A DOCTORAL DISSERTATION

Transcription and Chromosomal Location of the Active Retroelements HERV-K and L1Hs in Human Cells

Department of Biology

Graduate School
Cheju National University

Sang Hyun Han

December, 2003

인간 세포에서 Retroelements HERV-K와 L1Hs의 전사와 염색체상의 위치

지도교수 오 문 유

한 상 현

이 논문을 이학 박사학위 논문으로 제출함

2003년 12월

한상현의 이학 박사학위 논문을 인준함

심사위원장 _____

위 원 _____

위 원 _____

위 원 _____

위 원 _____

제주대학교 대학원

2003년 12월

Transcription and Chromosomal Location of the Active Retroelements HERV-K and L1Hs in Human Cells

Sang Hyun Han

(Supervised by Professor **Moon You Oh)**

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Date Approved:

_____

_____

_____

_____

_____

_____

Department of Biology

Graduate School, Cheju National University

December, 2003

ABSTRACT

Transposable elements play an important role in genome function. They can transpose their chromosome locations using transposition and recombination. The resulting repetitive sequences provide numerous opportunities for genetic recombination. Moreover, the changes in the chromosomal locations of transposable elements induce genomic instability, transformation, and differentiation of cells. The human genome contains several retro-transposable elements. The representative retroelements, human endogenous retroviruses (HERVs) and long interspersed nuclear elements (LINEs), are functionally active. HERVs are estimated to constitute at least 0.5-3% of the human genome; they encode characteristic retroviral proteins, such as gag, pol, and env, and have flanking promoter-like long terminal repeats (LTRs). HERV LTRs and solitary LTRs found in the human genome play an important role in cell biology. LTRs have various elements for regulating the expression of HERVs and neighboring host genes. LINE-1 elements of human (L1Hs) have two open reading frames: one encodes an enzyme for reverse transcription and the other one for RNA-protein complex formation. This study investigated the transcriptions and chromosomal locations of the transcripts of HERV-K and L1Hs in various human cells (two normal lymphocytes, two normal fibroblast, and sixteen cancer cell lines). In addition, this study examined the effects of selected chemical carcinogens on gene expression. RNA transcripts of the retroelements were detected in all cell lines tested, but the relative amounts differed according to cell type. For example, the levels of transcription varied at different developmental stages in fibroblasts as well as in leukemia cells. In addition, the expression patterns of HERV-K genes (HERV-K10 LTR, HERV-K(T47D) LTR, HERV-K(C4) *env*, and HERV-K10 *env*), L1Hs *p40*, and regulatory protein genes (*BAX*, *c-MYC*, *IL-6*, and *p53*) were

changed after treated with carcinogens. Since cell type-specific activation of HERV-K genes has been reported in various cells, it is thought that HERV-K genes may have functions in proliferation, apoptosis, differentiation, cancer development, and anti-viral protection strategies against exogenous retroviruses. L1Hs transcription was not distinguishable in the cell lines tested as a cell-type independent, but varied in response to carcinogenic and stress factors. Although RNA transcripts of retroelements were detected in the cell lines tested, evidences for a direct association between retroelement expression and human cancers could not find. The results suggest that the altered activity levels of the retroelements contribute to cell transformation, based on changes in the expression of the HERV-K genes and L1Hs *p40* in responses to carcinogens. Some carcinogens induced the transposition activity. Genome-scale amplification has shown the altered amplification patterns using transposon-specific ITR and arbitrary random ITR as a amplification primer. After average 5 passages, the distinguishable fragments have seen in an amplification profile. Sequence analysis following a BLAST search and phylogenetic analysis showed that RNA transcripts of the retroelements were transcribed from various locations on most chromosomes. RNA transcripts of HERV-K10 LTR elements were located from the centromere to the telomere in all chromosomes, except chromosomes 21 and Y. HERV-K10 LTRs might be introduced into the human genome several times before and after mammalian speciation. L1Hs element transcripts were found on all but chromosome 21. Active L1Hs elements in at least 140 positions of the haploid human genome were identified by similarity search and putative protein translation of DNA sequences. In addition, the putative active L1Hs elements formed two clusters in the strict consensus tree. Based on chromosomal locations, retroelement transcripts were transcribed from many fragile sites, suggesting their association with chromosome instability, cancer-related breakpoints, and genetic disorder-related loci. After retroelement transcripts were introduced into human genome, they may expand

introns of the genomes and produce many pseudogenes. Consequently, this study suggests that many retroelements are transcribed and transposed in the host genome, and they can respond to various stimuli in the extra- and intracellular environments. In addition, HERV-K and L1Hs retroelements are widely dispersed throughout the human genome, maintaining their retrotransposition potential.

Key words: cancer, carcinogen, chromosomal instability, chromosome, chromosome rearrangement, env, fragile site, gene expression, genetic disorder, genome, genomic instability, HERV-K, HERV-K10 LTR, LTR, LINE-1, L1Hs p40, mutation, recombination, retroelement, retroposon, retrotransposition, retrovirus, transcription, transposable element, transposition, transposon

CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

The human genome contains a number of repetitive sequences that originated from retroviruses. Recent studies have reported that about one-third of the human genome sequence is derived from different transposable elements (Miki, 1998; Smit, 1999). Of those, endogenous retrotransposable elements have been distributed by retrotransposition during the course of evolution of the genome. The process of retrotransposition is mediated by RNA transcripts, as in human immunodeficiency virus (HIV). This large family of transposable elements can be classified into several groups according to differences in their lengths, structures, and protein encoding genes (Moran *et al*., 1996; Hohjoh and Singer, 1997; Malik *et al*., 1999). They contain short- and long interspersed nuclear elements (SINEs and LINEs), such as *Alu* repeats and L1 retroposon, and human endogenous retroviruses (HERVs) (Fig. 1). Although some retroelements are still active and produce deleterious mutations that cause hereditary diseases, most transposon activity in humans seems to be ceased (reviewed in Smit, 1999; Kazazian, 2000).

1. Human Endogenous Retrovirus-K

In 1973, it was discovered that primate placentas produce retrovirus-like particles. These particles not only resemble retroviruses morphologically, but also possess an enzyme activity that is indistinguishable from that of retroviral RNA-dependent DNA polymerase (reverse transcriptase) (Kalter *et al*. 1973; Strickland *et al*., 1973; Medstrand and Blomberg, 1993).

Vertebrate genomes contain many endogenous retroviral sequences that resulted, in all

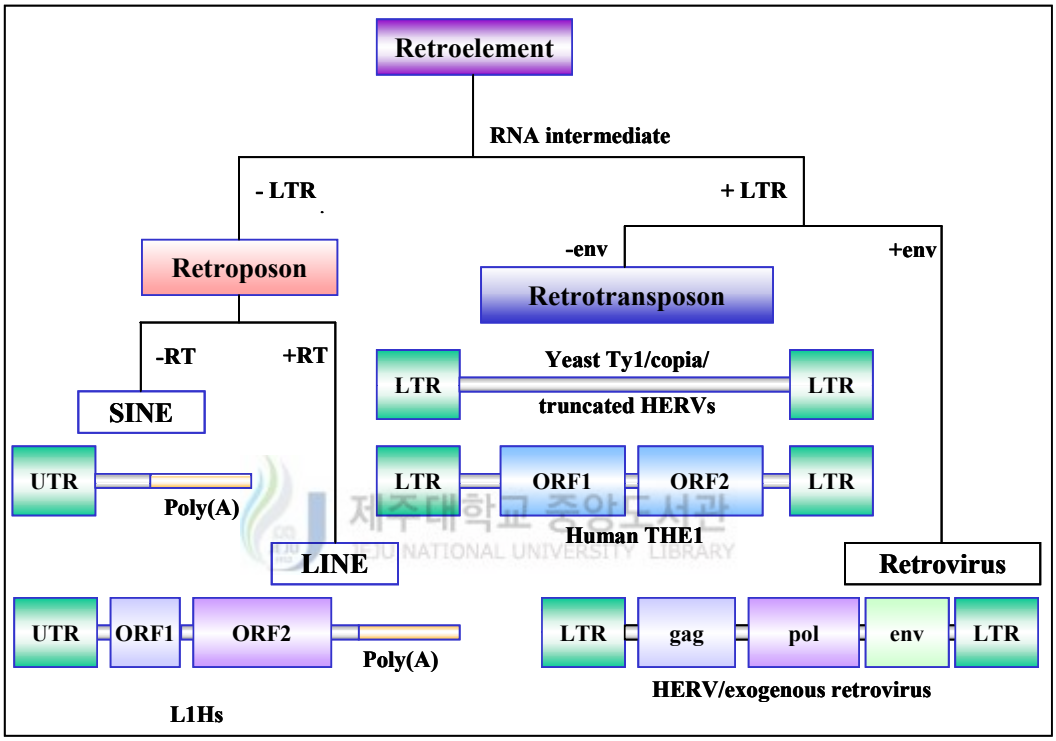


Figure 1. Classification of retroelements in human genome.

likelihood, from the integration of exogenous retroviruses into the germ line of a predecessor species. Such the events of inter-genomic integration are thought to be occurred frequently during evolution, and various families of endogenous retroviruses (ERVs) can be found in vertebrate genomes (Wilkinson *et al*., 1994; Benit *et al*., 1999). The expression of stably integrated proviral sequences and subsequent retrotranspositions lead to an increase in the copy number of proviral sequences in the host. Probably, HERV-K-like elements may enter into the primate genome about 33-40 million years ago. There are estimated to be ca. 25,000 copies of HERV-K long terminal repeats (LTRs) in the human genome (Leib-Mösch *et al*., 1993). These integrated proviral sequences have subsequently inherited to a stable Mendelian fashion in subsequent generations and transmitted to newly arising species, and then propagated due to retrotransposition events.

The retroviral genome consists of two, usually identical, ssRNA molecules ranging about 7-10 kb, with 5'-CAP and 3'-poly(A) tail. The order of the genes encoding structural proteins is invariably *gag*, *pol*, and *env*. *Gag* stands for group-specific antigen in the virion core, encoding four separate proteins; matrix protein (Pepinsky and Vogt, 1984), capsid protein (Dickson *et al*., 1982), nucleocapsid protein, and protease (Coffin *et al*., 1997). *Pol* encodes three enzymes: reverse transcriptase for synthesis of viral complementary DNA (Varmus, 1987, 1988; Coffin *et al*., 1997), RNase H, and integrase. *Env* encodes two envelope proteins: surface and transmembrane proteins (Dickson *et al*., 1982). The viral genome is bounded by LTRs. There is no apparent DNA sequence specificity for host integration sites although some preference has been noted (Wilson and Young, 1975).

Two major characteristics of LTRs make their important features in the genome: they contain regulatory elements and are probably inserted into transcriptionally active regions. Some HERV LTRs recruited during evolution as regulatory elements for gene

control, as they provide sequences for the initiation of transcription, for instance, promoters, enhancers, hormone-responsive elements, transcription-factor binding sites, and polyadenylation signals. Consequently, they can induce significant changes in the expression of neighboring cellular genes (Ting *et al*., 1992; Feuchter-Murthy *et al*., 1993; Knössl *et al*., 1999; Domansky *et al*., 2000; Schön *et al*., 2001). Reports suggest that LTRs frequently modulate the expression of nearby genes directly through transcription regulatory signals or regulation of translation (Kapitonov and Jurka, 1999; Kowalski *et al*., 1999; Mager *et al*., 1999; Sverdlov, 1998). Note that in the reported examples of LTR involvement in gene regulation, these elements are always incorporated in the 5'-proximal regions of genes. The first clones of HERV sequences were obtained by screening a genomic library with an African green monkey ERV probe (Martin *et al*., 1981). Since then, a number of HERV elements have been reported, and they are estimated to constitute at least 0.6-3% of the human genome (Wilkinson *et al*., 1994; Mayer *et al*., 1998; Zsiros *et al*., 1998).

The majority of these elements appear to be either rearranged or mutated such that they are in capable of producing proteins. Almost all HERV sequences are too unstable to encode functional proteins. It has been suggested that the HERV-K family is an important exception to this generality because it contains open reading frames (ORFs) and may be involved in tumor development. Full-length ORFs for *gag*, *prt*, *pol*, and *env* have been isolated from the human genome (Wilkinson *et al*., 1994; Löwer *et al*., 1996; Kitamura *et al*., 1996; Mayer *et al*., 1997).

The HERV-K family of human endogenous retroviral sequences was originally cloned from Syrian hamster intracisternal A type particles (Ono *et al*., 1986). RNA-transcripts of HERV-K have been detected in distinct cell lines and tissues (Brodsky *et al*., 1993; Medstrand and Blomberg, 1993; Vinogradova *et al*., 2001), and HERV-K particles have been identified in the placenta and supernatants of teratocarcinoma cell lines (Boller *et*

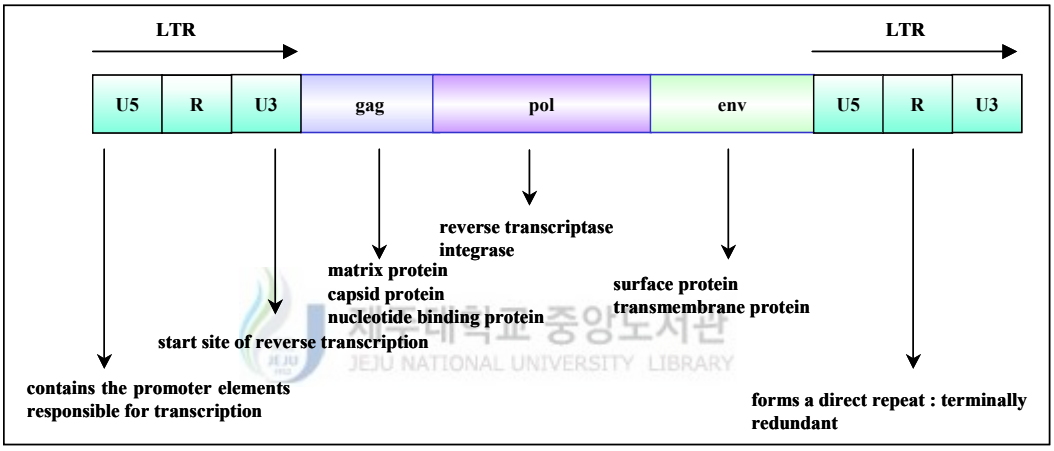


Figure 2. A representative structure of HERV-K and the functions of each gene.

al., 1993; Löwer *et al*., 1993). Transcripts containing HERV-K *pol* and *env* ORFs have been detected in human teratocarcinoma cell lines, and enzymatically active *gag* and *pol* gene products have been expressed *in vitro* (Löwer *et al*., 1993; Kitamura *et al*., 1996; Tönjes *et al*., 1997; Berkhout *et al*., 1999). Recent reports have suggested the possible involvements of HERV-K sequences in the initiation or progression of germ cell tumors, and Gag protein and *gag*-encoded retrovirus-like particles are present in tumor cells and derived cell lines, respectively (Sauter *et al*., 1995; Sauter *et al*., 1996). Possible roles of HERV-K elements in human diseases have been considered in insulin dependent diabetes mellitus (Conard *et al*., 1997), seminoma (Sauter *et al*., 1995), and HERV-K-T47D mammary carcinoma (Seifarth *et al*., 1998). According to reports of HERV-K elements in several cancer cells, HERV-K LTRs and HERV-W *env* sequences have been found in RNA transcripts (Zsiros *et al*., 1998; Kim *et al*., 2001; Yi *et al*., 2001). This suggests that HERV-K elements are active in cancer cells and may be involved in carcinogenesis. In addition, HERV-K Gag and Env antibodies have been detected in germ cell tumors and trophoblastic tumors (Müeller-Lantzsch *et al*., 1993; reviewed in Mayer *et al*., 1997; Herbst *et al*., 1999).

2. Long Interspersed Nuclear Elements-1

The LINE-1 element is a retroposon that is found in all mammalian genomes, and belongs to the class of retroposons lacking LTRs. LINE-1 elements account for about 15% of the human genome and have achieved this abundance by retrotransposition (Fanning and Singer, 1987; Hutchison *et al*., 1989; Singer *et al*., 1993). Active human L1 sequences (L1Hs) can cause mutations, diseases, genetic variations, and

polymorphisms, and even their remainders can be involved in genetic recombination and chromosomal rearrangements (Gilbert *et al*., 2002; Kazazian and Goodier, 2002; Sankoff *et al*., 2002; Chi *et al*., 2003). At least 3-4% of L1Hs are full-length, while the others are truncated to varying degrees (Hwu *et al*., 1986; Smit, 1999).

