTGATTTTGTTTGT-3'	5'-AAAAAAATTTAACCCATTACC-3'	49	
	1753	1st	5'-GGGAGTGTTAAAGATTTTTTTTGGG-3'	5'-ACTCCAATAACCTACTTTCCC-3'	55	*Taq*I
	(*gag*)	2nd	5'-TTTATTTTTTAAGGTTTGGAGGAG-3'	5'-TTAAAAATCCAAATCTAACAAACCC-3'	55	
	2988	1st	5'-GTTAAAAAGGTTAATGGAATTTGG-3'	5'-CCTCTAAAAATAATAATAAATCCTC-3'	52	*Taq*I
	(*pol*)	2nd	5'-GGGTTTTTTGATTTGTTTAGTTTG-3'	5'-AAACTTACTAAAAAAATATCATCC-3'	51	
	4187	1st	5'-GGGTGAAATTGTGTAGTTTTGTAGG-3'	5'-CCTATTTTCAAACGAATCTACCTCC-3'	57	*Acc*II
	(*pol*)	2nd	5'-GTGATTAGTAGGGTATTTGTGAGAG-3'	5'-ATTATCACAAAAATCATTCCCCC-3'	52	
	5151	1st	5'-GGTATTATTTTAAGTTTTTTGG-3'	5'-CTCCAATTATAAAAATACAACAAC-3'	46	*Taq*I
	(*pol*)	2nd	5'-GTTAGTGGAAAGGATTATAGGAGG-3'	5'-AACTTACCCATAATATTAAAAATC-3'	51	
	6113	1st	5'-GGATTTATTGTTTTGATTTTTAG-3'	5'-CTTTACATAATCCTCCTTACTCCC-3'	51	*Taq*I
	(*env*)	2nd	5'-GGATTTATTGTTTTGATTTTTAG-3'	5'-CCCAAAACAAAAAATCAAAACC-3'	53	
	7258	1st	5'-GAGGTGGYGTTTTTTTTTTTGG-3'	5'-CCTTAAAAATCTTAAAAATTCTC-3'	47	*Taq*I
	(pX)	2nd	5'-AAGGATAGTAAATYGTTAAGTATAG-3'	5'-CCCAAATAATCTAATACTCTAAAC-3'	50	
	8342	1st	5'-YGATGGTAYGTTTATGATTTTYGGG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	57	*Taq*I
	(3'-LTR)	2nd	5'-YGATGGTAYGTTTATGATTTTYGGG-3'	5'-AACTCCTACTAATTTATTAAACC-3'	52	
	5'-LTR[b]		5'-GCTTTGCCTGACCCTGCTTGC-3'	5'-AAGATTTGGCCCATTGCCTAGGG-3'	63	
	env		5'-TGCCAGCCTCTCCACTTGGCACG-3'	5'-ATGGAGCCGGTAATCCCGCCAGC-3'	64	
	pX		5'-AAGGATAGCAAACCGTCAAGCACAG-3'	5'-CCCAGGTGATCTGATGCTCTGGAC-3'	63	
	3'-LTR		5'-CCCCTCATTTCTACTCTCACACGGC-3'	5'-TGGGTGGTTCTTGGTGGCTTCCC-3'	64	

[a] Nucleotide position corresponding to that of ATK. This number means the cytidine of CpG sites analysed.
[b] For ChIP assay, we used primers to amplify the indicated regions.

are shown in Table 1. The amplified PCR products were purified and subcloned into pGEM-T Easy vectors (Promega). For each region, at least 10 clones were sequenced using Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied BioSystems, Foster City, CA) and ABI3100 autosequencer (Applied Biosystems).

Combined bisulfite restriction analysis (COBRA)

For COBRA, eight different regions of HTLV-I provirus were amplified with sodium bisulfite treated genomic DNAs using each primer sets as shown in Table 1. The nested PCR reactions were performed using FastStart Taq DNA Polymerase (Roche) with the following condition: 5 minutes at 95°C for denaturation, 40 cycles of 30 sec at 95°C, 30 sec at each annealing temperature (Table 1), 30 sec at 72°C, and 2 min at 72°C for final extension. The PCR products were digested for at least 4 hrs with an appropriate restriction enzyme (TaqI or AccII) that had a single recognition site within each product [38]. When CpG site within amplified region was methylated, it was resistant to sodium bisulfite treatment, resulting in digestion by these enzymes. On the other hand, since unmethylated CpG was converted to UG by sodium bisulfite treatment, these enzymes could not digest the amplified DNAs. The digested PCR products were separated in a 3% Nusieve 3:1 agarose (BMA, Rockland, ME) gel. The intensity of each fragment was determined using ATTO Densitograph Ver. 4.0 (ATTO, Tokyo, Japan), and the extent of DNA methylation was calculated as follows: % methylation = 100 × (digested PCR products/undigested+ digested PCR products).

Southern blot analyses

To determine the number of integrated HTLV-I provirus, we performed Southern blot method using HTLV-I probe as described previously [10]. In brief, 5 μg of DNA were digested with EcoRI, separated by electrophoresis in a 0.7% agarose gel, and transferred to nylon membrane (Hybond N+, Amersham Biosciences, Piscataway, NJ). The membrane was hybridized to the alkaline phospatase labeled pX probes. 0.9 kb PCR product of HTLV-I pX region derived from HTLV-I clone λ23-3 was used as probe [39]. DNA probe was labeled, and hybridized to the membrane with Gene Images AlkPhos Direct Labelling and Detection system (Amersham Biosciences).

Inverse-long PCR

To check the HTLV-I integration in PBMCs of carriers, we analyzed the genomic DNAs from carriers by inverse-long PCR method as described previously [36]. In brief, genomic DNA was digested with EcoRI, and then ligated with T4 DNA ligase. Circularized DNA was digested with MluI that cut the provirus at pX region to prevent amplification of provirus itself. Then, treated genomic DNA was amplified with primers as follows: Long-IPCR-F: 5'-TGCCTGACCCTGCTTGCTCAACTCTACGTCTTTG-3', Long-IPCR-R 5'-AGTCTGGGCCCTGACCTTTTCAGACTTCTGTTTC-3'. PCR condition was as follows: 2 min at 98°C for denaturation, 5 cycles (30 sec at 98°C, 10 min at 64°C), followed by 35cycles (30 sec at 94°C, 10 min at 64°C) and 15 min at 72°C for final extension. The PCR products were subcloned into plasmid DNA and their sequences were determined.