Full-length L1Hs are about 6 kbp in length and contain two nonoverlapping open reading frames (ORFs), a 5' untranslated region (UTR) and 3' poly(A) tail (Fig. 3). ORF1 encodes an RNA-binding protein p40, which forms an RNA-protein (RNP) complex, and has been observed in breast cancer, teratocarcinoma, normal breast tissue, germ cell tumors lines, and in several kinds of tumor cells (Bratthauer and Fanning, 1992; 1993; Bratthauer *et al*., 1994; Asch *et al*., 1996). L1Hs p40 has no significant sequence homology to *gag* and *gag*-like proteins observed in endogenous retroviruses (Hohjoh and Singer, 1997). ORF2 encodes a 149-kDa protein, which has endonuclease and reverse transcriptase activities. Examination of the contribution of various protein regions to the RT activity has demonstrated that deletions and mutations in the ORF2 protein domains eliminate or significantly decrease the RT activity (Dombroski *et al*., 1994; Feng *et al*., 1996). Furthermore, mutation and deletion analyses of the L1Hs element have shown that both the p40 and the EN and RT regions of the ORF2 protein are required for efficient retrotransposition in cultured cells (Feng *et al*., 1996; Dhellin *et al*., 1997; Clements and Singer, 1998).

L1Hs elements are found in all chromosomes including sex chromosomes X and Y. Many L1Hs copies have accumulated in AT-rich regions of human genome (Smit, 1999; Pavlicek *et al*., 2002). Active L1Hs elements have the capacity to cause mutations, diseases, genetic variations, and polymorphisms, which their inactive copies appear to be involved in chromosome recombination and rearrangement during cell division (Brouha *et al*., 2002). There are about 100,000 L1Hs elements in the human genome, but only about 30-60 of the full-length L1Hs are capable of autonomous retrotransposition, and

capable of synthesizing new copies at different positions on chromosomes containing the L1Hs integrating target sequence. The transcript is reverse transcribed and integrated into a new genomic location by a process termed 'target-primed reverse transcription' (Luan *et al*., 1993). On integration, the L1Hs element typically is flanked by target-site duplications of 6-20 bp; the integration site has a weak consensus sequence (5'-Py-Py/AAAA-3') (Cost and Boeke 1998). In addition, L1Hs elements insert at random locations in the genome (Hutchison *et al*., 1989; Ovchinnikov *et al*. 2001). The haploid human genome has approximately above 7×10^5 putative target sites for L1Hs target sequences.

Recently, new retrotranspositions of active human elements have been identified through systematic screening of disease genes (Kazazian *et al*., 1988; Morse *et al*., 1988; Miki *et al*., 1992; Narita *et al*., 1993; Miki, 1998). These developments have identified new insertions or deletions. For example, insertion in *β-globin* gene in mother and her daughter showing typical β-thalassemia trait (Kimberland *et al*., 1999), factor Ⅷ and Ⅸ genes in the patients with hemophilia A, B (Kazazian *et al*., 1988; Youssoufian *et al*., 1988; Woods-Samuels *et al*., 1989), in the *APC* gene in a colon cancer patient (Miki *et al*., 1992), in the retinitis pigmentosa-2 (*RP2*) gene of a X-linked retinitis pigmentosa and his carrier mother (Kimberland *et al*., 1999), and in the *dystrophin* gene of a muscular dystrophy (MD) patient (Holmes *et al*., 1994); deletions in the *PAX* gene in aniridia patients (Drechsler and Royer-Pokora, 1996) and in the *dystrophin* gene in severe muscular dystrophy patients (Sugimoto *et al*., 2001) have also been identified. The major sequences of L1Hs are integrated in human retinoblastoma (*Rb*), interleukine 2 (*IL-2*), factor VIII genes. Since *Rb* and factor VIII have a high frequency of chromosomal deletions in various cancers and haemophilia A, the universal integration of long, and homologous L1Hs segments in these genes and all chromosomes may promote abnormal DNA rearrangement (Kuo *et al*., 1998). L1Hs insertions into human

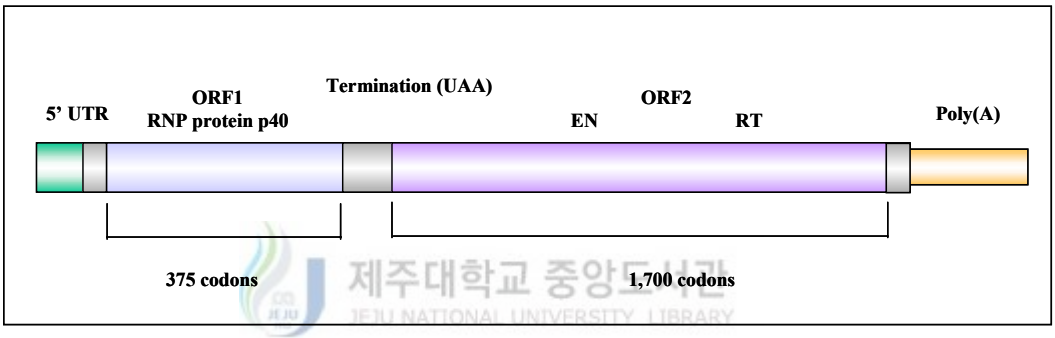


Figure 3. Schematic illustration of LINE1 element.

genes have also been described in cultured mammalian cells transfected with full-length elements, and large deletions due to L1Hs can occur via various pathways *in vitro* and *in vivo* (Moran *et al*., 1996; Kimberland *et al*., 1999; Gilbert *et al*., 2002).

3. Purposes of This Study

While it has been suggested that retroelement activity is involved in genome instability, as repetitive elements are abundant and widespread in the human genome, it is very difficult to elucidate a direct relationship between their activities and cancer development and genome instability. To understand the association between the activity of the autonomous retroelements (HERV-K and L1Hs) and human cancers, this study was performed.

At first, to investigate the difference of gene expression, transcription patterns of HERV-K and L1Hs retroelements were examined in normal fibroblast cells and cancer cells. In addition, to screen the effects on genomic instability, gene expression of these retroelements, and genome scale transposition activities were analyzed in normal fibroblast cells by carcinogen exposure. To determine the active retroelement sequences transcribing and transposing and its chromosomal locations in normal and cancer cells, DNA sequencing of transcripts from various human cells and following analyses on the chromosomal locations were performed. The associations between the activation of these retroelements and genetic disorders including cancers and diseases were investigated. Finally, evolutionary implications were tested by phylogenetic analysis.

Ⅱ. MATERIALS AND METHODS

1. Cell Lines and Cell Culture

Whole blood samples were collected from peripheral vessel in two healthy individuals. Two normal fibroblasts (Table 1) and sixteen cancer cell lines (Table 2) were used in this study. Cell lines were obtained from Korean Cell Line Bank (KCLB) and American Tissue and Cell Culture Co. (ATCC). Each cell was cultured in Dulbecco's modified Eagle's medium (Gibco-BRL, USA) or RPMI1640 medium, supplemented with 10% fetal bovine serum (FBS) and 1× antimycotic-antibiotics, under constant 37℃ and 5% CO_2 condition, as previously described.

2. Chemicals and Treatment

The normal fibroblast cell line CCD-25Lu was treated with benzidine (BZ), dimethylaminoazobenzene (DMBA), dimethyl sulfoxide (Me2SO), ethanol (EtOH), ethidium bromide (EtBr), *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine (MNNG), and 12-*O*-tetradecanoyl phorbol-13-acetate (TPA). The chemicals were dissolved in the appropriate solvent and diluted at least 100 fold with phosphate-buffered saline (PBS). The dosage of each chemical that corresponded to the maximum resistance concentration was determined by a DNA fragmentation assay and a microscopic observation.

Table 1. Normal human cells used in this study.

Cell line	Symbol	Age	Gender	Status	Origin
KYJ*	YJ	22	female	normal, lymphocyte	blood
KBG*	BG	30	male	normal, lymphocyte	blood
CCD-25Lu	CC	7	male	normal, fibroblast	lung
HEL-299	HE	embryonic	male	normal, fibroblast	lung

* whole blood from volunteer without clinical symptoms.

Table 2. Human cancer cell lines used in this study.

Cell line	Symbol	Status	Origin
CEM-CM3	CE	acute lymphoblastic leukemia	T lymphoblast
HL-60	HL	leukemia, acute promyelocytic	blood
Jurkat	JU	acute T cell leukemia	T lymphocyte
Molt-4	MO	acute lymphoblastic leukemia	T lymphoblast
U-937	U9	histocytic lymphoma	monocyte
A-431	A3	carcinoma	skin
A-498	A4	carcinoma	kidney
A-549	A5	carcinoma	lung
HCT-15	HC	adenocarcinoma	colon
Hep-G2	HG	hepatoblastoma	liver
MCF-7	MC	adenocarcinoma; pleural effusion	breast, mammary gland
MIA PaCa-2	MP	carcinoma	pancreas
SNU-16	SU	adenocarcinoma	stomach; ascites
SNU-C5	SC	adenocarcinoma	colon; colorectal
SK-N-SH	SN	neuroblastoma, metastasis to bone marrow	brain
SK-OV-3	SO	adenocarcinoma, malignant ascites	ovary

3. Total DNA Extraction

Total DNA was extracted from cultured cells and whole blood with slightly modified sucrose-proteinase K method. Before application, the whole blood cells were treated red cell lysis buffer (RCLB) to remove the erythrocytes. Leukocytes from whole blood and cultured cells were washed with PBS. All samples were lysed with 5 ㎖ of pre-warmed sucrose-proteinase K lysis buffer (27% sucrose, 1× SSC, 1% SDS, 1 mM EDTA, and 100 ㎍/ml proteinase K), separated twice with phenol/chloroform/iso-amylalcohol, centrifuged at 3,000 rpm for 10 min, precipitated in iso-propanol, and washed with ethanol, according to the procedure described by Birren *et al*. (1997). The DNA concentration ranged from 50 ng/㎕ depending on the sample, each DNA was used as a template for PCR amplification, DNA fragmentation assay, and transposition analysis.

4. Cytogenetic Analysis and Fluorescence *in situ* Hybridization (FISH)

FISH analysis was performed with metaphase chromosomes of peripheral leukocytes, KYJ. The leukocytes were isolated with a gradient centrifugation method using HistoPaque 1770 (Sigma, USA). The cells were cultured in McCoy' 5A medium containing 15% FBS, streptomycin-penicillin, and a mitogen phytohemagglutinin (Sigma, USA). On 3 day, a mitotic blocker, colcemid was treated to prepare the metaphase chromosomes. Harvesting and preparation of slides were performed according to standard cytogenetic procedures. Harvested cells were treated with hypotonic solution (0.075 M KCl) for 20 min at 37℃, and fixed at -20℃ in Carnoy's

fixative. For FISH and G-Banding analyses, cells were dropped on microscope slides and air dried. Slide preparations were allowed to age at room temperature for 1 day or 2 days before the analyses. Slides were stained with Giemsa. Image acquisition and analysis were performed by the cytogenetic analysis software Karyo-G Band ver. 5.0 (Imstar, France). A L1Hs p40 probe was used for FISH analysis of KYJ metaphase. Probes were labelled with purified PCR product using Dig-Labelling and Detection Kit (Roche, Germany). Each labelled probe was precipitated with Salmon sperm DNA and resuspended in hybridization mixture (50% formamide, 10% dextran sulfate, 2× SSC, pH 7.0). The metaphase preparations were pre-treated with RNase A and pepsin. Slides were denatured in 70% formamide, 2× SSC for 5 min at 70℃. After probe denaturation at 75℃ for 5 min, pre-annealing was performed at 45℃ for 30 min, The pre-annealed hybridization mixture was added to the denatured slides and allowed to hybridize for 16 hr. Post-hybridization washing were performed for 3× 5 min with 50% formamide/ 2× SSC at 45℃, and for 3× 5 min with 50% formamide/0.1× SSC at 60℃, both with agitation. After blocking with 5% bovine serum albumin in 2× SSC/0.1% Tween-20, the Digoxigenin-labelled probes were detected using anti-digoxigenin Fab fragment conjugated with Rhodamine (Roche, Germany). The air dried slides were counter-stained with DAPI (Sigma, USA) and mounted with antifade (2% DABCO in glycerol). Image acquisition and analysis were performed with the software mFISH ver. 5.0 (Imstar, France).

5. Transposition Analysis

Transposition activity had been screened on each treatment of cell cultures.

Table 3. List of primers used in transposition analysis.

Name	Gene	Sequence (5'→3')	Reference
HsMar	*Hsmar1*	TATTAGGTTGGTGCAAAAGT	Dematteі *et al*. (2000)
Mos1	*Mos1*	AGCTCGTTTA	Augé-Gouillou *et al*. (2001)
Pogo1	*Pogo1*	CAGGCATACCTC	Hugh (1996)
HSH103	AP*	GTGACGCCGC	*This study*
HSH147	AP*	GTGCGTCCTC	*This study*

* is arbitrary primer designed in this study.

Carcinogen treatment had maintained continuously for five culture passages, and total DNA and cellular RNA were extracted as described above.

To amplify the transposition-related region in the genome, two kinds of primers were designed; one is specific primers in flanking region of inverted terminal repeats (ITRs) from three transposons (*Mos1*, *Pogo1* and human *Hsmar1*-like element), and the other is random primers in arbitrary ITR sequences. Nucleotide sequences of the ITR primers were listed in Table 3. Amplification procedures were performed by the method of Misra *et al*. (2001). PCR products were separated on 8% non-denaturing polyacrylamide gels under 100 V constant condition for 4 hr. The bands were stained by EtBr, and then visualized with UV-illumination.

6. Extraction of Total RNA

Total RNA was extracted from cells and whole blood samples using High Pure RNA Isolation Kit (Roche, Germany) according to manufacturer's guide. Whole blood was treated with Red Cell Lysis Buffer to remove the erythrocyte prior to RNA extraction. The harvested cells from non-treated each culture of cell lines directly applied. RNA extraction from each treatment on fibroblast CCD-25Lu was performed on day 3 after treatment. Extracted RNA was dissolved in DEPC-water, and cellular DNA was removed by digestion with RNase-free DNase I (Roche, Germany).

To calculate the amount of RNA, the absorbance was measured at 260 nm and 280 nm wavelength. The RNA samples with A_{260}/A_{280} ratio over 1.8 were used for cDNA synthesis.

7. cDNA Synthesis and Semi-Quantitative RT-PCR

The first strand cDNA synthesized with 1 ㎍ DNase I-digested RNA was reverse-transcribed with the oligo$(dT)_{15}$ primer using ImpromII cDNA Synthesis Kit (Promega, USA) according to manufacturer's protocol. After cDNA synthesis, reaction mixture was diluted to final volume of 100 ㎕ by addition of 80 ㎕ DEPC-water, and the amount of synthesized cDNA was measured with absorbance test. Oligonucleotide primers used in RT-PCR amplification were listed in Table 4.

To amplify the transcripts, 100 ng of cDNA was used as a PCR template to amplify each gene, except for 500 ng of cDNA, HERV-K(T47D) LTR and HERV-K(*C4*) *env*. Amplification was performed with *Taq* DNA polymerase (Promega, USA) and appropriate primer pairs. The PCR mixture contained sterile distilled water, 2.5 ㎕ 10× PCR buffer (10mM Tris-HCl, pH 9.0, 2 mM $MgCl_2$, 50 mM KCl, and 0.1% Triton X-100), 10 mM dNTP, 2 units of *Taq* DNA polymerase, 20 nM primers, and cDNA. The PCR was performed with a GeneAmp PCR system 9600 (Perkin-Elmer Cetus, USA). The amplification was performed for 30 cycles with denaturation (45 sec at 94℃), annealing (45 sec at proper temperature according to each primer pair), and extension (1 min at 72℃). The PCR products were electrophoresed on 1.5% agarose gels.

The bands observed on a gel were quantified using Quantity-One image analysis software (Bio-Rad, USA). To examine the relative levels of each gene transcription, the intensity values of observed bands were normalized to those of the human *β-actin* gene. For cloning and DNA sequencing, RT-PCR product was purified with QIAquick PCR Purification Kit (Qiagen, USA) according to manufacturer's manual.

Table 4. List of primers used in RT-PCR analysis.