Table 2: Primer sets and annealing temperatures for genome specific PCR

	Case	Locus		Forward primer	Reverse primer	Anneal (°C)
Primers for case	Acute ATL 1	5q11.1	1st	5'-TTTGGAGAGGGAATTTTATATTG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	55
			2nd	5'-GGAGTGTAGAGATGTAGTTTTGG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	50
	Acute ATL 2	8p23.1	1st	5'-GAGAAATTTGTGTTGATTTTATTAG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	47
			2nd	5'-TTAGTGGTAGATTAAGTTAAAG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	45
	Acute ATL 3	1q31.1	1st	5'-GGTAGAAATTATAGGTTTTTGTAGG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	51
			2nd	5'-GTTATTTGTGAAGTAAGATGTTTTG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	53
	Acute ATL 21	15q24.3	1st	5'-GAGGTGGATTTTTATTTTATTG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	52
			2nd	5'-GGTTTTTGATTATATTTGGGGAG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	54
	Acute ATL 22	19q13.11	1st	5'-GTTAGTTGTTAGAGAGTTTTTTGG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	52
			2nd	5'-AAGATTATTTAGTTTTTTGGGG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	54
	Chronic ATL 1	1p22.1	1st	5'-GGGTTTGAAGTTTTTTTTGTAGG-3'	5'-ACCCCCTCCTAAACTATCTCC-3'	53
			2nd	5'-AAGATTATTTAGTTTTTTGGGG-3'	5'-ACCCCCTCCTAAACTATCTCC-3' (5'-LTR U3)	50
Primers for human genome		5q11.1	1st	5'-TTTGGAGAGGGAATTTTATATTG-3'	5'-CCCAAACTAATCTTCAACTCC-3'	52
			2nd	5'-GGAGTGTAGAGATGTAGTTTTGG-3'	5'-CCACCATAAAAAACCCTCCC-3'	54
		8p23.1	1st	5'-GAGAAATTTGTGTTGATTTTATTAG-3'	5'-AATATCACTATAACAATAACCAC-3'	46
			2nd	5'-TTAGTGGTAGATTAAGTTAAAG-3'	5'-CTCTCAACAAATTCCATCTTTCC-3'	49
		1q31.1	1st	5'-GGTAGAAATTATAGGTTTTTGTAGG-3'	5'-CACCATTAAACAAACTAAATTCTC-3'	51
			2nd	5'-GTTATTTGTGAAGTAAGATGTTTTG-3'	5'-CACATAAAAAAACCCACACAATC-3'	53
		15q24.3	1st	5'-GAGGTGGATTTTTATTTTATTG-3'	5'-ATCTACCTAAAAAACCCACCC-3'	52
			2nd	5'-GGTTTTTGATTATATTTGGGGAG-3'	5'-AAAAACCCACCCAAACAAACC-3'	57
		19q13.11	1st	5'-GTTAGTTGTTAGAGAGTTTTTTGG-3'	5'-CAACTCCCTAAACCCTCCTCC-3'	52
			2nd	5'-GTTTTTTGGTTAAGGTTATGGG-3'	5'-CTCCTACCACGAACCTACTCC-3'	54
		1p22.1	1st	5'-GGGTTTGAAGTTTTTTTTGTAGG-3'	5'-CAACAAAAACAATAAACAAAACC-3'	54
			2nd	5'-AAGATTATTTAGTTTTTTGGGG-3'	5'-CTTTACACCAATAAATTTAATACC-3'	50

DNA methylation in neighboring regions of HTLV-I integration sites

The integration sites of HTLV-I provirus has been determined by inverse long PCR, and DNA methylation of genomic DNAs neighboring integration sites was determined in both ATL cells and PBMCs. The nested PCR reactions were performed using FastStart Taq DNA Polymerase (Roche) with the following condition: 5 minutes at 95°C for denaturation, 40 cycles of 30 sec at 95°C, 30 sec at each annealing temperature (Table 2), 30 sec at 72°C, and 2 min at 72°C for final extension.

RT-PCR

Total RNA was isolated from PBMCs or lymph node cells using TRIzol Reagent (Invitrogen, Carlsbad, CA) and RT-PCR was performed using RNA LA PCR Kit (AMV) Ver. 1.1 (Takara Bio Inc., Otsu, Japan) according to the manufacturer's protocol. The tax and GAPDH gene transcripts were amplified using the following primers: RPX2 5'-CCGGCGCTGCTCTCATCCCGGT-3' and RPX5 5'-GGCCGAACATAGTCCCCCAGAG-3' (for tax), GAPDH1 5'-ATGGGGAAGGTGAAGGTCGGAGTC-3' and GAPDH1a 5'-CCATGCCAGTGAGCTTCCCGTTC-3' (for GAPDH) under following conditions: 2 minutes at 95°C for denaturation, 35 cycles of 30 sec at 95°C, 30 sec at 62°C, 30 sec at 72°C (for tax), 25 cycles of 30 sec at 95°C, 30 sec at 55°C, 30 sec at 72°C (for GAPDH) and 2 min at 72°C for final extension.

Chromatin immunoprecipitation (ChIP) assay

ChIP assays were performed as described previously [40]. Briefly, ATL cell lines and fresh ATL cells from ATL patients (5×10^5 cells/antibody) were fixed with formaldehyde and then sonicated to obtain soluble chromatin. The chromatin solutions were immunoprecipitated with anti-acetyl-Histone H3 or anti-acetyl-Histone H4 (Upstate Biotechnology), or normal rabbit IgG, overnight at 4°C, and the immunoprecipitates were then collected with 50% protein A and G-Sepharose slurry preabsorbed with 0.1 mg/ml sonicated salmon sperm DNA. The resulting purified DNAs were subjected to PCR reactions using primer sets specific for 5'-LTR, env, pX and 3'-LTR. The sequences of the primers are shown in Table 1. To distinguish 5' and 3'-LTR, we used primers specific for gag and R region of LTR for amplification of 5'-LTR, and primers for pX region and U3 region were used for amplification of 3'-LTR. The PCR reactions were performed using FastStart Taq DNA Polymerase (Roche) with the following condition: 5 minutes at 95°C, 35 or 37 cycles of 30 sec at 95°C, 30 sec at each annealing temperature (Table 1), 30 sec at 72°C, and 2 min at 72°C. The PCR products were electrophoresed in an agarose gel and the results were analyzed using ATTO Densitograph Ver. 4.0. Values were calculated as the signal intensity of each sample normalized by that of the whole cell extract.

Statistical analyses

Statistical analyses were performed using the Mann-Whitney's U-test and Student's t-test.

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

YT conceived this project and carriers out most of experiments in Figs. 1, 2, 3, 5 and 6. KN established COBRA assay and performed experiments in Figs. 1 and 2. JY performed experiments in Fig. 7. MM established most of HTLV-I transformed cell lines, and analyzed experiments in Fig. 4. AO and NM provided sequential DNA samples from seroconverters, and analyzed the data. M.Matsuoka directed and supervised the experiments and interpretations All authors read and approved the final manuscript.

Acknowledgements

We thank Shinjiro Hino for valuable suggestions.

References

1. Takatsuki K, Yamaguchi K, Matsuoka M: ATL and HTLV-I-related diseases. In *Adult T-cell leukemia* Edited by: Takatsuki K. New York: Oxford University Press; 1994:1-27.
2. Matsuoka M: **Human T-cell leukemia virus type I and adult T-cell leukemia.** *Oncogene* 2003, **22(33):**5131-5140.
3. Yoshida M: **Multiple viral strategies of HTLV-I for dysregulation of cell growth control.** *Annu Rev Immunol* 2001, **19:**475-496.
4. Franchini G, Fukumoto R, Fullen JR: **T-cell control by human T-cell leukemia/lymphoma virus type 1.** *Int J Hematol* 2003, **78(4):**280-296.
5. Jin DY, Spencer F, Jeang KT: **Human T cell leukemia virus type I oncoprotein Tax targets the human mitotic checkpoint protein MAD1.** *Cell* 1998, **93(1):**81-91.
6. Jin DY, Giordano V, Kibler KV, Nakano H, Jeang KT: **Role of adapter function in oncoprotein-mediated activation of NF-kappaB. Human T-cell leukemia virus type I Tax interacts directly with IkappaB kinase gamma.** *J Biol Chem* 1999, **274(25):**17402-17405.
7. Jacobson S, Shida H, McFarlin DE, Fauci AS, Koenig S: **Circulating CD8+ cytotoxic T lymphocytes specific for HTLV-I pX in patients with HTLV-I associated neurological disease.** *Nature* 1990, **348(6298):**245-248.
8. Bangham CR: **Human T-lymphotropic virus type I (HTLV-I): persistence and immune control.** *Int J Hematol* 2003, **78(4):**297-303.
9. 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(7):**346-352.
10. Tamiya S, Matsuoka M, Etoh K, Watanabe T, Kamihira S, Yamaguchi K, Takatsuki K: **Two types of defective human T-lymphotropic virus type I provirus in adult T-cell leukemia.** *Blood* 1996, **88(8):**3065-3073.
11. 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(4):**987-993.
12. Koiwa T, Hamano-Usami A, Ishida T, Okayama A, Yamaguchi K, Kamihira S, Watanabe T: **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(18):**9389-9397.
13. Takeda S, Maeda M, Morikawa S, Taniguchi Y, Yasunaga J, Nosaka K, Tanaka Y, Matsuoka M: **Genetic and epigenetic inactivation of tax gene in adult T-cell leukemia cells.** *Int J Cancer* 2004, **109(4):**559-567.
14. Verma M: **Viral genes and methylation.** *Ann N Y Acad Sci* 2003, **983:**170-180.
15. Han Y, Lassen K, Monie D, Sedaghat AR, Shimoji S, Liu X, Pierson TC, Margolick JB, Siliciano RF, Siliciano JD: **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(12):**6122-6133.
16. Doi K, Wu X, Taniguchi Y, Yasunaga JI, Satou Y, Okayama A, Nosaka K, Matsuoka M: **Preferential selection of human T-cell leukemia virus type I (HTLV-I) provirus integration sites in leukemic versus carrier states.** *Blood* 2005.
17. Seiki M, Eddy R, Shows TB, Yoshida M: **Nonspecific integration of the HTLV provirus genome into adult T-cell leukaemia cells.** *Nature* 1984, **309(5969):**640-642.
18. Jordan A, Bisgrove D, Verdin E: **HIV reproducibly establishes a latent infection after acute infection of T cells in vitro.** *Embo J* 2003, **22(8):**1868-1877.
19. Seiki M, Inoue J, Hidaka M, Yoshida M: **Two cis-acting elements responsible for posttranscriptional trans-regulation of gene expression of human T-cell leukemia virus type I.** *Proc Natl Acad Sci U S A* 1988, **85(19):**7124-7128.
20. Zhang W, Nisbet JW, Albrecht B, Ding W, Kashanchi F, Bartoe JT, Lairmore MD: **Human T-lymphotropic virus type I p30(II) regulates gene transcription by binding CREB binding protein/p300.** *J Virol* 2001, **75(20):**9885-9895.
21. Nicot C, Dundr M, Johnson JM, Fullen JR, Alonzo N, Fukumoto R, Princler GL, Derse D, Misteli T, Franchini G: **HTLV-I-encoded p30II is a post-transcriptional negative regulator of viral replication.** *Nat Med* 2004, **10(2):**197-201.
22. Gaudray G, Gachon F, Basbous J, Biard-Piechaczyk M, Devaux C, Mesnard JM: **The complementary strand of the human T-cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription.** *J Virol* 2002, **76(24):**12813-12822.
23. Miyoshi H, Smith KA, Mosier DE, Verma IM, Torbett BE: **Transduction of human CD34+ cells that mediate long-term engraftment of NOD/SCID mice by HIV vectors.** *Science* 1999, **283(5402):**682-686.
24. Gatlin J, Padgett A, Melkus MW, Kelly PF, Garcia JV: **Long-term engraftment of nonobese diabetic/severe combined immunodeficient mice with human CD34+ cells transduced by a self-inactivating human immunodeficiency virus type 1 vector.** *Hum Gene Ther* 2001, **12(9):**1079-1089.
25. Pion M, Jordan A, Biancotto A, Dequiedt F, Gondois-Rey F, Rondeau S, Vigne R, Hejnar J, Verdin E, Hirsch I: **Transcriptional suppression of in vitro-integrated human immunodeficiency virus type 1 does not correlate with proviral DNA methylation.** *J Virol* 2003, **77(7):**4025-4032.
26. 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(4):**521-529.
27. Larocca D, Chao LA, Seto MH, Brunck TK: **Human T-cell leukemia virus minus strand transcription in infected T-cells.** *Biochem Biophys Res Commun* 1989, **163(2):**1006-1013.
28. Takai D, Jones PA: **Comprehensive analysis of CpG islands in human chromosomes 21 and 22.** *Proc Natl Acad Sci U S A* 2002, **99(6):**3740-3745.
29. Nosaka K, Maeda M, Tamiya S, Sakai T, Mitsuya H, Matsuoka M: **Increasing methylation of the CDKN2A gene is associated with the progression of adult T-cell leukemia.** *Cancer Res* 2000, **60(4):**1043-1048.
30. Furuta RA, Sugiura K, Kawakita S, Inada T, Ikehara S, Matsuda T, Fujisawa J: **Mouse model for the equilibration interaction between the host immune system and human T-cell leukemia virus type I gene expression.** *J Virol* 2002, **76(6):**2703-2713.
31. Hanabuchi S, Ohashi T, Koya Y, Kato H, Hasegawa A, Takemura F, Masuda T, Kannagi M: **Regression of human T-cell leukemia virus type I (HTLV-I)-associated lymphomas in a rat model: peptide-induced T-cell immunity.** *J Natl Cancer Inst* 2001, **93(23):**1775-1783.
32. Soengas MS, Capodieci P, Polsky D, Mora J, Esteller M, Opitz-Araya X, McCombie R, Herman JG, Gerald WL, Lazebnik YA, *et al.*: **Inactivation of the apoptosis effector Apaf-1 in malignant melanoma.** *Nature* 2001, **409(6817):**207-211.
33. Yasunaga J, Taniguchi Y, Nosaka K, Yoshida M, Satou Y, Sakai T, Mitsuya H, Matsuoka M: **Identification of aberrantly methylated genes in association with adult T-cell leukemia.** *Cancer Res* 2004, **64(17):**6002-6009.
34. Bachman KE, Park BH, Rhee I, Rajagopalan H, Herman JG, Baylin SB, Kinzler KW, Vogelstein B: **Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene.** *Cancer Cell* 2003, **3(1):**89-95.

35. Lemasson I, Polakowski NJ, Laybourn PJ, Nyborg JK: **Transcription regulatory complexes bind the human T-cell leukemia virus 5' and 3' long terminal repeats to control gene expression.** *Mol Cell Biol* 2004, **24(14)**:6117-6126.
36. Etoh K, Tamiya S, Yamaguchi K, Okayama A, Tsubouchi H, Ideta T, Mueller N, Takatsuki K, Matsuoka M: **Persistent clonal proliferation of human T-lymphotropic virus type I-infected cells in vivo.** *Cancer Res* 1997, **57(21)**:4862-4867.
37. Yasunaga J, Sakai T, Nosaka K, Etoh K, Tamiya S, Koga S, Mita S, Uchino M, Mitsuya H, Matsuoka M: **Impaired production of naive T lymphocytes in human T-cell leukemia virus type I-infected individuals: its implications in the immunodeficient state.** *Blood* 2001, **97(10)**:3177-3183.
38. Xiong Z, Laird PW: **COBRA: a sensitive and quantitative DNA methylation assay.** *Nucleic Acids Res* 1997, **25(12)**:2532-2534.
39. Clarke MF, Gelmann EP, Reitz MS Jr: **Homology of human T-cell leukaemia virus envelope gene with class I HLA gene.** *Nature* 1983, **305(5929)**:60-62.
40. Hino S, Fan J, Taguwa S, Akasaka K, Matsuoka M: **Sea urchin insulator protects lentiviral vector from silencing by maintaining active chromatin structure.** *Gene Ther* 2004, **11(10)**:819-828.
41. Seiki M, Hattori S, Hirayama Y, Yoshida M: **Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA.** *Proc Natl Acad Sci U S A* 1983, **80(12)**:3618-3622.