Name	Gene	Sequence (5'→3')	Source of primer	Reference
KSH45 KDS14	HERV-K10 LTR	GTATTGTCCAAGGTTTCTCCC GTGCTGTGCTTTTGGATATGC	M14123	Kim *et al*. (2001)
K1505 K1506	HERV-K(T47D)LTR	TCCCCTTGGAATACTCCTGTTTT CATTCCTTGTGGTAAAACTTTCCA	U39936	Patience *et al*. (1996)
KenvS KenvA	HERV-K10 *env*	GTAGAAGTACCTACTGTCAG CTCACCAGCAGAATACGGTG	AF490464	Sugimoto *et al*. (2001)
KC4S KC4AS	HERV-K(*C4*) *env*	ATGTACATCTCTGATCACACTATGG TAATTTGGAGTGTTTAGGGCCTA	U07586	Schneider *et al*. (2001)
Lp40S Lp40A	L1Hs *p40*	AGGCCAACGTTCAGATTCAG GCTGGTACCGGTTGTTCC	U93564	*This study*
BAS BAAS	*β-actin*	ATGGGTCAGAAGGATTCCTATG CAGCTCGTAGCTCTTCTCCA	NM_001101	Gunning *et al*. (1983)
BAXS BAXA	*BAX*	ACCAAGAAGCTGAGCGAGTGT ACAAAGATGGTCACGGTCTGCC		
IL6S IL6AS	*IL-6*	AATTCGGTACATCCTCGACG GCGCAGAATGAGATGAGTTG		
MYCS MYCA	*c-Myc*	TCTGGATCACCTTCTGCTGG GCTCCTCTGCTTGGACGGAC		
TP53S TP53A	*p53*	GTGGAAGGAAATTTGCGTGT TCTGAGTCAGGCCCTTCTGT		*This study*

8. Cloning and Sequencing

To determine the nucleotide sequences of transcripts of retroelements, DNA sequencing was performed. The purified PCR product and RT-PCR product were ligated into a pCR2.1-TOPO vector and cloned using TOPO^{TM} TA Cloning kit (Invitrogen, USA). Ligation mixture contained 2 $\mu\ell$ sterile distilled water, 2 $\mu\ell$ insert DNA, 1 $\mu\ell$ TOPO vector, and 1 $\mu\ell$ salt solution (1.2 M NaCl, 0.06 M $MgCl_2$). Ligation mixture was incubated at room temperature for 15 min. The conditions for transformation and cloning were followed by the manufacturer's manual slightly modified. Plasmid was isolated with Wizard DNA MiniPrep (Promega, USA) and confirmed the insertion of DNA by restriction enzyme digestion or nested PCR. Sequencing reaction was performed in ALFexpress DNA Sequencing System (Pharmacia Biotech, Sweden) with Cy5-labelled vector inner primers (M13-40 and M13 reverse) and Cy5-AutoCycle Sequencing Kit (Pharmacia Biotech, Sweden). The nucleotide sequences of HERV-K10 LTR and L1Hs *p40* obtained by DNA sequencing were compared with previously reported sequences using the BLAST search program in GenBank.

9. SignalScan of the Signal Responsive Sequences of HERV-K LTRs

Analysis of putative signals was screened by SignalScan ver. 4.1 (Prestridge, 1991) with *Mammal and Virus* option using TRANSFAC database (Wingender, 1994). Signal analysis of LTRs of HERV-K10 and HERV-K(T47D) was performed using the sequence from M14123 (Ono *et al*., 1986) and from U39936 (Patience *et al*.,

1996) previously reported in GenBank, respectively.

10. Database Search of the Retroelement Sequences in Human Genome Resources

To analyze the associations between chromosome locations of transcriptionally active retroelements and the human genetic disorders, LOCUSLINK was used to determine the chromosomal map locations in Human Genome Resources. Chromosomal fragile sites previously reported were also screened using LOCUSLINK. OMIM (Online Mendelian Inheritance in Man) was used to search the locus-related genetic disorders and cancers.

11. Protein Sequence Analysis of L1Hs p40

The standard active L1Hs p40 protein sequences retrieved from U93573 (Sassaman *et al*., 1997) were analyzed using PROSITE database (Bairoch *et al*., 1997) in Network Protein Sequence (http://npsa-pbil.ibcp.fr). To scan a sequence for sites/signatures PROSCAN carried out against PROSITE database, and a helix-turn-helix motif did HTH (Dodd and Egan, 1990). Phylogenetic analysis on putative protein sequences obtained from the result of translation of L1Hs *p40* nucleotide sequences were carried out using CLUSTAL W (Thompson *et al*., 1994) for multiple alignment and PHYLIP package (Felsenstein, 1993) for phylogenetic relationship. Protein distance was calculated with the program PROTDIST.

12. Phylogenetic Analysis of Retroelement Transcripts

DNA sequences were multiple aligned using CLUSTAL W (Thompson *et al*., 1994), and were analyzed using PHYLIP program packages (Felsenstein, 1993). The sequence boundaries of each gene were determined by comparison with those previously reported (Table 3). The multiple alignment of obtained sequences and those previously reported was performed using CLUSTAL W program with the default settings, and then alignment was adjusted manually in order to align several regions of conserved sequences. Genetic distances among sequences were calculated using DNADIST program with the option of two-parameter method (Kimura, 1980) in PHYLIP. A bootstrap analysis was done using 1,000 replicating data sets generated with the SEQBOOT program. Then, CONSENSE program was used to construct a strict consensus tree.

Ⅲ. RESULTS

Until now, expression of retroelements has been observed in several normal tissues and cancers. To examine the associations between the expression of functionally active retroelements and human cancers, this study has been analyzed the transcription patterns, the effects of carcinogen treatment on gene expression and transposition activity, transcript sequences, and the chromosomal locations of active retroelements HERV-K and L1Hs in human cells.

1. Human endogenous retrovirus-K

1) Cell type-specific transcription of the HERV-K genes

In normal fibroblasts, the pattern of HERV-K transcription was examined using RT-PCR analysis. Different transcription levels were observed (Fig. 4). After standardization to human *β-actin* mRNA, the relative levels of HERV-K *env* genes and HERV-K LTRs transcription were different between CCD-25Lu and HEL-299 cells. The relative transcription of K10 LTR, K10 *env*, and T47D LTR was higher in CCD-25Lu cells than in HEL-299 cells, while that of C4 *env* expression was lower. Lung carcinoma A-549 showed high expression in HERV-K LTRs and C4 *env*, but not of K10 *env*, compared to normal lung fibroblasts (CCD-25Lu and HEL-299 cells). In the leukemia cell lines, T lymphoblast cells (Molt-4 and CEM-CM3) expressed relatively higher levels of HERV-K genes than did HL-60, U-937, or Jurkat. The transcription of K10 *env* gene was relatively higher in several cancer cells and fibroblasts. The C4 *env* gene was also

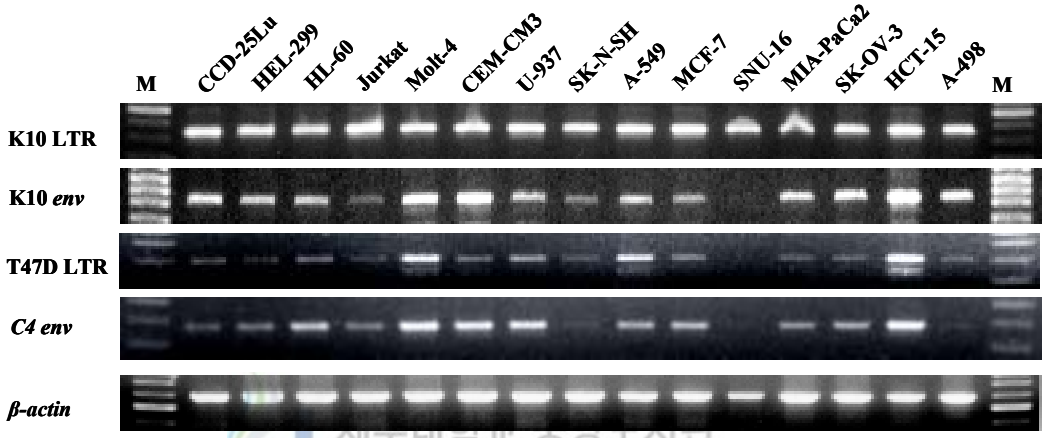


Figure 4. RT-PCR analysis of the HERV-K genes in human cells. 100 ng of cDNA was used as a PCR template for amplifying the K10 LTR and K10 *env*, and 500 ng of cDNA for T47D LTR and C4 *env*, respectively. M is the DNA size marker, GenRuler DNA ladder.

transcribed in a variety of cells except for SNU-16, but the relative transcription levels were different remarkably (Fig. 5). In the analysis, no differences in the relative transcription levels of normal fibroblasts and cancer cells was detected. While the relative transcription of each gene differed in the cell line tested. In particular, the transcription of HERV-K10 *env*, HERV-K(C4) *env*, and HERV-K(T47D) LTR showed cell-type specificity, as described previously (Majors, 1990; Ting *et al*., 1992; Kjellman *et al*., 1999; Sugimoto *et al*., 2001; Andersson *et al*., 2002).

2) Various signals can effect on the activation of HERV-K

A conserved region was identified by BLAST search and multiple alignment of HERV-K LTRs. As reported, putative signal responsive elements on LTR include promoter, enhancer, and hormone responsive elements (Baust *et al*., 2001; Schön *et al*., 2001). A computer analysis of HERV-K10 LTR sequences was performed using the TRANSFAC database (Wingender, 1994). Figure 6 shows putative signal response sites in HERV-K10 LTR. The viral TATA box and enhancers were found, and the TATA box is the potential region for binding of various transcription factors (TMF, UBP-1, TFIID, TFIIF, TRF, TBP, and NF-E2). In addition, various transcription factor response elements (AP-1, CAC binding protein, GATA-1, Sp1 and TGT3; hormone response elements, GR for glucocorticoid and window 4 and 10 for the human epidermal growth factor) and the oncoprotein c-MYC binding element were also detected.

HERV-K(T47D) LTR has the signal responsive elements, for example, binding sequences for transcription factors (AP-1, AP-2, NF-1, NF-2, NF-E, TFIID, TBP, Sp1, and Pit-1), oncoproteins (v-JUN and c-MYC), and hormone (GR) (Fig. 7).

3) Carcinogens induced the various effects on gene expression of the HERV-K and regulatory proteins in CCD-25Lu

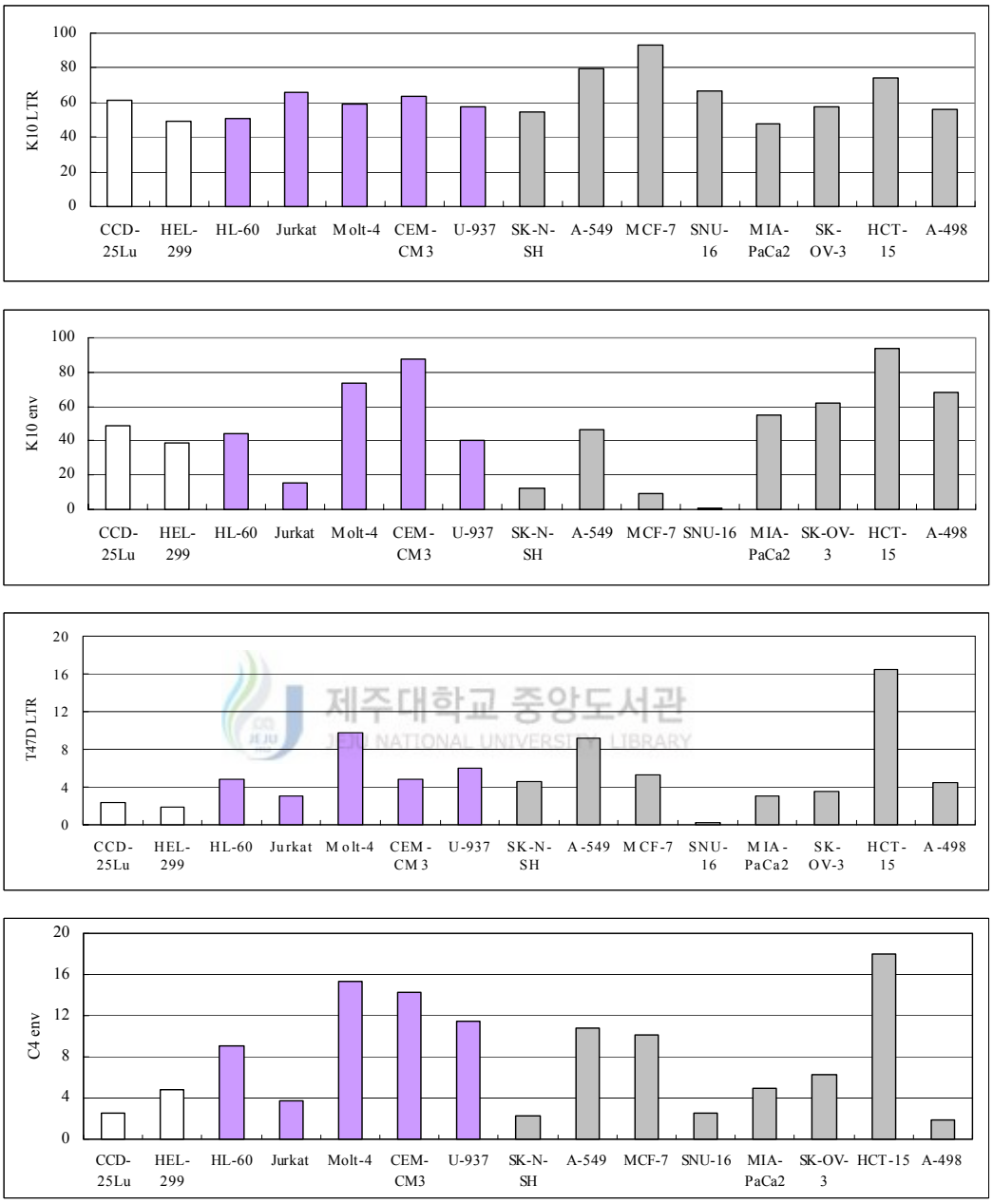


Figure 5. Relative levels of the HERV-K gene transcription in human cells. 100 ng of cDNA was used as a template for amplifying the K10 LTR and K10 *env*, and 500 ng for T47D LTR and C4 *env*. RT-PCR bands were standardized by comparing the ratio of each HERV-K gene to that of the human *β-actin* with Quantity One software (Bio-Rad).

Figure 6. Signal scan of the HERV-K10 LTR using the TRANSFAC signal database with human and virus option.

Figure 7. Signal scan of the HERV-K(T47D) LTR using the TRANSFAC signal database with human and virus option.

To screen the effects on transcription of HERV-K by carcinogens, CCD-25Lu cells were treated with various chemical carcinogens. The carcinogen treatments had various effects on CCD-25Lu cells. The cytotoxic effects by DNA fragmentation assays also were different. High concentration of MNNG and EtBr induced cell death. Degraded DNA fragments produced in cellular processes of the activated apoptotic pathway were observed in fibroblasts treated with carcinogens, which the pattern of DNA degradation varied according to the carcinogen. Using RT-PCR, the result showed that carcinogens alter the expression of HERV-K genes (Fig. 8) and regulatory protein genes. Cellular mRNA contained more RNA transcripts of HERV-K genes in all treatments, except HERV-K10 LTR in EtOH, EtBr, and MNNG treatment (Fig. 9). Of the regulatory genes, *BAX* mRNA was increased by all the carcinogens, *c-MYC* mRNA reduced by EtBr, and *IL-6* was decreased by MNNG (Fig. 10). The expression of *p53* was reduced by DMBA, EtOH, EtBr, MNNG and TPA in CCD-25Lu fibroblasts (Fig. 11).

4) Carcinogens induced the transposition activity in CCD-25Lu

Screening of transposition events has been examined using random amplification of the genomic DNAs and cDNAs, which extracted from the CCD-25Lu fibroblasts exposed to each carcinogen for 7 passages. After average 5 passages, an distinguishable amplification patterns have been shown (Fig. 12 and 13). Amplification with the arbitrary and transposon specific ITR primers showed that altered patterns were dependent on culture passages. In addition, RT-PCR analysis on cDNA using the same primers has also been shown the differential fragments on gels at 5 passages of carcinogen treatments. Transcription patterns had also been altered by carcinogen treatment in CCD-25Lu fibroblasts. Comparing to saline treated control, carcinogen specific bands were found in the treatments, and relative levels of each amplified band has also been remarkably different (Fig. 14).

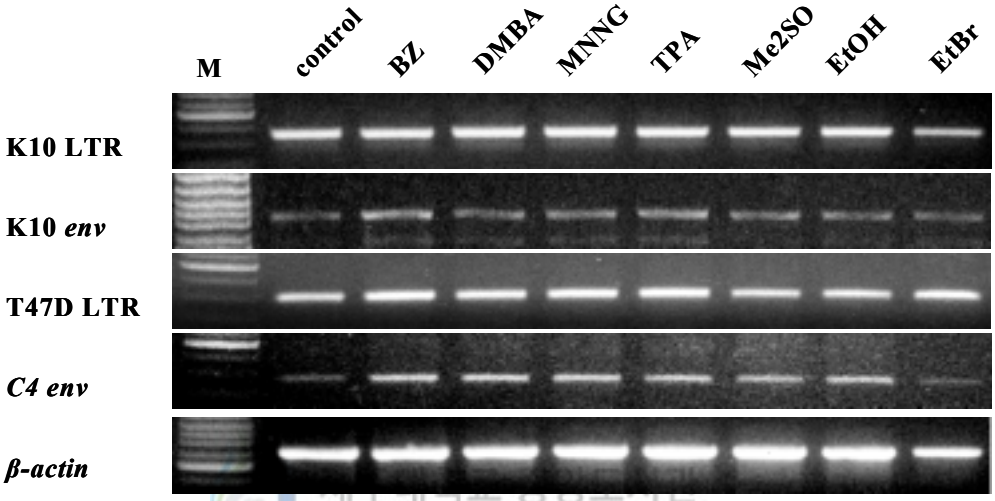


Figure 8. RT-PCR analysis of the HERV-K gene transcription in CCD-25Lu cells treated with carcinogens.100 ng of cDNA was used as a RT-PCR template for amplifying the K10 LTR and K10 env, and 500 ng of cDNA for T47D LTR and C4 env, respectively. M is the DNA size marker, GenRuler DNA ladder.

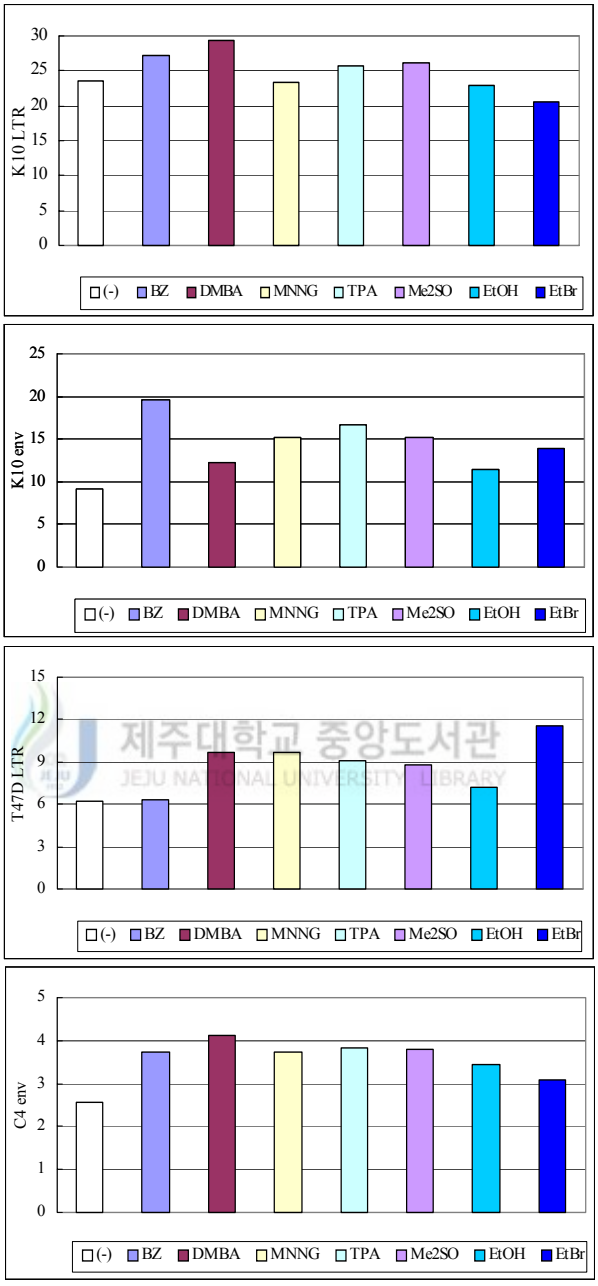


Figure 9. Relative levels of the HERV-K gene transcription in CCD-25Lu cells treated with carcinogens. 100 ng of cDNA was used as a template for amplifying the K10 LTR and K10 *env*, and 500 ng for T47D LTR and *C4 env*. Standardization for the transcription level of each gene was performed as described in Fig. 5.

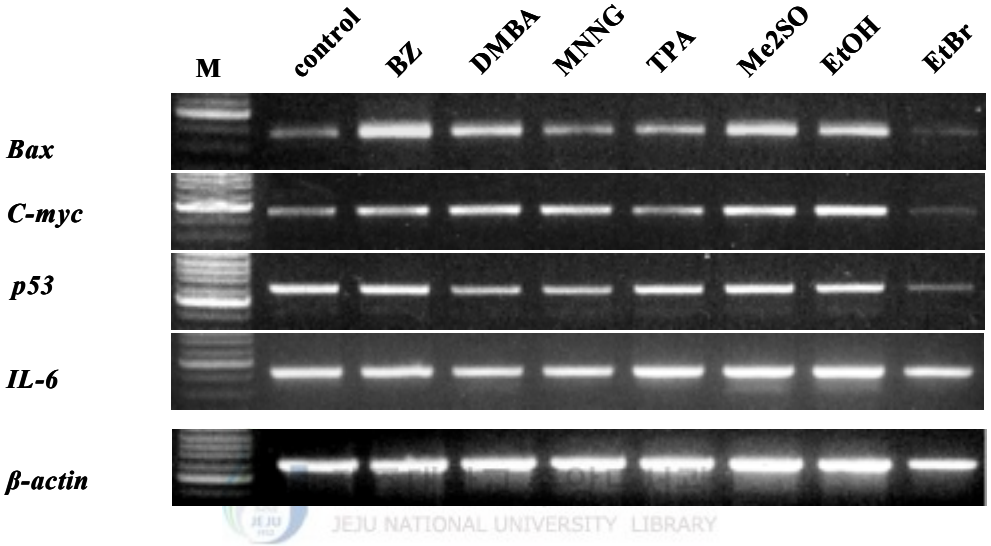


Figure 10. RT-PCR analysis of the regulatory protein gene transcription in CCD-25Lu cells treated with carcinogens. 100 ng of cDNA was used as a RT-PCR template for amplifying each gene. M is the DNA size marker, GenRuler DNA ladder.

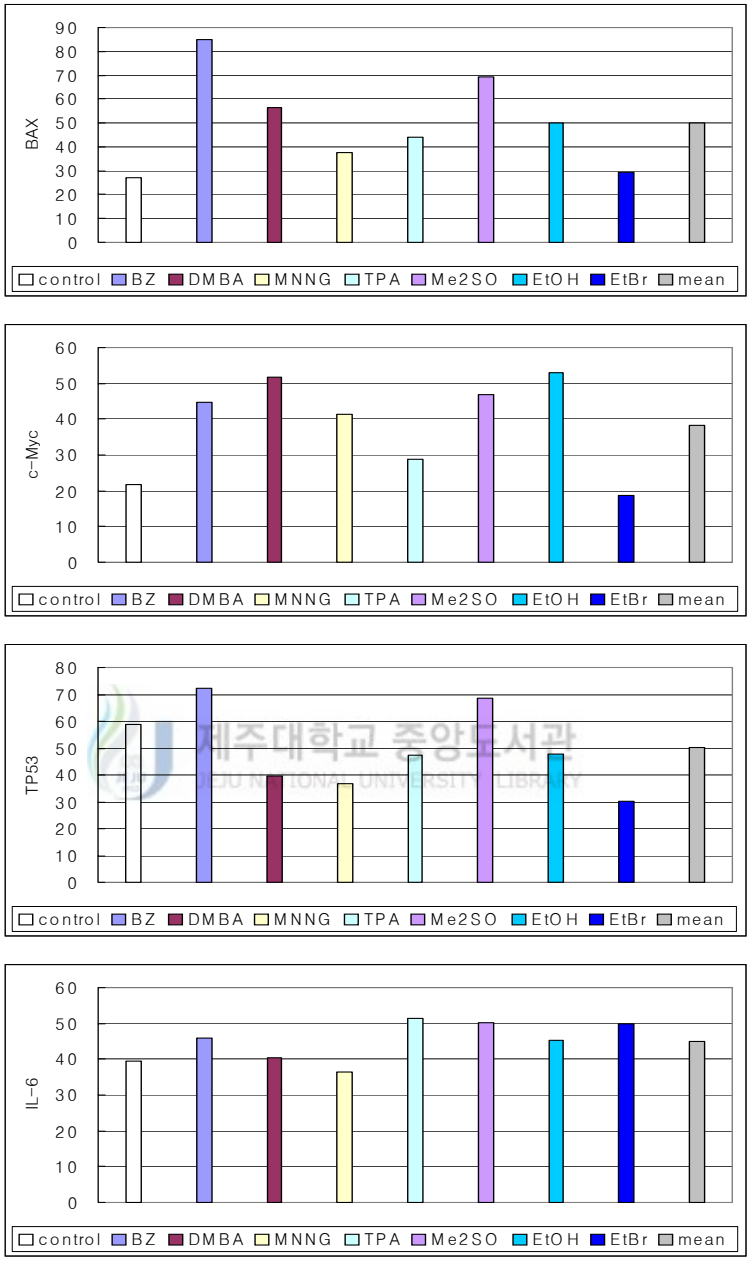


Figure 11. Relative levels of the regulatory gene transcription in CCD-25Lu cells treated with carcinogens. 100 ng of cDNA was used as a RT-PCR template for amplifying each gene. Standardization for the transcription level of each gene was performed as described in Fig. 5.

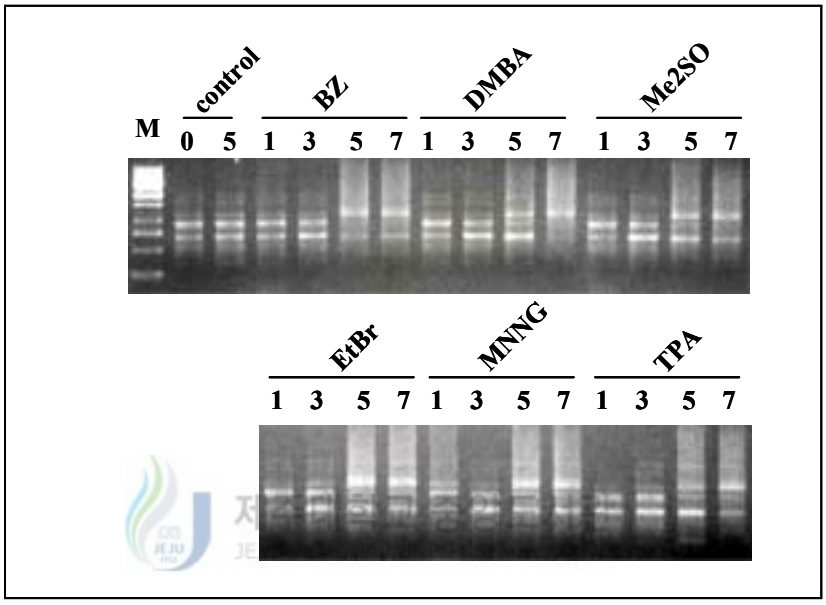


Figure 12. Transposition analysis on genomic DNAs of CCD-25Lu cells treated with carcinogens for 7 culture passages. Amplification was performed using arbitrary random priming ITR, HSH103. The number on the top of each lane indicates the culture passage. M is the DNA size marker, 1kb ladder plus.

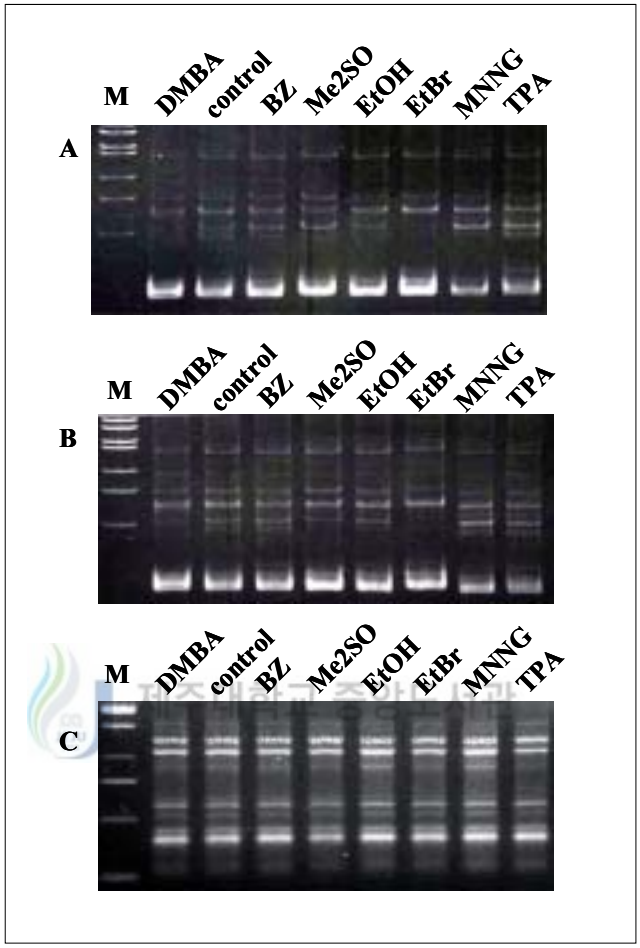


Figure 13. Transposition analysis on the carcinogen treatments after 5 culture passages. A, Amplification patterns on CCD-25Lu cells using HSH103; B, those on HL-60 cells using HSH103; C, those on SK-N-SH cells with *Hsmar1*. M is DNA size marker, 1kb ladder plus.

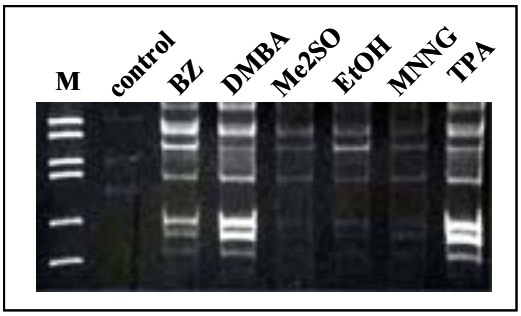


Figure 14. Transposition analysis on CCD-25Lu cDNAs treated with carcinogens after 5 culture passages using transposon specific ITR primer *Pogo1*. 100 ng of cDNA was used as a template for amplifying. M is DNA size marker, 100 bp ladder plus.

5) HERV-K10 LTRs transcribed from various locations on chromosomes

To investigate active HERV-K10 LTR, the RT-PCR products from the cell lines were sequenced. The characteristics and chromosomal locations of HERV-K10 LTR elements transcribed from human cells are shown in Table 5. The BLAST result confirmed that the HERV-K10 LTR sequences from various human chromosomes were identical. 'K10 LTR groups' have been defined as identical sequences, appearing at more than two loci on different chromosomes or accession numbers. For example, HERV-K10 LTR group 1 occurs in twenty-four identical sequences at fourteen different locations on nine chromosomes (Table 6 and Table 7). The sixty clones sequenced from normal cells were made up of sixteen elements. RNA intermediates of HERV-K10 LTR should be transcribed from as least sixteen positions in the haploid human genome. They had lengths of 297 or 298 bp, and the sequence similarity ranged from 97-100%. From cancer cells, a hundred and fifty clones were sequenced. Their lengths were more vary than those of elements from normal cells, *i. e.*, from 287 to 306 bp, and the sequence similarity ranged from 95-100%. Several HERV-K10 LTR elements were identical to those previously reported by Yi *et al*. (2001).