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)

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 κB (NF-κB), 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.

provirus. First, genomic DNA (1.5 μg) was digested with a restriction enzyme (*Hin*dIII, *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 μM), $MgCl_2$ (2.5 mM), and deoxynucleoside triphosphates (dNTPs, 0.4 mM) were mixed (total 20 μL) 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 μg), 10 × LA buffer (5 μL), and LA Taq (0.4 μL) were added (total 30 μL). 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 *Hin*dIII-digested samples, primer 3: 5′-TGGGCAGGATTGCAGGGTTTAGAGTGG-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′-TCATTCACGACTGACTGCCGG-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 ×) 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-1 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-1 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 alphoid 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 alphoid sequences in ATL (2/59; 3.4%). Since alphoid sequences were estimated to compose 3% to 5% of the human genome, we used 5% for the statistical

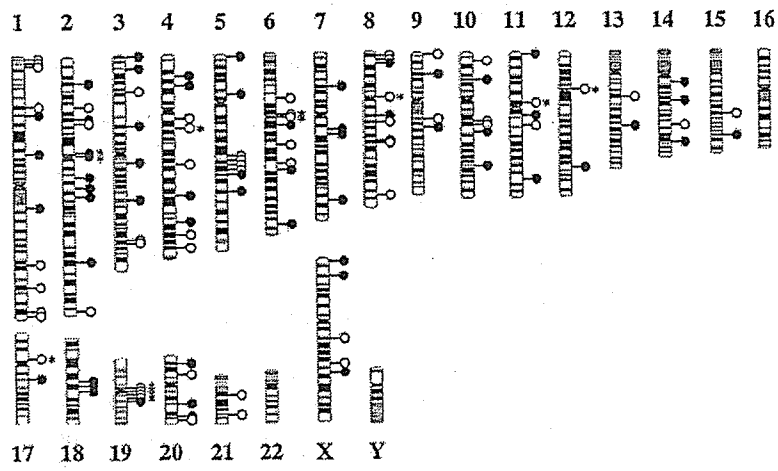

Figure 1. Distribution of HTLV-I provirus integration sites in human chromosomes. The integration sites of HTLV-I provirus in ATL patients (●) and HTLV-I carriers (○) are demonstrated. Integration within alphoid repetitive sequence is marked by an asterisk.

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 alphoid 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 alphoid 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 alphoid repetitive sequences, whereas the frequency of integration into alphoid sequences was significantly decreased in leukemic cells. When HTV-1 provirus is integrated into alphoid 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.	P	No.	P	no.
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.

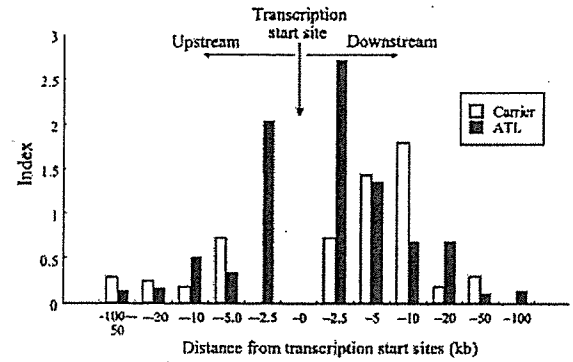

Figure 2. Relationship between HTLV-I provirus integration sites and transcriptional start sites. The distances from the transcriptional start sites are shown. The index was defined by the following calculation: (observed number of integrations) × 100 / (total number) × (window size).

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 alphoid 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 alphoid 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 alphoid 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†
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 alphoid 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 alphoid 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

1. Hayward WS, Neel BG, Astrin SM. Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. Nature. 1981;290: 475-480.
2. 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.
3. 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.
4. 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.
5. 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.
6. Matsuoka M. Human T-cell leukemia virus type I and adult T-cell leukemia. Oncogene. 2003;22: 5131-5140.
7. 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.
8. 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.
9. 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.
10. 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.
11. 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.
12. 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.
13. 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.
14. Kent WJ. BLAT—the BLAST-like alignment tool. Genome Res. 2002;12:656-664.
15. 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.
16. Ohshima K, Ohgami A, Matsuoka M, et al. Random integration of HTLV-1 provirus: increasing chromosomal instability. Cancer Lett. 1998;132: 203-212.
17. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409:860-921.
18. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. Science. 2001; 291:1304-1351.
19. 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.
20. 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.
21. 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.
22. 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.
23. Jordan A, Bisgrove D, Verdin E. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. EMBO J. 2003;22:1868-1877.
24. 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. Predominant 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, Gallay 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.

Retrovirology

BioMed Central

Review

Open Access

Human T-cell leukemia virus type I (HTLV-I) infection and the onset of adult T-cell leukemia (ATL)

Masao Matsuoka*

Address: Institute for Virus Research, Kyoto University, Kyoto 606-8507, Japan

Email: Masao Matsuoka* - mmatsuok@virus.kyoto-u.ac.jp

* Corresponding author

Published: 26 April 2005

Retrovirology 2005, 2:27 doi:10.1186/1742-4690-2-27

Received: 29 March 2005

Accepted: 26 April 2005

This article is available from: http://www.retrovirology.com/content/2/1/27

Abstract

The clinical entity of adult T-cell leukemia (ATL) was established around 1977, and human T-cell leukemia virus type I (HTLV-I) was subsequently identified in 1980. In the 25 years since the discovery of HTLV-I, HTLV-I infection and its associated diseases have been extensively studied, and many of their aspects have been clarified. However, the detailed mechanism of leukemogenesis remains unsolved yet, and the prognosis of ATL patients still poor because of its resistance to chemotherapy and immunodeficiency. In this review, I highlight the recent progress and remaining enigmas in HTLV-I infection and its associated diseases, especially ATL.

Background

In 1977, Takatsuki et al. reported adult T-cell leukemia (ATL) as a distinct clinical entity [1-3]. This disease is characterized by its aggressive clinical course, infiltrations into skin, liver, gastrointestinal tract and lung, hypercalcemia and the presence of leukemic cells with multilobulated nuclei (flower cell)(Figure 1). In 1980, Poiesz et al. discovered a human retrovirus in a cell line derived from a patient with ATL, and designated it human T-cell leukemia virus type I (HTLV-I) [4,5]. The linkage between ATL and HTLV-I was proven by Hinuma et al., who demonstrated the presence of an antibody against HTLV-I in patient sera [6]. Thereafter, Seiki et al. determined the whole sequence of HTLV-I and revealed the presence of a unique region, designated pX [7]. The pX region encodes several accessory genes, which control viral replication and the proliferation of infected cells [8]. In this review, I describe the recent advances in the field of HTLV-I and ATL research, with particular focus on the mechanism of leukemogenesis and therapeutic aspects.