6) Several locations of HERV-K10 LTR transcripts were specific for cancer cells

The putative chromosomal locations were investigated using a BLAST search of the Human Genome database in GenBank. Forty-four elements were thought to be derived from single location on one chromosome, while the others had multiple origins. The HERV-K10 LTR was found from the centromere to the telomere. Four HERV-K10 LTRs (2F7-1, MC-5, MP-4, and SC-1) might be transcribed from the telomere regions of the four chromosomes (chromosome 4, 8, 9, and 15). Cell-type specific HERV-K10 LTR elements were found, and they were transcribed from different loci of chromosomes. For example, nine clones of HERV-K10 LTRs transcribed from 1q21 were observed in six

Table 5. Chromosomal locations of the HERV-K10 LTR transcripts from the human cells

Cell line	Clone	Length	Similarity (%)	The most similar sequence
KYJ	YJ-1	297	99	K10 LTR group 4; AC023162.41(3q21), AL449213.4(3q21)
	YJ-2	297	99	AC114933.2(5p14)
	YJ-3	297	99	K10 LTR group 1; AB023523.1(ND*), AB059364.1(ND), AF276875(ND) AL139421.12(1p22) AC006039.3(7p15) AC100821.2(8q11) AL354855.26(9q34) AK075315.1(11q13), NM_173578.1(11q13), AP003385.2(11q13), AP000812.4(11q13), AP000776.5(11q22), AC023281.13(12p11), AC048344.44(12p11) AC025420.26(12q13) AC107977.13(15q25), AC129907.3(15q25), AC109513.14(15q25), AC140725.3(15q25) AC012175.7(16p13), AC002400.1(16p12), AC008870.8(16q12), AC009132.9(16q23) AL031654(20q12)
	YJ-4	298	99	Group 3
	YJ-5	297	99	Group 3
KBG	BG-1	297	100	AL049570.11(1p35)
	BG-2	297	98	BC036118.1(8q24)
	BG-3	297	97	AP005203.3(18p11)
	BG-4	297	98	Group 1
	BG-5	297	100	AL392107.17(10q24)
	BG-6	297	98	Group 2
CCD-25Lu	CC-1	297	99	Group 3
	CC-2	297	99	K10 LTR group 2; AL049570.11(1p35), AL512637.18(1q44) AC084198(3q12) AF235103.5(8q25), BC036118.1(8q25) AL135901.23(13q14) AC026803.7(19q13)
	CC-3	297	98	K10 LTR group 3; AB023542.1(ND), AB023539.1(ND) AC016839.12(18q11), X54171.1(18q11)
	CC-4	297	99	AL135901.23(13q14)
	CC-5	298	99	AC025674.10(8q11)
	CC-6	297	99	AL135901.23(13q14)
2F-7	2F7-1**	291	100	Z95704.1(4p16)
A-431	A3-1**	297	99	AC019155.4(7q32)
	A3-2**	297	98	AL356019.5(14q11)
	A3-3**	297	99	AC105049.4(8q24)
A-498	A4-1	297	99	K10 LTR group 6; AC079753.7(2q14), BN000002.1(2q14)
	A4-2	298	100	AC025674.10(8q11)
	A4-3	297	98	Group 6
A-549	A5-1	297	100	AL392107.17(10q24)
	A5-2	297	99	AL592284.10(1q21)
	A5-3	294	98	AC020924.8(5q12)
	A5-4	297	99	AC105049.4(8q24)

Table 5 (Continued)

Cell line	Clone	Length	Similarity (%)	The most similar sequence
HCT-15	HC-1	297	97	AC008784.7(5q11)
	HC-2	297	99	K10 LTR group 7; AC104150.6(3q26), AC104640.3(3q26) AC080023.6(11p15), AP003716.4(11q12)
	HC-3	294	95	K10 LTR group 8; AF017229.1(ND), Y17833.1(ND) AC112702.3(19q12), AC010508.8(19q12)
	HC-4	294	96	Group 8
	HC-5	297		AL356019.5(14q11)
Hep-G2	HG-1	297	99	AL392107.17(10q24),
	HG-2	298	99	K10 LTR group 9; BC041646.1(ND), BC011367.1(ND), AF164610.1(ND), K090528.1(ND) AL353807.18(1q22) AC021133.5(6p12), AL121969.12(6p12)
	HG-3	297	99	K10 LTR group 10; BC001407.1(ND)* AL713888.10(10q21)
	HG-4**	297	99	K10 LTR group 11; AF394944.1(Xq26), AL034407.1(Xq26)
	HG-5**	287	100	AC004979(4q31)
HL-60	HL-1	294	100	AL671879.2(6p21)
	HL-2	306	99	AC011447.7(19p13)
	HL-3	299	95	AC053544.5(4q13)
	HL-4	298	99	AL592284.10(1q21)
	HL-5	298	100	AC007283.3(2q33)
Jurkat	JU-1	293	99	AL590543.8(6q25)
	JU-2	297	98	AC109327.10(17q25)
	JU-3	290	97	K10 LTR group 12; BX546487.5(ND) AC128674.3(1q21) AC127384.4(1q21), AL954711.3(1q21)
	JU-4	298	97	AL049742.8(1q21)
	JU-5	297	98	K10 LTR group 13; AC109322.16(8q24), AC105049.4(8q24) AL445931.29(9q34), AC002103(9q34), AC000387.1(9q34) AC000401.1(9q34), AL591386.18(9q34)
	JU-6**	298	99	AL592284.10(1q21)
	JU-7**	297	99	AL391244.11(1p36)
MCF-7	MC-1**	297	100	AL139404.9(10q11)
	MC-2**	297	100	AC087533.7(8p11)
	MC-3**	297	98	Group 7
	MC-4**	297	98	AC000387(9q34)
	MC-5**	297	100	K10 LTR group 14; AC000387(9q34), AL591386.18(9q34)
MIA-PaCa-2	MP-1**	287	100	AC004979(4q31)
	MP-2**	298	100	AC007283.3(2q33)
	MP-3	298	97	AL592284.10(1q21)
	MP-4	297	99	AC140725.3(15q26)
	MP-5	298	98	AL592284.10(1q21)
	MP-6	288	98	AC011322.29(3q29)
	MP-7	297	99	AL391244.11(1p36)

Table 5 (Continued)

Cell line	Clone	Length	Similarity (%)	The most similar sequence
SNU-C5	SC-1	297	98	Group 2
	SC-2	296	99	AC003976(17q21)
SNU-16	SU-1	297	99	AC019155.4(7q32)
	SU-2	295	100	K10 LTR group 5 AP000346.1(22q11), AP000345.1(22q11)
	SU-3	296	98	Group 6
	SU-4	294	95	AC104301.2(17q21)
	SU-5	298	98	AL592284.10(1q21)
U-937	U9-1	297	99	AL513550.9(6q16)
	U9-2	298	98	AL049742.8(1q21)
	U9-3**	297	100	Group 1
	U9-4**	297	99	Group 1
	U9-5**	297	99	AC072054.10(7p22)
	U9-6	298	99	AL391244.11(1p36)

* not determined in Human Genome Resources in GenBank.

** is previously reported by Yi *et al*. (2001).

Table 6. HERV-K10 LTR groups found in the similarity search

Group	No. of identical sequence	No. of location*	No. of chromosome**
K10-1	24	14	10
K10-2	7	6	5
K10-3	4	1	1
K10-4	2	1	1
K10-5	2	1	1
K10-6	2	1	1
K10-7	4	3	2
K10-8	4	1	1
K10-9	7	2	2
K10-10	2	1	1
K10-11	2	1	1
K10-12	4	1	1
K10-13	6	2	2
K10-14	2	1	1

* is determined by BLAST search in GenBank database.

** is determined by LOCUS search in Human Genome Resources in GenBank.

Table 7. HERV-K10 LTR groups and identical sequence clones from the cell lines tested

Cell line	No. of clones tested	No. of K10 LTR group*														No. of clones found in K10 LTR groups
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	
YJ	20	7	4	1	-	-	2	1	-	-	-	-	-	1	-	16
BG	20	8	3	1	1	-	1	1	-	-	-	-	-	1	-	16
CCD-25Lu	20	6	2	1	-	-	-	2	-	2	-	-	-	1	-	14
A-498	15	6	2	-	-	-	1	-	1	1	-	1	-	-	-	12
A-549	15	3	4	-	1	-	-	1	-	1	-	-	-	1	-	11
HCT-15	15	4	2	1	1	-	-	1	-	-	1	-	-	-	-	10
Hep-G2	15	3	3	1	-	-	-	-	-	2	-	-	1	1	1	12
HL-60	15	4	5	-	-	-	-	-	-	-	-	-	1	-	-	10
Jurkat	15	3	3	-	-	-	-	-	-	2	-	-	-	2	-	10
MIA-PaCa-2	15	3	-	2	-	-	-	2	1	-	-	-	1	1	-	10
SNU-C5	15	6	2	1	-	-	1	-	-	1	-	-	1	1	-	13
SNU-16	15	2	4	1	1	1	-	1	-	1	-	-	-	-	-	11
U-937	15	3	3	1	-	-	-	2	-	1	-	-	-	1	1	12
Total	210	58	37	10	4	1	5	11	2	11	1	1	4	10	2	157

* group numbers are given in Table 6.

cancer cells (A-549, HL-60, Jurkat, MIA-PaCa-2, SNU-16, and U-937) and three clones from 1p36 in three cancer cells (Jurkat, MIA PaCa-2, and U-937), but not in normal cells (Table 8). In addition, HERV-K10 LTR transcripts from the eight chromosomal locations (2q14, 2q33, 4q31, 7q32, 8p11, 10q11, 14q11, and 17q21) were shared two kinds of cancers. For example, HERV-K10 LTR clones transcribed from 2q14 were observed in skin carcinoma A-431 and stomach adenocarcinoma SNU-16. The six locations (1q21, 2q14, 9q34, 13q14, 18q11, and 19q13) showed more than two clones from each single cell line. In these, three clones (JU-3, JU-4, and JU-6) transcribed from the same location 1q21 in jurkat. However, some transcripts from the six locations (1p35, 3q31, 5p14, 13q14, 18p11, and 18q11) were specific for normal cells. Those from 8q11 and 10q24 were observed in normal cells as well as in cancer cells. Except for HERV-K10 LTR groups, cell line specific transcripts have not found in chromosomes 11, 12, 16, 20, 21, and Y (Table 8).

7) HERV- K10 LTR elements divided into several clusters

Sixty-two different HERV-K10 LTRs obtained from human cells were subjected to phylogenetic analysis, along with twenty-eight elements, previously reported from the human genome, nine from chimpanzees (*Pan troglodytes*: AB023912, AB023913, AB023915, AB023916, AB023917, AB023918, AF311937, and M57949), two from macaques (*Macaca mulatta*: AB024015 and AB024016), two from gorillas (*Gorilla gorilla*: AB023923 and AB023921), and one from common squirrel monkey (*Saimiri sciureus*: AB023924). The neighbor-joining (NJ) tree constructed using DNA distances had three main clusters; a branch containing the common ancestral elements of primates, elements shared by human and macaques, and human-specific type I elements (Fi. 15). The human-macaques cluster was further subdivided into human-specific type II elements and other elements. In addition,

Table 8. Cell line-specific transcripts of the HERV-K10 LTR and its chromosomal locations

Chromosomal location	Symbol of clone*	No. of clone
1p36	JU-7, MP-7, U9-6	3
1p35	BG-1	1
1q21	A5-2, HL-4, JU-3, JU-4, JU-6, MP-3, MP-5, SU-5, U9-2	9
2q14	A4-1, A4-3, SU-3	3
2q33	HL-5, MP-2	2
3q21	YJ-1	
3q29	MP-6	1
4q13	HL-3	1
4q31	HG-5, MP-1	2
5p14	YJ-2	1
5q11	HC-1	1
5q12	A5-3	1
6p21	HL-1	1
6q16	U9-1	1
6q25	JU-1	1
7p22	U9-5	1
7q32	A3-1, SU-1	2
8p11	A3-2, MC-2	2
8q11	CC-5, A4-2	2
8q24	A3-3, A5-4, BG-2	3
9q34	MC-4, MC-5	2
10q11	HG-3, MC-1	2
10q21	HG-1	1
10q24	BG-4, A5-1	2
13q14	CC-4, CC-6	2
14q11	A3-2, HC-5	2
15q26	MP-4	1
17q21	SC-2, SU-4	2
17q25	JU-2	1
18p11	BG-3	1
18q11	YJ-4, YJ-5, CC-1, CC-3	4
19q12	HC-3, HC-4	2
19q13	HL-2	1
22q11	SU-2	1
Xq26	HG-4	1

* is given in Table 5.

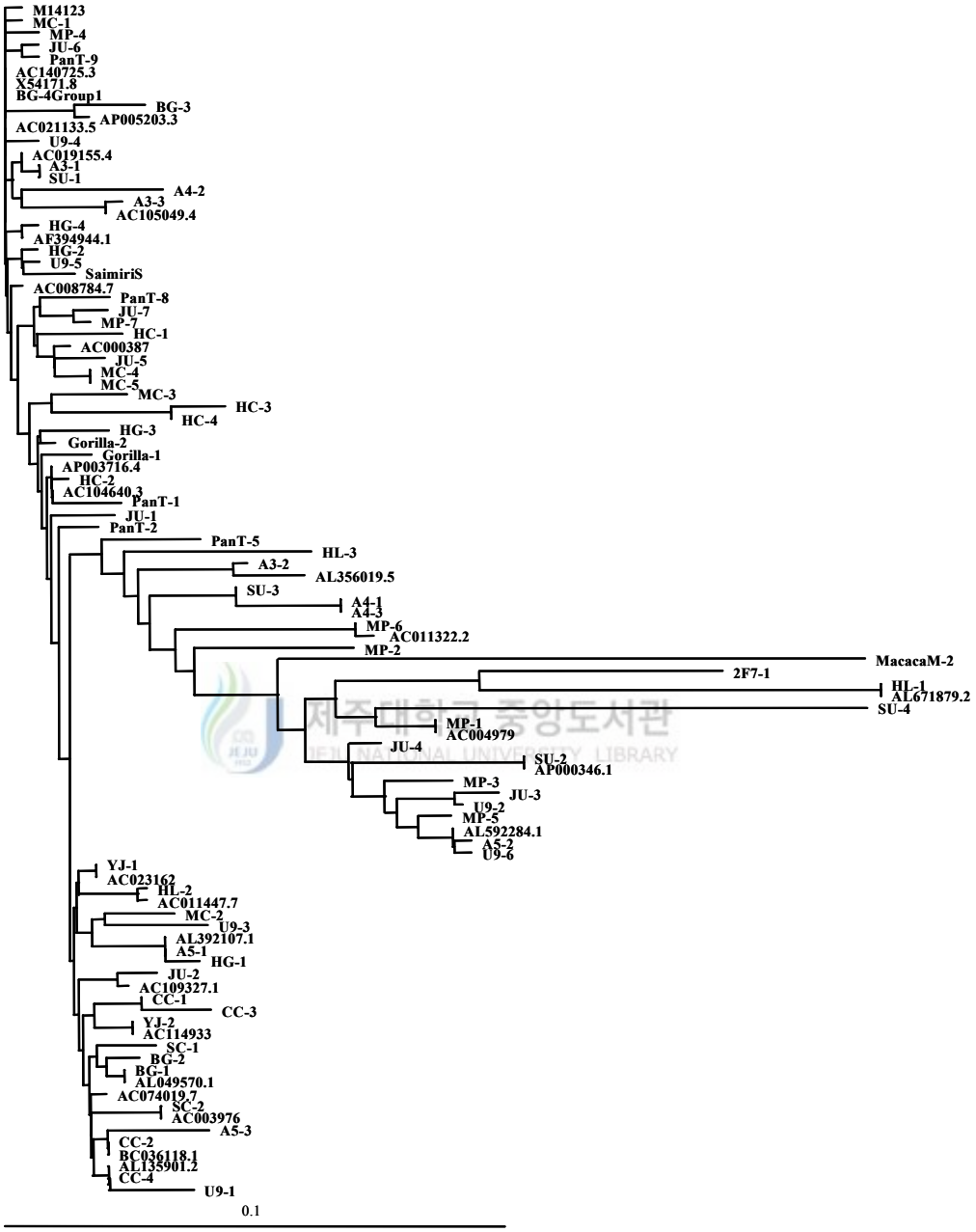


Figure 15. Neighbor-joining tree based on the sequences of the HERV-K10 LTR transcripts in the human cells. Abbreviations of each HERV-K10 LTR clone were represented in Table 4. Gorilla, *Gorilla gorilla*; MacacaM, *Macaca mulatta*; Pant, *Pan troglodytes*; SaimiriS, *Saimiri sciureus*.

the cluster of common ancestral elements was subdivided into two subclusters: human-chimpanzee-gorilla and human-chimpanzee-common squirrel monkey.

2. Human long interspersed nuclear element-1

1) L1Hs *p40* elements were transcribed in various human cells

In RT-PCR analysis, L1Hs *p40* transcripts were found in all cell lines tested, however, the transcripts could not be clearly distinguished among the cell lines tested (Fig. 16). Cellular mRNA for two normal fibroblasts (CCD-25Lu and HEL-299) also contained L1Hs *p40* transcripts, but the relative levels of transcription had not significantly distinguished. The transcription levels of A-549, MCF-7, HCT-15 and SNU-16 were higher than those of leukemic cells and fibroblasts (Fig. 17).

2) Carcinogens changed the level of L1Hs *p40* transcription in CCD-25Lu

RT-PCR analysis of mRNAs extracted from cultures treated with carcinogens were shown higher transcription levels, compared to control treated with saline (Fig. 18). Figure 19 shows the relative levels of L1Hs *p40* transcription in each carcinogen treatment. Compared to *β-actin*, DMBA treatment caused the highest transcription rate, and EtBr treatment the lowest. Different expression patterns have been observed in the EtBr and MNNG treatments, and resulted in cell death.

3) L1Hs p40 transcribed from various locations on the chromosomes

Using RT-PCR products from four normal cells, fourteen new L1Hs elements were

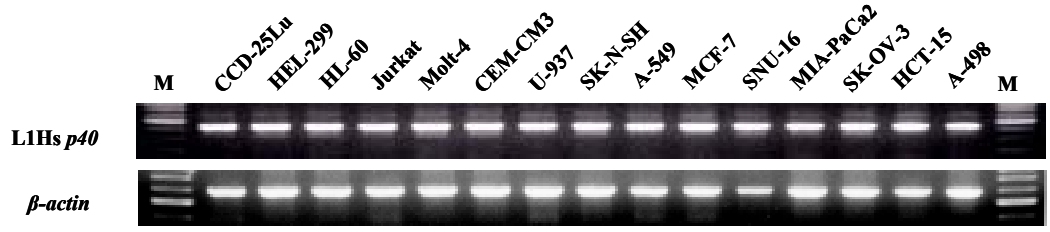


Figure 16. RT-PCR analysis of the L1Hs *p40* transcription in the human cells. 100 ng of cDNA was used as a RT-PCR template for amplifying L1Hs *p40*. M is the DNA size marker, GenRuler DNA ladder.

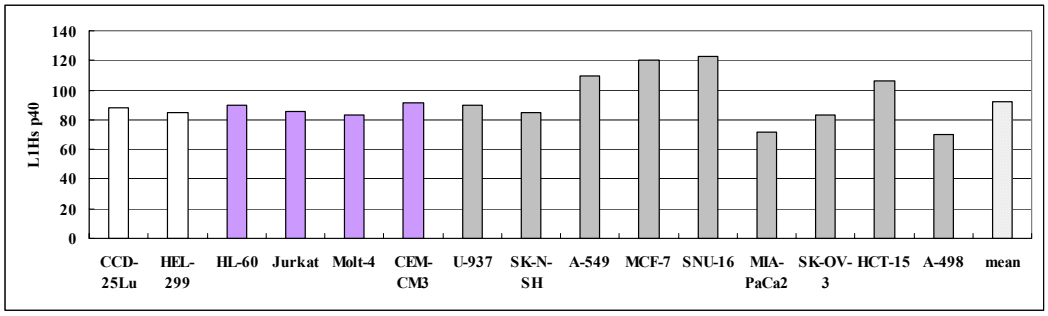


Figure 17. Relative levels of the L1Hs *p40* transcription in the human cells. 100 ng of cDNA was used as a RT-PCR template for amplifying L1Hs *p40*. RT-PCR bands were standardized by comparing the ratio of the L1Hs *p40* to that of the human *β-actin* with Quantity One software (Bio-Rad) as described in Fig. 5.

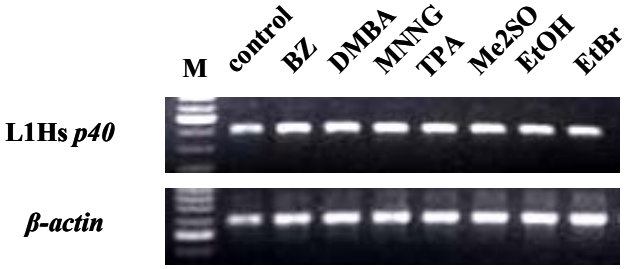


Figure 18. RT-PCR analysis of the L1Hs *p40* transcription in CCD-25Lu cells treated with carcinogens. 100 ng of cDNA was used as a PCR template for amplifying each gene. M is the DNA size marker, 100 bp ladder plus.

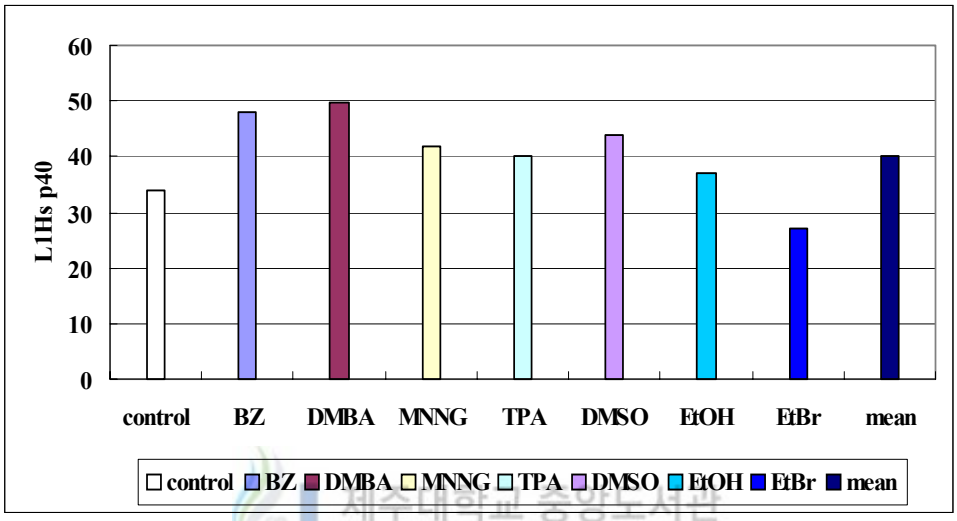


Figure 19. Relative levels of the L1Hs *p40* transcription in CCD-25Lu cells treated with carcinogens. 100 ng of cDNA was used as a PCR template for amplifying L1Hs *p40*. Standardization for the transcription level of each gene was performed as described in Fig. 17.

identified. Length variation was found in several sequences, which ranged from 376 to 382 bp (Table 9). The BLAST analysis showed that many of the determined L1Hs element sequences were identical to previously reported sequences (Table 10 and Table11). Furthermore, several copies of some identical sequences were found on several chromosomes. These were defined as 'L1Hs group', and thirteen distinct L1Hs groups were found (Table 10). For instance, in the human genome, L1Hs group 1 contains sixty-five identical sequences at fifty-one locations on twenty chromosomes excluding chromosomes 17, 18, and 21. Of the L1Hs from normal cells, ten elements are transcribed from a single locus on one chromosome. The chromosome locations were distributed over all chromosomes, excluding chromosome 21, and ranged from the centromere to the telomere. The length of the L1Hs *p40* gene, determined from various cancer cells, varied from 363 to 439 bp. Many L1Hs elements had the typical length of 379-bp. Clones HL-1 and JU-4 contained 5' additional sequences, and MP-2 had a 40-bp tandem duplications in the 5' region (Appendix 2). Forty-three sequences from cancer cells were transcribed from a single genome location.

To confirm the chromosomal locations and distributions of L1Hs elements, FISH analysis was performed with normal leukocyte cells. FISH analysis of L1Hs in the metaphase showed that it was a multi-copy retroelement, distributed on all chromosomes. FISH analysis using an L1Hs probe showed that there were over five thousand L1Hs element loci in a single metaphase preparation (Fig. 20)

4) Cell line specific L1Hs elements were transcribed from different chromosomal locations

In several cell lines, the specific transcripts of L1Hs *p40* elements were observed. For instance, several L1Hs elements (transcribed from 1p13, 1q42, 3p24, 4p16, 5p13, and Xq13) were transcribed in normal cells but not in cancer cells tested. In these, three

Table 9. Chromosomal locations of the L1Hs *p40* transcripts from the human cells

Cell line	Clone	Length	Similarity (%)	The most similar sequence
CCD-25Lu	CC-1	378	98	AL139400.9(Xq13)
	CC-2	379	98	L1Hs group1; U93564.1(ND*), U93567.1(ND), U93573.1(ND), U93574.1(ND) AL354949.10(1p31), AL683886.8(1p13), AL357045.10(1p11), L021069(1q24), AL162431.17(1q25), AL592494.4(1q31), AC096569.1(2p13), AC015971.4(2p11), AC105762.4(2q14), AC104088.6(2q31), AC104820.4(2q31), AC012486.9(2q33), AF312225(2q34), AC138658.1(2q35) AC099664.2(3p12), AC046144.15(3q13), AC109583.2(3q21), AC011597.27(3q22) AC114806.5(4p15), AC093607.3(4p15), AC013724.8(4q12), AC106884.5(4q13), AC092663.2(4q13), AC004704(4q25), AC096898.6(4q28), AC125376.3(4q32), AC020599(4q32), AC010450.6(5p15), AC091872.3(5p15), AC008948.8(5q13) AL031768.9(6p25), AL512428.13(6p22), AL591051.7(6q22), AL589693.3(6q25), AC004694.1(7p15), AC079855.8(7q11), NG_000004(7q21), AC006992.2(7q35) AC090150.9(8p21) AL160001.10(9q31) AL391260.13(10q24) AC139427.3(11q13) AC024934.3(12q12), AC025263(12q14), AC095347.6(12q21), AC093025.5(12q21) AL138694.18(13q34) AL160234.3(14q23), AL137191.5(14q23) AC073940.4(15q15) AC007347(16q12) AL161938.6(20p12), AL096677.21(20p12) AF421375(22q11), AL078622.7(22q23) AL663118.2(Xp11), AL356317.8(Xq11), Z95325.2(Xq21), AL022308(Xq21), AL590306.7(Xq22) AC010081.4(Yp11)
	CC-3	379	98	L1Hs group 5; AC019201.6(2q24) AC140125.3(5q35) AL354951.7(10q25)
	CC-4	379	99	L1Hs group2; AL133320(1p31), AC104454(1p22) AC092505.2(3p24), AC107028(3p11) AC098597(4p15), AC108516(4q22), AC093729(4q22), AC018680(4q31) AC109494(5p12), AC093255(5q21), AC008467(5q22), AC011413(5q31) AC004200(6p21), AL356601(6q15), AL356438(6q21), AP006196(6q21), AL583853(6q22) AC093775(7p11), AC092595(7q22) AC084257(8q21), AC020783(8q21), AC067844(8q23) AP004242(11q19), AP001876(11q21) AC005885(12q24) AC044907(15q25) AC009063(16q23), AC106814(16q23) AC005939(17q35) AP001028(18p11), AC002980(Xp22), AF148856(Xp22), AC004554(Xp22), AC004000(Xq24)
HEL-299	HE-1	378	98	L1Hs group 11; AC104456(1p22), AL451010(1p21)
	HE-2	379	97	AL596133(1q42)
	HE-3	376	94	AL161450(9p24)
	HE-4	382	100	AC092795.2(3p24)

Table 9 (Continued)

Cell line	Clone	Length	Similarity (%)	The most similar sequence
KYJ	YJ-1	379	100	AL160171.27(1p13)
	YJ-2	378	97	AC096740.3(4p16)
KBG	BG-1	379	98	L1Hs group 4; AC004514.1(8q21) AC040174.5(16q21), AC119049.1(16q21), AC040159.4(16q21)
	BG-2	378	96	L1Hs group 10; AC009319.19(3q27) AC090179.4(15q24), AC090751.9(15q24)
	BG-3	379	94	AL096710.8(6p11)
	BG-4	378	99	AC114967.2(5p13)
	BG-5	377	96	AL160171.27(1p13)
	BG-6	378	97	AL160171.27(1p13)
A-498	A4-1	377	98	AL121772.19(20p11)
	A4-2	379	96	L1Hs group 8; AC068206.10(15q25), AC020687.6(15q25)
	A4-3	380	98	AL096710.8(6p11)
	A4-4	379	100	AL157820.27(13q34)
	A4-5	379	97	AL096710.8(6p11)
A-549	A5-1	380	99	L1Hs group 12; AC079075.6(15q26), AC013787.15(15q26)
	A5-2	379	97	AC007000.2(7q11)
CEM-CM3	CE-1	383	95	AL139320.18(13q14)
	CE-2	378	93	AC079228.7(4q32)
HCT-15	HC-1	378	98	AC034229.4(5p15)
	HC-2	379	99	L1Hs group 3; AC094098.4(5q32), AC016572.6(5q35), AC022413.4(5q35) AL590095.5(10p15) AL117375.12(20p11), AL590285(Xp21), AL161778.19(Xq26)
	HC-3	379	99	L1Hs group 9; AC073326.6(7q31) AL078623.28(20p12)
	HC-4	375	98	L1Hs group 13; AC007992.12(8q22) AC015998.8(8q22), AP003465.2(8q22)
	HC-5	380	98	AL031176.8(Xq23)
HL-60	HL-1	399	98	AC125236.4(4p11)
Jurkat	JU-1	376	97	AC078816.16(3q24)
	JU-2	379	97	AL590131.8(Xq26)
	JU-3	378	100	AL049831.2(14q11)
	JU-4	385	99	AC004140.2(7q21)
	JU-5	378	99	AC093388.4(2q32)
MCF-7	MC-1	376	99	Z72001.3(Xq26)
	MC-2	380	98	AC131151.2(3q11)
	MC-3	379	99	AC131151.2(3q11)
	MC-4	379	99	AC104648.2(2q14)
MIA-PaCa-2	MP-1	380	99	AC106750.3(5q22)
	MP-2	439	96	AC093733.2(4q22)
	MP-3	379	99	AC025477.6(5q12)
	MP-4	379	99	AL356284.11(9p24)

Table 9 (Continued)

Cell line	Clone	Length	Similarity (%)	The most similar sequence
Molt-4	MO-1	379	97	AC087377.7(11p11)
	MO-2	379	98	AC144523.2(ND)
	MO-3	379	98	AC113385.2(5q21)
	MO-4	379	99	AL159154.16(13q22)
	MO-5	377	98	AL096710.8(6p11)
SK-N-SH	SN-1	379	99	AC022537.8(10q11)
	SN-2	379	99	AC022537.8(10q11)
	SN-3	379	98	AC022537.8(10q11)
	SN-4	379	99	AC004694.1(7p15)
	SN-5	379	99	AC004694.1(7p15)
SK-OV-3	SO-1	379	98	AC073196.3(4q12)
	SO-2	379	99	AC011155.13(18q21)
	SO-3	377	97	AC093783.2(4q28)
SNU-16	SU-1	379	98	L1Hs group 6; AC084264.7(2p22), AC011242.8(2p22) AL356794.10(6q24) AL160253.16(13q32)
	SU-2	379	96	AL139136.17(1q31)
	SU-3	385	91	AC079033.12(12q13)
	SU-4	379	97	L1Hs group 7; AC104623.4(2p24) AC091936.2(5q21) AL121852.3(14q12)
	SU-5	384	100	AC097515.5(4p16)
U-937	U9-1	379	100	AC010365.5(19q12)
	U9-2	379	100	AC090060.15(12q21)
	U9-3	379	98	M80340(ND)

* not determined the chromosomal location in Human Genome Resources in GenBank.

Table 10. L1Hs groups found in the similarity search

Group	No. of identical sequence	No. of location*	No. of chromosome**
L1Hs-1	65	51	20
L1Hs-2	34	38	14
L1Hs-3	7	6	4
L1Hs-4	4	2	2
L1Hs-5	3	3	3
L1Hs-6	4	3	3
L1Hs-7	3	3	3
L1Hs-8	2	1	1
L1Hs-9	2	2	2
L1Hs-10	3	2	2
L1Hs-11	2	2	1
L1Hs-12	2	1	1
L1Hs-13	3	1	1

* is determined by BLAST search in GenBank database.

** is determined by LOCUS search in Human Genome Resources in GenBank.

Table 11. L1Hs groups and the number of identical sequence clones from the cell lines tested

Cell line	No. of clones tested	No. of L1Hs group*													Total No. of clones found L1Hs group
		1	2	3	4	5	6	7	8	9	10	11	12	13	
KYJ	20	10	4	1		1		1				1			18
KBG	20	9	4				1								14
CCD-25Lu	20	9	3			1	1				1			1	16
HEL-299	20	8	6			2									16
A498	15	5	2		1	1					1				10
A549	15	5	3	2			1						1	1	13
CEM-CM3	15	4	1	2			2	1	1		1			1	13
HCT-15	15	3	1	1	1	1				1			1	1	10
HL-60	15	8	2	1		1		1					1		14
Jurkat	15	3	2		1	1			1	1		1			10
MCF-7	15	1	3	2		2	1		1		1				11
MIA-PaCa-2	15	4	1	1	1	1			1			1	1		11
Molt-4	15	5	1			1		1		1		1			10
SK-N-SH	15	4	3	1		2									10
SK-OV-3	15	5	4		1		1				1				12
SNU-16	15	3	3	1	2		1								10
U-937	15	6	1	1	1		1	1	1						12
Total	275	92	44	13	8	14	9	5	5	3	5	4	4	4	210

* group numbers are given in Table 10.