1. History of humans and HTLV-I

HTLV-I is a member of the Deltaretroviruses, which include HTLV-II, bovine leukemia virus and simian T-cell leukemia virus (STLV). The latter two viruses also cause lymphoid malignancies in the host, similar to the case with HTLV-I. HTLV and STLV are thought to originate from common ancestors, and share molecular, virological and epidemiological features. Therefore, they have been designated primate T-cell leukemia viruses (PTLVs). Phylogenetical analyses have revealed that HTLV-Ic first diverged from simian leukemia virus around 50,000 ± 10,000 years ago, while the spread of PTLV-I in Africa is estimated to have occurred at least 27,300 ± 8,200 years ago. Subsequently, HTLV-Ia, which is the most common subtype in Japan, diverged from the African strain 12,300 ± 4,900 years ago [9]. Thus, these viruses have had a long history with humans after the interspecies transmission. In contrast, human immunodeficiency virus type 1 (HIV-1) is thought to originate from simian immunodeficiency virus in chimpanzees (SIV_{CPZ}) [10], and the interspecies

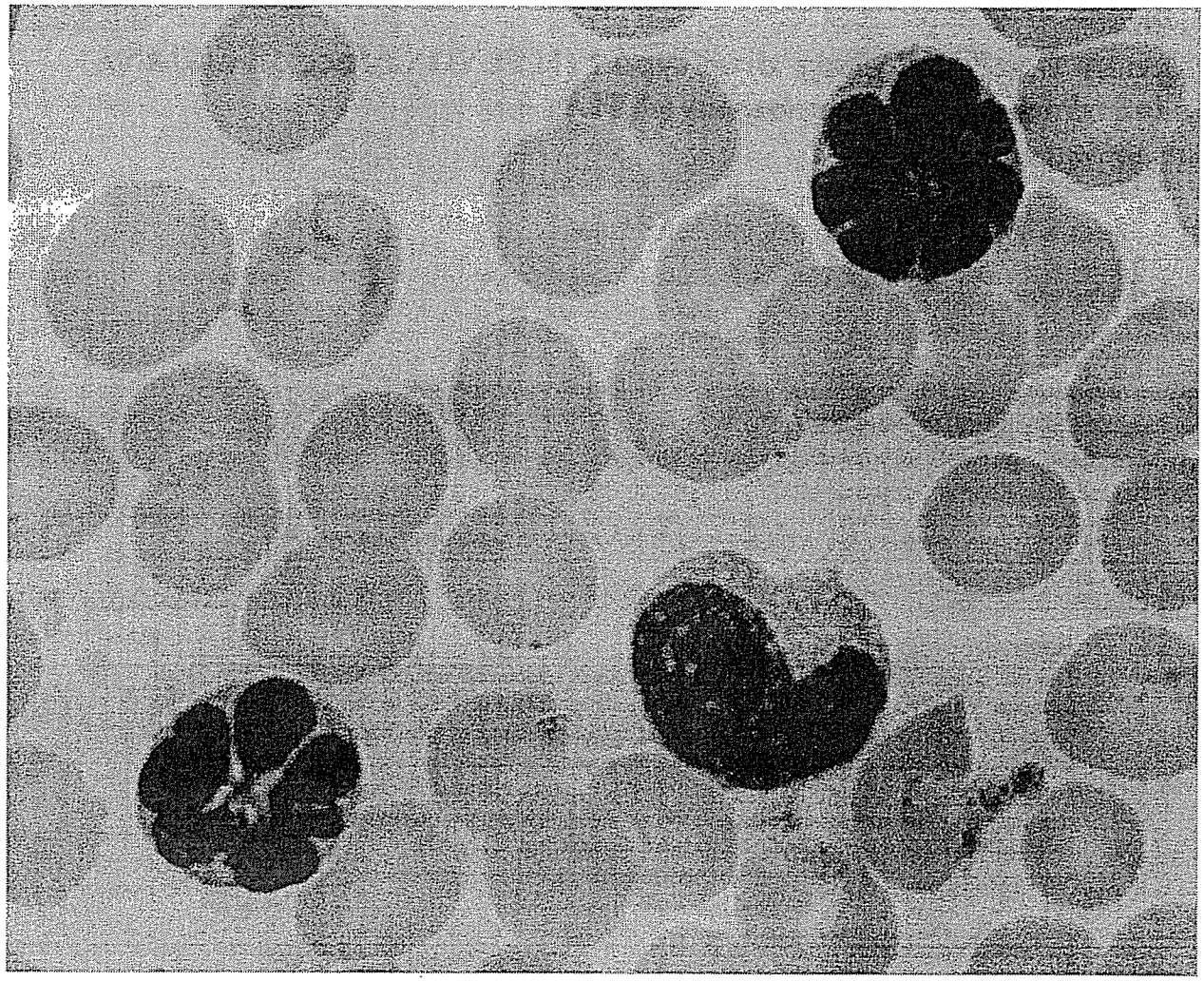

Figure 1
Typical "flower cell" in the peripheral blood of an acute ATL patient. In the peripheral blood of an acute ATL patient, leukemic cells with multilobulated nuclei.

transmission to humans is estimated to have occurred recently.

2. How does HTLV-I spread in humans?

There are approximately 10–20 million HTLV-I carriers in the world [11]. In particular, HTLV-I is endemic in Japan, parts of central Africa, the Caribbean basin and South America. In addition, epidemiological studies of HTLV-I have revealed high seroprevalence rates in Melanesia, Papua New Guinea and the Solomon islands, as well as among Australian aborigines [12]. In Japan, approximately 1.2 million individuals are estimated to be infected by HTLV-I, and more than 800 cases of ATL are diagnosed each year [13]. Moreover, this virus also causes the neurodegenerative disease, HTLV-I-associated myelopathy/tropical spastic paraparesis (HAM/TSP) [14,15]. The cumulative risks of ATL among HTLV-I carriers in Japan are estimated to be about 6.6% for men and 2.1% for women, indicating that most HTLV-I carriers remain asymptomatic throughout their lives [16].

3. How does HTLV-I replicate and increase its copy number?

The HTLV-I provirus has a similar structure to other retroviruses: a long terminal repeat (LTR) at both ends and internal sequences such as the *gag*, *pol* and *env* genes. A

characteristic of HTLV-I is the presence of the pX region, which exists between *env* and the 3'-LTR. This region encodes several accessory genes, which include the *tax*, *rex*, *p12*, *p21*, *p30*, *p13* and *HBZ* genes. Among these, the *tax* gene plays central roles in viral gene transcription, viral replication and the proliferation of HTLV-I-infected cells. Tax enhances viral gene transcription from the 5'-LTR via interaction with cyclic AMP responsive element binding protein (CREB). Tax also interacts with cellular factors and activates transcriptional pathways, such as NF-κB, AP-1 and SRF [8,17-20]. For example, activation of NF-κB induces the transcription of various cytokines and their receptor genes, as well as anti-apoptotic genes such as *bcl-xL* and *survivin* [21-23]. The activation of NF-κB has been demonstrated to be critical for tumorigenesis both *in vitro* and *in vivo* [24,25]. On the other hand, Tax variant without activation of NF-κB has also been reported to immortalize primary T-lymphocytes *in vitro* [26], suggesting that mechanisms of immortalization are complex. In addition to NF-κB, activation of other transcriptional pathways such as CREB by Tax should be implicated in the immortalization and leukemogenesis.

Tax also interferes with the functions of p53, p16 and MAD1 [27-30]. These interactions enable HTLV-I-infected cells to escape from apoptosis, and also induce genetic instability. Although inactivation of p53 function by Tax is reported to be mediated by p300/CBP [27,28,31] or NF-κB activation [32], Tax can still repress p53's activity in spite of loss of p300/CBP binding or in cells lacking NF-κB activation [33], indicating the mechanism of p53 inactivation by Tax needs further investigation.

Although Tax promotes the proliferation of infected cells, it is also the major target of cytotoxic T-lymphocytes (CTLs) *in vivo*. Moreover, excess expression of Tax protein is considered to be harmful to infected cells. Therefore, HTLV-I has redundant mechanisms to suppress Tax expression. Rex binds to Rex-responsive element (RxRE) in the U3 and R regions of the 3'-LTR, and enhances the transport of the unspliced *gag/pol* and the singly spliced *env* transcripts. By this mechanism, double-spliced *tax/rex* mRNA decreases, resulting in suppressed expression of Tax [34]. On the other hand, p30 binds to *tax/rex* transcripts, and retains them in the nucleus [35]. The *HBZ* gene is encoded by the complementary strand of HTLV-I, and contains a leucine zipper domain. HBZ directly interacts with c-Jun or JunB [36], or enhances their degradation [37], resulting in the suppression of Tax-mediated viral transcription from the LTR.