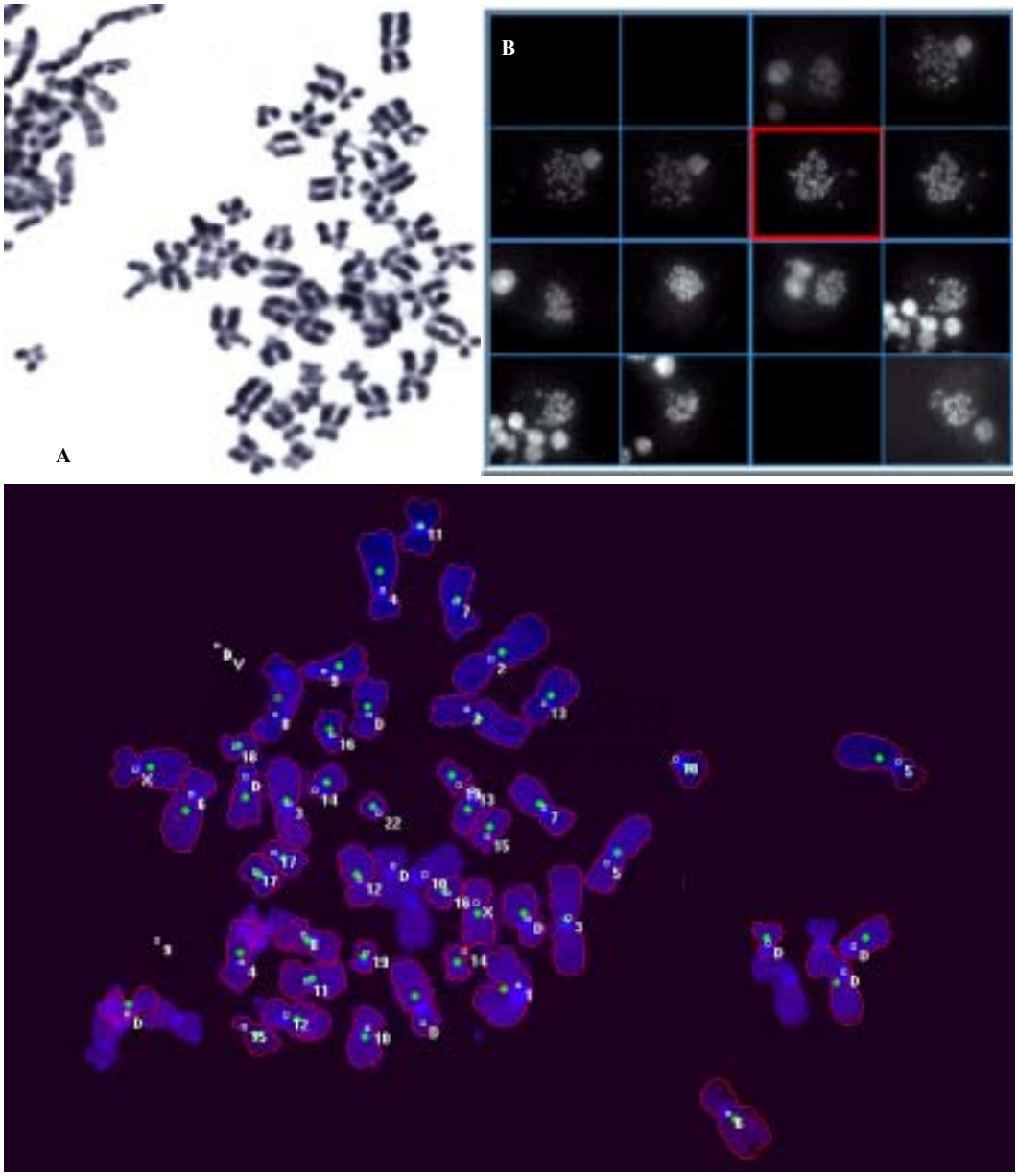


Figure 20. Chromosomal distribution of the L1Hs elements detected by cytogenetic analysis. A, the metaphase chromosomes stained with Giemsa; B, acquired FISH image panel from various metaphases; C, marked chromosomes using mFISH program (Imstar), open circles indicate the centromere of each chromosome, filled-green circles do midsts, and symbol 'D' does un-defined chromosome.

transcripts (transcribed from 1q42, 3p24, and Xq13) were specific for leukocytes. In addition, some elements were specific for each cells, for example, L1Hs elements from 10q11 and 7p15 was specific for neuroblastoma SK-N-SH cells, and those from 3q11 did breast cancer MCF-7 (Table 12). Transcripts from 6p11 were shared a normal leukocyte KBG and two cancer cell lines (A-498 and Molt-4). In addition, four L1Hs clones were located on the chromosome X; a clone from normal fibroblasts and three from cancer cells.

5) L1Hs p40 protein sequences may be involved in protein interactions in various cellular metabolism

Putative protein sequences of L1Hs p40 have converted from the DNA sequences. Fourteen from RNA transcripts and twenty-four from GenBank Accessions had active protein sequences. In addition, twelve L1Hs groups possessed the active protein sequences. The motif analysis has shown that the L1Hs p40 protein has two N-glycosylation sites, a cAMP- and cGMP-dependent protein kinase phosphorylation site, seven protein kinase C (PKC) phosphorylation sites, ten casein kinase II phosphorylation sites, an amidation site, a microbodies C-terminal targeting signal, a luecine zipper, and an N-myristoylation site (GKKQNRKTG). In addition, a helix-turn-helix DNA-binding motif (ALKELLKEALNMERNNQYQLLQ) have also found (Fig. 21). Multiple alignment of active L1Hs p40 protein sequences showed various missense mutations including in protein interaction positions (Appendix 3). In addition, nonsense mutations by nucleotide substitution have also been found in the transcripts from various cells (data not shown).

6) The sequences of transcripts and proteins of active L1Hs p40 divided into several clusters

Table 12. Cell line-specific transcripts of the L1Hs *p40* gene and its chromosomal locations

Chromosomal location	Symbol of clone*	No. of clone
1p13	YJ-1, BG-5, BG-6	3
1q31	SU-2	1
1q42	HE-2	1
2q14	MC-4	1
2q32	JU-5	1
3p24	HE-4	1
3q11	MC-2, MC-3	2
3q24	JU-1	1
4p16	YJ-2, SU-5	2
4p11	HL-1	1
4q12	SO-1	1
4q22	MP-2	1
4q28	SO-3	1
4q32	CE-2	1
5p15	HC-1	1
5p13	BG-4	1
5q12	MP-3	1
5q21	MO-3	1
5q22	MP-1	1
6p11	BG-3, A4-3, A4-5, MO-5	4
7p15	SN-4, SN-5	2
7q11	A5-2	1
7q21	JU-4	1
8q22	HC-4	1
9p24	HE-3, MP-4	2
10q11	SN-1, SN-2, SN-3	3
11p11	MO-1	1
12q13	SU-3	1
12q21	U9-2	1
13q14	CE-1	1
13q22	MO-4	1
13q34	A4-4	1
14q11	JU-3	1
15q25	A4-2	1
15q26	A5-1	1
18q21	SO-2	1
19q12	U9-1	1
20p11	A4-1	1
Xq13	CC-1	1
Xq23	HC-5	1
Xq26	JU-2, MC-1	2

* Symbol of clone is given in Table 9.

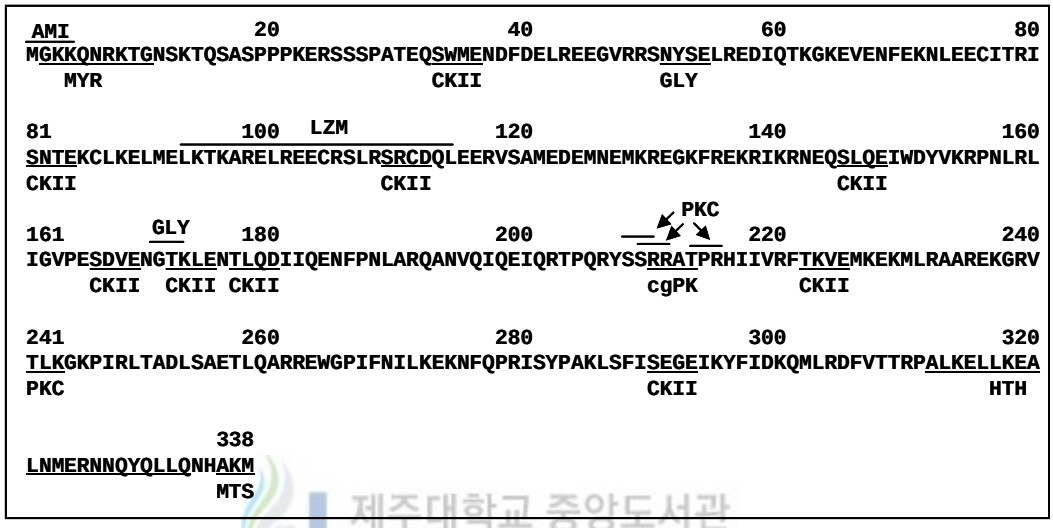


Figure 21. Protein motifs of the L1Hs p40 protein. AMI, amidation site; CKII, casein kinase II phosphorylation site; cgPK, cAMP-and cGMP dependent protein kinase phosphorylation site; GLY, N-glycosylation site; HTH, helix-turn-helix motif; LZM, leucine zipper; MTS, microbodies C-terminal targeting signal; MYR, N-myristoylation site, PKC, protein kinase C.

Sixty-two L1Hs DNA sequences identified from human cells were subjected to phylogenetic analysis with those previously reported as active forms of L1Hs (Sassaman *et al*., 1997) and each representative member of L1Hs groups (Appendix 2). The L1Hs elements formed three clusters on a consensus tree, constructed after 1,000 bootstrap replications of the gene sequences. Interestingly, six L1Hs groups were located in the cluster L1C, only L1Hs-11 was in the cluster L1B, and the other six groups were in the L1A (Fig. 22).

Phylogenetic tree based on L1Hs p40 protein sequences showed that three main branches (p40A, p40B, and p40C) appear and resemble those of L1Hs *p40* genes. However, p40A and p40C were human-specific, and p40B contained those of mice, gorilla, chimpanzees (Fig. 23). Active forms of L1Hs p40 have found at least a hundred and sixteen elements in the haploid human genome.

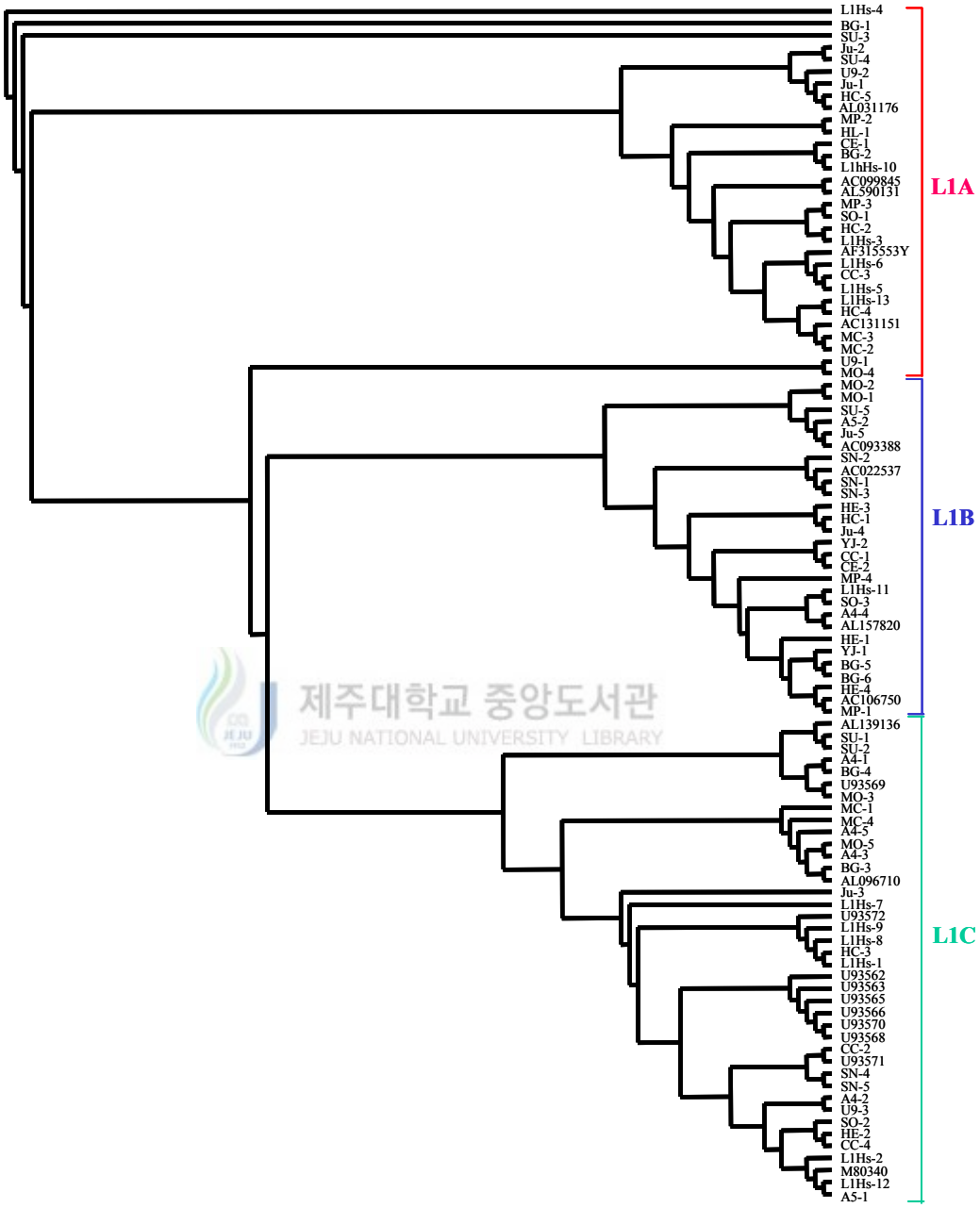


Figure 22. The strict consensus tree derived from the most parsimonious trees based on the L1Hs *p40* gene sequences obtained from various human cells and those reported previously in GenBank. The information of each clone is given in Table 9, and that of each group does in Table 11.

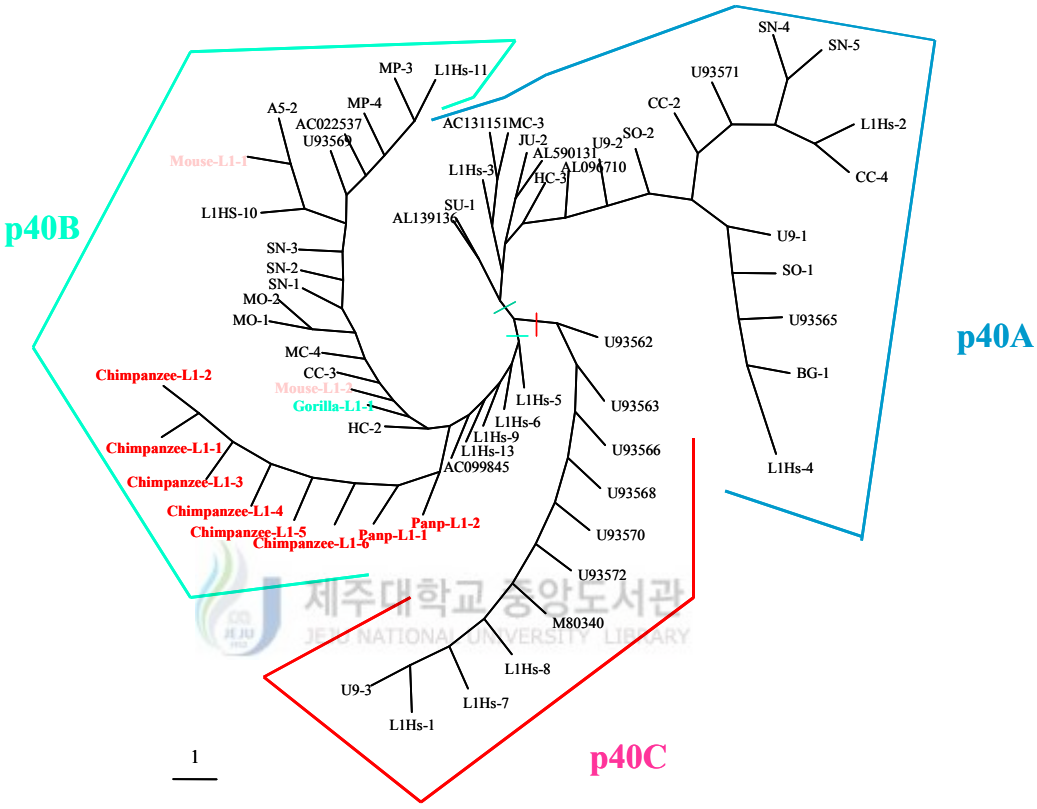


Figure 23. The strict consensus tree of active L1Hs p40 proteins from human cells and those converted the DNA sequences previously reported in GenBank. The information of each clone is given in Table 9, and those of groups do in Table 11. Gorg, *Gorilla gorilla*; Mouse, *Mus musculus*; Panp, *Pan paniscus*; Pant, *Pan troglodytes*.