Transforming growth factor-β (TGF-β) is an inhibitory cytokine that plays important roles in development, the immune system and oncogenesis. Since TGF-β generally suppresses the growth of tumor cells, most tumor cells acquire escape mechanisms that inhibit TGF-β signaling, including mutations in its receptor and in the Smad molecules that transduce the signal from the receptor. Tax has also been reported to inhibit TGF-β signaling by binding to Smad2, 3 and 4 or CBP/p300 [38,39]. Inhibition of TGF-β signaling enables HTLV-I-infected cells to escape TGF-β-mediated growth inhibition.

ATL cells have been reported to show remarkable chromosomal abnormalities [40], which should be implicated in the disease progression. Tax has been reported to interact with the checkpoint protein MAD1, which forms a complex with MAD2 and controls the mitotic checkpoint. This functional hindrance of MAD1 by Tax protein causes chromosomal instability, suggesting the involvement of this mechanism in oncogenesis [30]. Recently, Tax has been reported to interact with Cdc20 and activate Cdc20-associated anaphase-promoting complex, an E3 ubiquitin ligase that controls the metaphase-to-anaphase transition, thereby resulting in mitotic abnormalities [41].

In contrast to HTLV-I, HTLV-II promotes the proliferation of CD8-positive T-lymphocytes *in vivo*. Although it was first discovered in a patient with variant hairy cell leukemia, HTLV-II is less likely to have oncogenic properties since there is no obvious association between HTLV-II infections and cancers. Regardless of the homology of their *tax* sequences, the oncogenic potential of Tax1 (HTLV-I Tax) is more prominent than that of Tax2 (HTLV-II Tax). The most striking difference is that Tax2 lacks the binding motif at C-terminal end to PDZ domain proteins, while Tax 1 retains it [42]. When the PDZ domain of Tax1 is added to Tax2, the latter acquires oncogenic properties in the rat fibroblast cell line Rat-1, indicating that this domain is responsible for the transforming activity of HTLV-I [43].

To understand the pleiotropic actions of Tax protein more clearly, transcriptome analyses are essential. The transcriptional changes induced by Tax expression have been studied using DNA microarrays, which revealed that Tax upregulated the expression of the mixed-lineage kinase MLK3. MLK3 is involved in NF-κB activation by Tax as well as NIK and MEKK1 [44]. In addition to transcriptional changes, Tax is also well known to interact with cellular proteins and impair or alter their functions. For example, proteomic analyses of Tax-associated complexes showed that Tax could interact with cellular proteins, including the active forms of small GTPases, such as Cdc42, RhoA and Rac1, which should be implicated in the migration, invasion and adhesion of T-cells, as well as in the activation of the JNK pathway [45].

4. How does HTLV-I transmit and replicate in vivo?

Receptor and transmission of HTLV-I

HTLV-I can infect various types of cells, such as T-lymphocytes, B-lymphocytes, monocytes and fibroblasts [46]. Glucose transporter 1 (GLUT-1) has been identified as a receptor for HTLV-I and this receptor is ubiquitously expressed on cell surfaces [47]. However, the HTLV-I provirus is mainly detected in CD4-positive lymphocytes, with about 10% in CD8-positive T-lymphocytes [48]. This situation possibly arises because Tax mainly induces the increase of CD4-positive T-lymphocytes *in vivo* by enhanced proliferation and suppressed apoptosis.

In HTLV-I-infected individuals, no virions are detected in the serum. In addition, the infectivity of free virions is very poor compared with that of infected cells. These findings suggest that HTLV-I is spread by cell-to-cell transmission, rather than by free virions. *In vitro* analyses of HTLV-I-infected cells revealed that HTLV-I-infected cells form "virological synapses" with uninfected cells. Contact between an infected cell and a target cell induces the accumulation of the viral proteins Gag and Env, viral RNA and microtubules, and the viral complex subsequently transfers into the target cell [49]. HTLV-I also spreads in a cell-to-cell manner via such virological synapses *in vivo*.

HTLV-I is mainly transmitted via three routes: 1) mother-to-infant transmission (mainly through breast feeding) [50]; 2) sexual transmission (mainly from male-to-female); and 3) parenteral transmission (blood transfusion or intravenous drug use) [12]. In either route, HTLV-I-infected cells are essential for transmission. This was supported by the findings that fresh frozen plasma from carriers did not cause transmission [51] and freeze-thawing of breast milk reduced vertical transmission [52].

Provirus load and transmission

The provirus load varies more than 1000-fold among asymptomatic carriers [53]. Since most infected cells are considered to have one copy of the provirus, the provirus load indicates the percentage of infected cells among lymphocytes. The provirus load is relatively constant during the latent period [53]. Analysis of naive individuals who seroconvert after marrying an HTLV-I-seropositive spouse demonstrated that the proviral gp46 sequences are identical among married couples. This finding confirmed that HTLV-I is transmitted from a seropositive individual to an uninfected spouse. The provirus loads frequently differ between couples despite infection by the same HTLV-I virus, indicating that the number of infected cells is determined by host factors rather than virus itself [54].

Why does HTLV-I increase the number of infected cells by the pleiotropic actions of Tax? The provirus load in peripheral blood mononuclear cells (PBMCs) is well correlated with that in breast milk, and a higher provirus load in breast milk increases the risk of vertical transmission of HTLV-I [55,56]. Similarly, a higher provirus load in PBMCs may be associated with a higher risk of sexual transmission. Thus, an increase in the number of infected cells by the actions of accessory genes, especially *tax*, facilitates transmission. Therefore, HTLV-I has strategies that increase the number of HTLV-I-infected cells via the action of accessory gene products, thereby increasing the chance of transmission.

Clonal expansion of HTLV-I-infected cells

After HTLV-I infection, viral proteins such as Tax protein promote the proliferation of infected cells and also inhibit apoptosis by their pleiotropic actions. Since the HTLV-I provirus is randomly integrated into the host genome, the identification of integration sites enables to identify each infected clone, and to trace the kinetics of infected cells *in vivo*. Analyses using inverse PCR, which can identify the integration sites of the HTLV-I provirus, revealed that the proliferation of infected cells is oligoclonal, and that infected cells persistently survive *in vivo* [57-59]. Importantly, such clonal expansion in carriers is directly associated with the onset of ATL [60]. Thus, the viral strategies to increase the number of HTLV-I-infected cells work efficiently in most carriers without any adverse effects. However, the increased number of infected cells causes an excess immune reaction, leading to inflammatory diseases, HAM/TSP, infective dermatitis [61] or HTLV-I-associated uveitis [62]. Moreover, such prolonged proliferation of infected CD4-positive T-lymphocytes results in the onset of ATL in some carriers after a long latent period.

Inactivation of Tax expression in ATL cells

As mentioned above, Tax expression confers advantages and disadvantages on HTLV-I-infected cells. Although the proliferation of infected cells is promoted by Tax expression, CTLs attack the Tax-expressing cells since Tax is their major target [63]. In HTLV-I-infected cells, Rex, p30 and HBZ suppress Tax expression. On the other hand, loss of Tax expression is frequently observed in leukemic cells. Three mechanisms have been identified for inactivation of Tax expression: 1) genetic changes of the *tax* gene (nonsense mutations, deletions or insertions) [64,65]; 2) DNA methylation of the 5'-LTR [65,66]; and 3) deletion of the 5'-LTR (Figure 2) [67]. Among fresh leukemic cells isolated from ATL patients, about 60% of cases do not express the *tax* gene transcript. Interestingly, ATL cells with genetic changes of the *tax* gene expressed its transcripts, suggesting that ATL cells do not silence the transcription when the *tax* gene is abortive [65]. Loss of Tax expression gives ATL cells advantage for their survival since they can escape from CTLs.

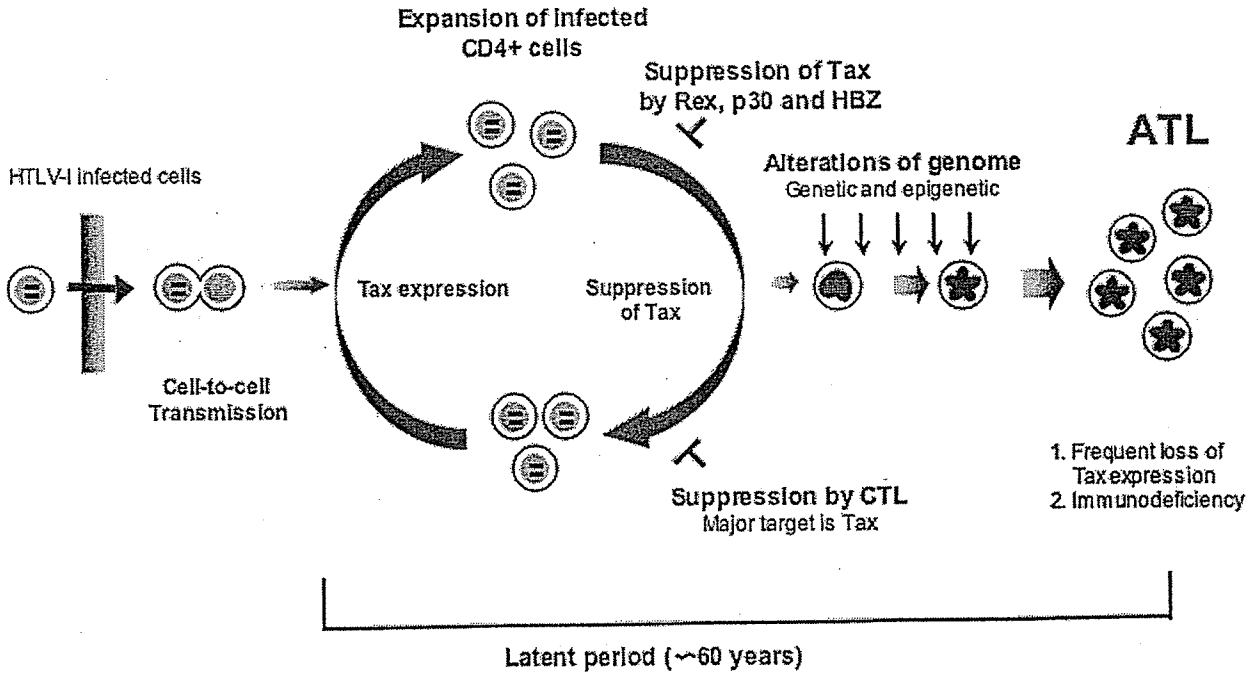

Figure 2
Natural course of HTLV-I infection to onset of ATL. HTLV-I is transmitted via three routes, and infected cells are necessary in all three. After infection, HTLV-I promotes clonal proliferation of infected cells by pleiotropic actions of Tax. Tax expression is suppressed by viral accessory gene products, such as Rex, p30 and HBZ proteins. Proliferation of HTLV-I infected cells is controlled by cytotoxic T-cells *in vivo*. After a long latent period, ATL develops in about 5% of asymptomatic carriers. The expression of Tax is inactivated by several mechanisms, suggesting that Tax is not necessary in this stage. Alternatively, alternations in the host genome accumulate during the latent period, finally leading to onset of ATL.

Longer lifespan of HTLV-I-infected cells and cancer

Lymphoid malignancy with a T-cell origin is rare compared with B-cell malignancy. ATL shares hematological, pathological and immunological features with cutaneous T-cell lymphoma (CTCL; Sezary syndrome and Mycosis fungoides). The frequency of CTCL in Japan is estimated to be one/million/year. On the other hand, the frequency of ATL among carriers is estimated to be 1000/million/year. From these data, HTLV-I infection is estimated to increase the risk of T-cell malignancy by up to 1000-fold in carriers.

HTLV-I infection confers a long lifespan on the infected cells due to the pleiotropic actions of Tax, resulting in increased numbers of infected cells. Such infected cells are essential for the transmission of HTLV-I. This strategy to increase the number of infected cells *in vivo* is thought to increase the incidence of cancer in T-cells. What is the mechanism for this oncogenesis? DNA methylation is known to be associated with aging. Some genes are hypermethylated in older people, indicating that DNA hypermethylation is a physiological phenomenon in some genes. Under normal circumstances, T-lymphocytes

survive for several years, and long-lived T-lymphocytes with disordered methylation should be replaced. However, HTLV-I-infected T-cells are considered to survive and accumulate abnormal methylation. The process of oncogenesis is similar to that of evolution [68]. The infected cells that are suitable for survival should be selected *in vivo*, and epigenetic and genetic changes of the genome play critical roles in this selection. Accumulating alterations of the host genome transform the HTLV-I-infected cells into ATL cells, and also enable ATL cells to proliferate in the absence of Tax expression (Figure 2). In the provirus, DNA methylation of the 5'-LTR silences viral transcription in leukemic cells, which facilitates the escape of ATL cells from the host immune system [65].

5. Somatic alterations in ATL cells

As described, some ATL cells can proliferate without functional Tax protein, suggesting that somatic (genetic and epigenetic) alterations cause transcriptional or functional changes to the host genes. The *p53* gene is frequently mutated in various cancers, and these mutations are associated with disease progression and a poor prognosis. The mutation rate of the *p53* gene in ATL cells has been reported to be 36% (4/11) and 30% (3/10) [69-71]. The *p16* gene is an inhibitor of cyclin-dependent kinase 4/6, and blocks the cell cycle. Genetic changes in this gene (deletion in most cases) have been described in many types of cancer cells. Deletion of the *p16* gene has also been reported in ATL cells [72]. Moreover, DNA methylation of the promoter region of the *p16* gene is also implicated in the suppression of p16 [73]. In addition, genetic changes in the $p27^{KIP1}$, *RB1/p105* and *RB2/p130* genes have been reported in ATL, although they are relatively rare: 2/42 (4.8%) for the $p27^{KIP1}$ gene; 2/40 (5%) for the *RB1/p105* gene; and 1/41 (2.4%) for the *RB2/p130* gene) [74]. The fact that higher frequencies of genetic changes in these tumor suppressor genes are observed among aggressive forms of ATL suggests that such genetic changes are implicated in disease progression.

Fas antigen was the first identified death receptor. It transduces the death signal by binding of its ligand, Fas ligand (FasL). ATL cells highly express Fas antigen on their cell surface [75], and are highly susceptible to death signals mediated by agonistic antibodies to Fas antigen, such as CH-11. Genetic changes of *Fas* gene in ATL cells, which confer resistance to the Fas-mediated signal, have been reported [76,77]. Normal activated T-lymphocytes express FasL as well as Fas antigen. Apoptosis induced by autocrine mechanisms is designated activation-induced cell death (AICD) and this controls the immune response [78]. Although ATL cells express Fas antigen, they do not produce FasL, thereby enabling ATL cells to escape from AICD. Attempts to isolate hypermethylated genes from ATL cells identified the *EGR3* gene as a hypermethylated gene compared to PBMCs from carriers [79]. EGR3 is a transcriptional factor with a zinc finger domain, that is essential for transcription of the *FasL* gene [80]. The finding that *EGR3* gene transcription is silenced in ATL cells could account for the loss of FasL expression, and the escape of ATL cells from AICD. Thus, alterations of the *Fas* (genetic) and *EGR3* (epigenetic) genes are examples of ATL cell evolution *in vivo*.

Disordered DNA methylation has been identified in the genome of ATL cells compared with that of PBMCs from carriers: hypomethylation is associated with aberrant expression of the *MEL1S* gene [81], while hypermethylation silences transcription of the *p16* [73], *EGR3* and *KLF4* genes as well as many others [79]. It is reasonable to consider that other currently unidentified genes are involved in such alterations of the genome in ATL cells, and play roles in leukemogenesis.

Transcriptome analyses using DNA microarrays have revealed transcriptional changes that are specific to ATL cells. Among 192 up-regulated genes, the expressions of the *tumor suppressor in lung cancer 1 (TSLC1)*, *caveolin 1* and *prostaglandin D2 synthase* genes were increased more than 30-fold in fresh ATL cells compared with normal CD4+ and CD4+, CD45RO+ T-cells [82]. TSLC1 is a cell adhesion molecule that acts as a tumor suppressor in lung cancer. Although TSLC1 is not expressed on normal T-lymphocytes, all acute ATL cells show ectopic TSLC1 expression. Enforced expression of TSLC1 enhances both the self-aggregation and adhesion abilities to vascular endothelial cells in ATL cells. Thus, TSLC1 expression is implicated in the adhesion or infiltration of ATL cells. By screening a retrovirus cDNA library from ATL cells, a gene with oncogenic potency was identified in NIH3T3 cells, and designated the *Tgat* gene [83]. Ectopic expression of the *Tgat* gene is observed in aggressive forms of ATL, and *in vitro* experiments showed that its expression is associated with an invasive phenotype.

6. Immune control of HTLV-I infection

The host immune system, especially the cellular response, against HTLV-I exerts critical control over virus replication and the proliferation of infected cells [84]. CTLs against the virus have been extensively studied, and Tax protein was found to be the dominant antigen recognized by CTLs *in vivo* [63]. HTLV-I-specific CD8-positive CTLs are abundant and chronically activated. The paradox is that the frequency of Tax-specific CTLs is much higher in HAM/TSP patients than in carriers. Since the provirus load is higher in HAM/TSP patients, this finding suggests that the CTLs in HAM/TSP cannot control the number of infected cells. One explanation for this is that the CTLs in HAM/TSP patients show less efficient cytolytic activity toward infected cells, whereas CTLs in carriers can suppress the

proliferation of infected cells [85]. Hence, the gene expression profiles of circulating CD4+ and CD8+ lymphocytes were compared between carriers with high and low provirus loads. The results revealed that CD8+ lymphocytes from individuals with a low HTLV-1 provirus load show higher expressions of genes associated with cytolytic activities or antigen recognition than those from carriers with a high provirus load [86]. Thus, CD8+ T-lymphocytes in individuals with a low provirus load successfully control the number of HTLV-I-infected cells due to their higher CTL activities. Thus, the major determinant of the provirus load is thought to be the CTL response to HTLV-I.

As mentioned above, the provirus load is considered to be controlled by host factors. Considering that the cellular immune responses are critically implicated in the control of HTLV-I infection, human leukocyte antigen (HLA) should be a candidate for such a host genetic factor. From analyses of HAM/TSP patients and asymptomatic carriers, HLA-A02, and Cw08 are independently associated with a lower provirus load and a lower risk of HAM/TSP. In addition, polymorphisms of other genes (*TNF-α*, *SDF-1*, *HLA-B54*, *HLA-DRB-10101* and *IL-15*) are also associated with the provirus load, although their associations are not as significant compared with HLA-A02, and Cw08 [87,88]. Regarding the onset of ATL, only a polymorphism of *TNF-α* gene was reported to show an association [89]. However, familial clustering of ATL cases is a well-known phenomenon, strongly suggesting that genetic factors are implicated in the onset of ATL [90-92].

Spontaneous remission is more frequently observed in patients with ATL than those with other hematological malignancies [90,93]. Usually, this phenomenon is associated with infectious diseases, suggesting that immune activation of the host enhances the immune response against ATL cells. If the immune response against HTLV-I is implicated in spontaneous remission, this suggests the possibility of immunotherapy for ATL patients by the induction of an immune response to HTLV-I [94], for example via antigen-stimulated dendritic cells.

Immunodeficiency in ATL patients is pronounced, and results in frequent opportunistic infections by various pathogens, including *Pneumocystis carinii*, cytomegalovirus, fungus, *Strongyloides* and bacteria, due to the inevitable impairment of the T-cell functions [95]. To a lesser extent, impaired cell-mediated immunity has also been demonstrated in HTLV-I carriers [96]. Such immunodeficiency in the carrier state may be associated with the leukemogenesis of ATL by allowing the proliferation of HTLV-I-infected cells. A prospective study of HTLV-I-infected individuals found that carriers who later develop ATL have a higher anti-HTLV-I antibody and a low anti-Tax antibody level for up to 10 years preceding their diagnosis. This finding indicates that HTLV-I carriers with a higher anti-HTLV-I titer, which is roughly correlated with the HTLV-I provirus load, and a lower anti-Tax reactivity may be at the greatest risk of developing ATL [97]. The anti-HTLV-I antibody and soluble IL-2 receptor (sIL-2R) levels are correlated with the HTLV-I provirus load [53], and a high antibody titer and high sIL-2R level are risk factors for developing ATL among carriers [98]. Taken together, these findings suggest that a higher proliferation of HTLV-I-infected cells and a low immune response against Tax may be associated with the onset of ATL. Given these findings, potentiation of CTLs against Tax via a vaccine strategy may be useful for preventing the onset of ATL [99].

EBV-associated lymphomas frequently develop in individuals with an immunodeficient state associated with transplantation or AIDS. This has also been reported in an ATL patient [100]. Does such an immunodeficient state influence the onset of ATL? Among 24 patients with post-transplantation lymphoproliferative disorders (PT-LPDs) after renal transplantation in Japan, 5 cases of ATL have been reported. Considering that most PT-LPDs are of B-cell origin in Western countries, this frequency of ATL in Japan is quite high. Although the high HTLV-I seroprevalence is due to blood transfusion during hemodialysis, the immunodeficient state during renal transplantation apparently promotes the onset of ATL [101]. In addition, when experimental allogeneic transplantation was performed to 12 rhesus monkeys and immunosuppressive agents (cyclosporine, prednisolone or lymphocyte-specific monoclonal antibodies) were administered to prevent rejection, 4 of the 7 monkeys that died during the experiment showed PT-LPDs. Importantly, the STLV provirus was detected in all PT-LPD samples [102]. These observations emphasize that transplantation into HTLV-I-infected individuals or from HTLV-I positive donors require special attention.

Although the mechanism of immunodeficiency remains unknown, some previous reports have provided important clues. One mechanism for immunodeficiency is that HTLV-I infects CD8-positive T-lymphocytes, which may impair their functions [48]. Indeed, the immune response against Tax via HTLV-I-infected CD8-positive T-cells renders these cells susceptible to fratricide mediated by autologous HTLV-I-specific CD8-positive T-lymphocytes [103]. Fratricide among virus-specific CTLs could impair the immune control of HTLV-I. Another mechanism for immunodeficiency is based on the observation that the number of naive T-cells decreases in individuals infected with HTLV-I via decreased thymopoiesis [48]. In addition, CD4+ and CD25+ T-lymphocytes are classified as immunoregulatory T-cells that control the host immune system. Regulatory T-cells suppress the immune reaction via the