Ⅳ. DISCUSSION

1. Transcription and Chromosomal Locations of the HERV-K

HERV-K genes were transcribed in both human cancer cells and normal cells, and the levels of transcription increased in fibroblasts treated with carcinogens. By RT-PCR analysis, RNA transcripts of HERV-K genes were observed as polyadenylated cellular transcripts, and the relative transcription levels varied in each cell line. HERV expression was found in the human cells including various cancers, placenta, and leukocytes of healthy individuals (Medstrand *et al*., 1992; Brodsky *et al*., 1993; Simon *et al*., 1994; Gotzinger *et al*., 1996; Yi *et al*., 2001). Cell type-specificity was also observed for K10 LTR, T47D LTR, K10 *env*, and *C4 env*. Cell type-specificity and tissue specific expression were reported previously (Majors, 1990; Ting *et al*., 1992; Medstrand and Blomberg, 1993; Kjellman *et al*., 1999; Sugimoto *et al*., 2001; Andersson *et al*., 2002). The cell-type dependent expression of retroviruses is due to specific regulatory sequences located in the U3 region, which contains a number of transcription-factor responsive sequences (Majors, 1990). For example, T47D LTR can induce expression of a reporter gene displaying distinct activity in different cell lines (Schön *et al*. 2001). The promoter analysis on HERV-K LTRs using SIGNALSCAN supported the results previously reported. Transcription factor Oct-1 and C/EBP play important roles in the cell-type specific expression. However, neither the Oct-1 binding site nor the CAAT-box was required for such expression of K10 LTR. Instead, the 3' terminus of U3 includes a TATA box sufficient for specificity (Baust *et al*., 2000; Schön *et al*., 2001; Vinogradova *et al*., 2001). The basic mechanery at the TATA box of K10 LTR may be determined the

cell-type specific expression (Lee *et al*., 2003). In addition, various mutations in K10 LTR sequences have also been observed in signal response elements (Appendix 1). Mutations in TATA box consequently can inhibit the transcription of genes. The other mutations in signal responsive elements can change the affinity of binding to transcription factors such as AP-1, c-MYC, GATA-1, Sp1, TGT3, and others. Since the 5'-LTR of HERV-K plays key roles in HERV-K transcription, this result indicates the sequence variations can also be responsible for the cell-type dependent patterns of HERV-K gene expressions and potential to retrotransposition.

HERV-K LTRs were expressed much more strongly than *env* genes, suggesting that there are many solitary LTRs distributed throughout the genome. In particular, the transcription patterns of *env* gene and LTRs were distinct in normal fibroblasts. Depil et al. (2002) suggested that HERV-K relative overexpression in leukemia cells might be specifically associated with tumor development, since there was no significant variation of HERV-K gene expression in normal peripheral blood mononuclear cells after exposure to different factors (PHA, gamma irradiation, 5-azacytidine). Since the CCD-25Lu and HEL-299 cell lines were derived from lung, the expression of HERV-K genes in these lines may be correlated with the stage of cellular development or the degree of cell differentiation. Reports suggested that HERVs expression can be involved in differentiation and morphogenesis of cells (Larsson *et al*., 1996; Herbst *et al*., 1999; Lin *et al*., 1999; 2000; Mi *et al*., 2000; Feng *et al*., 2001). Herbst *et al*. (1999) described the differential expression of HERV-K and an association of HERV-K in pathogenesis of germ cell tumors. HERV-R mRNA increased in cellular differentiation of monocyte U-937 by inducer molecules including TPA, vitamin-D1 and retinoic acid (Larsson *et al*., 1996), as a consequence of increased immune activity rather than causative of distinct diseases (Johnston *et al*., 2001). In addition, retinoic acid increased the expression of HERV-K in embryonal carcinoma (Caricasole *et al*., 2000) and in human

breast cancer (Feng *et al*., 2001). Sichangi *et al*. (2002) described the expression of HERV in male reproductive tissues speculating possible roles in spermatogenesis, sperm maturation, or tumor formation.

The expression of *C4 env* in these two cell lines was different, probably due to differences in the expression of *Complement 4* (*C4*). HERV-K(C4) is located in intron 9 of the *C4* gene in the reverse orientation to the *C4* coding sequence, and appears as a 6.4-kb insertion in 60% of human *C4* gene. The *C4* expression leads to production of an anti-sense RNA, which might protect against exogenous retroviral infections (Chu *et al*., 1995; Schneider *et al*., 2001).

DNA sequencing and subsequent BLAST search showed that K10 LTR was transcribed from all chromosomes, excluding chromosomes 21 and Y. This supports previous reports of the localization and distribution of entire HERV-K proviruses and its derived solitary LTRs in the human genome (Leib-Mösch *et al*., 1993; Seifarth *et al*., 1998; Zsiros *et al*., 1998; Kim *et al*., 2001; Sugimoto *et al*., 2001; Mayer and Meese, 2002). The expression and distribution of multiple HERV elements in various cancers (Feng et al., 2001; 2003; Yi et al., 2001). However, the K10 LTR elements on chromosome 21 and Y were inactive in all cell lines tested. Many K10 LTR elements had functionally active forms and were distributed on several chromosomes or also at fragile sites (Table 13), suggesting associations with chromosomal instability, cancers (Table 14), and genetic disorders (Table 15).

In this study, at least forty-nine HERV-K LTR elements from different loci on the chromosomes transcribed in the cells (Table 6). Except for two loci (8q11 and 10q24), the locations of transcriptionally active K10 LTR elements of normal cells were distinct from cancer cells. The cancer cell-specific and the normal cell-specific loci were observed. In addition, seven of cancer cell-specific loci shared different cancer cells, for example, eight K10 LTR elements transcribed from 1q21 in A-549, HL-60, Jurkat,

Table 13. Chromosomal location of detected retroelement transcripts and fragile sites

Locus	Retroelement detected		Fragile site	
	K10 LTR	L1Hs	Marker name	Class**
1p21		+*	FRA1E	A
1p22	+	++	FRA1D	A
1p36	+++		FRA1A	A
1q21	+++++++++		FRA1F	A
1q25		+	FRA1G	A
1q42		+	FRA1H	Z
1q44	+		FRA1I	A
2p11		+	FRA2L	A
2p13		+	FRA2E	A
2p22		++	FRA2K	F
2p24		+	FRA2C	A
2q31		++	FRA2G	A
2q33	+	+	FRA2I	A
3p24		++	FRA3A	A
3q27		+	FAR3C	A
4p15		+++	FRA4D	A
4p16	+	+	FRA4A	A
4q12		++	FRA4B	B1
4q31	++	+	FRA4C	A
5p13		+	FRA5A	B1
5p14	+		FRA5E	A
5q21		+++	FRA5F	A
5q31		+	FRA5D, B	A
5q35		+++	FRA5G	F
6p22		++	FRA6C	A
6p25		+	FRA6B	A
6q15		+	FRA6G	A
6q21		+++	FRA6F	A
7p11		+	FRA7A	F
7q21		++	FRA7E	A
7q22		+	FRA7F	A
7q31		+	FRA7G	A
7q32	++		FRA7H	A
8q22		+++	FRA8A, B	F, A
8q24	++++		FRA8C, D, E	A
10q11	+	+++	FRA10G	A
10q21	+		FRA10C	B1
10q25		+	FRA10B	B1

Table 13 (Continued)

Locus	Detection of retroelement		Fragile site	
	K10 LTR	L1Hs	Marker name	Class
11p15	+		FRA11C, I	A
11q13	++++	+	FRA11A, H	F, A
12q13	+	+	FRA12A	F
12q21		++	FRA12B	A
12q24		+	FRA12C	B2
13q32		+	FRA13D	A
14q23		++	FRA14B	A
16p12	++		FRA16E	D
16p13	+		FRA16A	F
16q23	+	++	FAR16D	A
18q21		+	FRA18B	A
19p13	+		FRA19B	F
20p11		++	FRA20A	F
20p12		+++	FRA20B	A
Xp22		+++	FRAXB	A
Xq22		+	FRAXC	A

* clone detected in LOCUSLINK search.

** is defined by Dhillon *et al*. (2003). Abbreviations: A1, common/aphidicolin; B1, common/BrdU; B2, rare/BrdU; D, common/distamycin-A; F, rare/folic acid; Z, common/ 5-azacytidine.

MIA-PaCa-2, SNU-16, and U-937 cells. 1q21 has a fragile site FRA1F (Table 13); and a chromosomal breakpoint in several cancers including chronic lymphocytic leukemia (CLL), bladder cancer (BC), and rectum cancer (RC) (Table 14). This finding presumes the association between the activation of provirus HERV-K and cancer development, as a transcriptional hotspot. Further studies on the association between them should be performed.

When K10 LTR elements were compared to the sequences of primates (chimpanzee, gorilla, macaques, and common squirrel monkey), three major clusters were made (Fig. 15). Human-specific subclusters and primates shared clusters have also found in the neighbor-joining tree. The time of divergence of the gorilla and human lineages, usually estimated to have occurred between 8 and 11 million years before present (YBP), and that of the chimpanzee and human lineages did about 6 million YBP (Ruvolo *et al*., 1991; Ovchinnikov *et al*., 2002). Therefore, this suggests that the proviruses had introduced into the common ancestors of primates at several times before and after speciation. HERV-K10 LTR elements could be classified into 14 classes (noted above as a 'K10 LTR group'), while Mold *et al*. (1997) reported four major classes based on HERV-K U3 region sequences.

2. Transcription and Chromosomal Locations of the L1Hs *p40*

RNA transcripts of the L1Hs elements have been found in human cells, including peripheral leukocytes, fibroblasts, and various human cancer cells, as cell type-independent (Minakami *et al*., 1992; Becker *et al*., 1993; Bratthauer and Fanning, 1993).

Table 14. Chromosomal location of retroelement transcripts, human cancer breakpoints and proto-oncogenes

Locus	Detection of retroelement		Cancer breakpoint	Proto-oncogene
	K10 LTR	L1Hs		
1p11		+[1]	ME	
1p13		++++	ChC*, NE	*NRAS*
1p21		+	OC, RC, RCC	
1p22	+	++	BC, CLL	
1p31		++	MA, RC	*RAB3B, VJUN*
1p36	+++		BC, CLL, ML, RC	*FGR, SRC, TNFR2*
1q21	+++++++++		BC, CLL, RC	
1q22	+			*SKI*
1q25		+	BC, ChC, PC, RC	*ARG, ABL2*
1q42		+	BC, NS, OC, PC	*RAB4*
2p11		+	BC, RC	
2p13		+	CLL	*REL*
2p22		++	CC, GL	
2p24		+	CLL, RC	*MYCN*
2q14	++	++		*LCO*
2q31		++	CLL	
2q33	+	+	CLL, RC, NS	
2q34		+	TLL	
3p11		+	RCC	
3p12		+	RCC	
3p24		++	CLL, LC, RCC	*RAB5, THRB*
3q21	++	+	OC, RCC	
3q24		+	AML, ME	
3q26	++		RCC	
3q27		+	BC, BLL, NS, OC	
4p11		+		*KIT*
4p15		+++	BC, OC	
4p16	+	+	CE	*RAF1P1*
4q12		++	GIST, TT	
4q31	++	+	CLL	
5p13		+	BC, NS, OC	
5p14	+		BC	*M1VI2*
5q21		+++	APC, BC, CC, EC, RC	*APC*
5q22		+++	APC, RCC	
5q31		+	BC, CLL	*AERB2, CTNNB1, TCTA*
5q32		+	GC, RCC	
5q35		+++	AML, BC	*CSF1R*

Table 14 (Continued)

Locus	Dectection of retroelement		Cancer breakpoint	Proto-oncogenes
	K10 LTR	L1Hs		
6p11		++++		*PKAS1*
6p21	+	+	CLL	*PIM1*, *TNAB*, *TNFA*
6p22		++	BC	
6q15		+	BC, ME	
6q21		+++	BC, ChC, CLL, US**	*FYN, MYB, ROS1, SYR*
6q25	+	+	LC	
7p11		+	BC	
7p15	+	+++		*MYCLK1*
7p22	+		CC, PC, RC	
7q11		++	PC	
7q21		++	BC	
7q22		+	BC, CLL	
7q31		+	BC, HCC, OC, RCC	*MET*
7q32	++		BC, CLL	*EPHT*
7q35		+		*BRAF1*
8q11	+++		BC, BR	*MOS*
8q22		+++	BC, OC, SCC	*MYBL1*
8q24	++++		BC, BL, OC, RCC	*BV1, MYC*
9p24		++		*OVC*
9q34	+++++++++		BC	*ABL1*
10q11	+	+++	MEN	*RET*
10q21	+			*PTC1*
10q24	++	+		*HOX11*
10q25		+	BC, PC	
11p11		+	PC	*SPI1*
11p15	+		BC, BR, LC	*HRAS, MTACR1*
11q12	+			
11q13	++++	+	BC, ChC, CLL, MEN, OC, OS	*BCL1, FGF3, FGF4, HSTF1, INT2, SEA*
11q14		+		*ARHD*
11q21		+	NE	
11q22	+		ACC, BR, LC	

Table 14 (Continued)

Locus	Dectection of retroelement		Cancer breakpoint	Proto-oncogenes
	K10 LTR	L1Hs		
12p11	++		RB	
12q12		+		*INT1*
12q13	+	+	CLL, BC, ChC, LP	*ERBB3, GLI, SAS, WNT*
12q14		++		*RAP1B*
12q21		++	BC	
12q24		+	BC, CLL, HA	
13q14	+++	+	BC, BR, RB	*BRCA2, RB1*
13q22		+	MY	
13q32		+	BC, CLL	
13q34		++	SCC	*RAP2A*
14q23		++	NS, SC	
15q25	+++	+++	BC	*FES*
15q26	+	++		*FPS*
16q23	+	++	CLL, EC, MY, RC, US	*MAF*
17q21	++		ChC, OC, RCC, SC, TLL	*BCL5, BRCA1, NEU, NEU*
17q25	+		SC	*ERBA2L*
18p11		+	APC	
18q11	+++		HCC	
18q21		+	BC, CLL, HCC, NS, RC	*BCL2, BCL3, SSAV1, YES*
19p13	+		BC, LL	*JUNB, JUND, LPSA, MEL, MELL1, RAB3A, VAV*
20p12		+++	BC	
20q12	+			*SRC*
22q11		+	BC, BR	*YESP*

Table 14 (Continued)

Locus	Dectection of retroelement		Cancer breakpoint	Proto-oncogenes
	K10 LTR	L1Hs		
Xp11		+	RCC, SS	*ARAF1*
Xp22		+++	US	*ELK1*
Xq11		+	PC	
Xq22		+	BC, CLL	*MYCL2*
Xq26	++	+++		*HRASP*

* childhood cancer not detailed.

** unspecified cancer.

Abbreviations: ACC, anal canal carcinoma; APC, adenomatous polyposis of the colon; AML, acute myeloid leukemia; BC, bladder cancer; BL, Burkitt's lymphoma; BLL, B lymphocytic leukemia; BR, Breast cancer; CC, colon cancer; CE, Cervical carcinoma; CLL, chronic lymphocytic leukemia; EC, esophageal cancer; GC, gastric cancer; GIST, gastrointestinal stromal tumor; GL glioma; HA, hepatic adenoma; HCC, hepatocellular carcinoma; LC, lung cancer; LL, lymphoid leukemia; LS, liposarcoma; MA, melanoma; ME, melanoma; MEN, multiple endocrine neoplasia; ML, myeloid cell leukemia; MTC, medullary thyroid carcinoma; MY, myeloma; NE, neuroblastoma; NS, non-small-cell lung cancer ; OC, ovary cancer; OS, oral squamous cell carcinoma; PC, prostate cancer; RB, retinoblastoma; RC, rectum cancer; RCC, renal cell carcinoma; SC, Schwannoma; SCC, squamous cell carcinoma; SS, synovial sarcoma; TLL, T lymphocytic leukemia; TT, testicular tumor.

In cellular RNA transcripts, various mutations led to the loss of function of retrotransposition. Several sequences had a missense mutation in the putative amino acid, or frame-shift mutations affecting their lengths. Length variation was usually caused by insertion/deletion mutations, although MP-2 had a 40-bp tandem duplication in the 5' region, and HE-1 and JU-2 had 5' additional sequences. Sequencing of L1Hs *p40* gene showed that 35 elements were the typical length. Sequence analysis provided information on the chromosomal locations of the active elements of the L1Hs. The putative active elements were determined by comparing the similarity, the typical lengths, and the putative protein translation of elements with previously reported active elements and the predicted proteins produced by L1Hs elements. Many L1Hs RNA transcripts appear to be transcriptionally active; twelve L1Hs groups and twenty-five solitary elements had potential activity for retrotransposition.

Protein motif analysis of L1Hs p40 using PROSCAN against PROSITE database showed various motifs for protein interactions (Fig. 20). L1Hs p40 has a luecine zipper motif that contributes to p40-p40 interactions and formation of multimeric complexes (Hohjoh and Singer, 1997). Various protein phosphorylation sites and modification sites had also been found in putative protein sequences. This predicts that activation of L1Hs p40 protein should be necessary to interact with cellular proteins and modify the protein p40 by cellular organelles. In addition, virus-related myristoylation site has been found, suggesting L1Hs p40 can be involved in maturation of the viral proteins and subsequent proteolytic processing (Towler *et al*., 1988; Grand, 1989). Until now, there are several hypotheses on L1Hs introduction via virus packaging. While, the origin of L1Hs has been predicted based on the comparison analysis of DNA sequences. There are no direct evidences on the origin of L1Hs from retrovirus. Therefore, this finding will support the retrovirus origin or the virus-related introduction (George *et al*. 1996; Malik *et al*. 1999). Moreover, this

means that L1Hs activation can be associated to viral pathogenesis in cells. Further study on their relationship will be needed.

Cell line-specific transcripts were observed in many chromosomal locations including three positions on chromosome X. Chromosome 4 had the most transcriptional active locations to transcribe L1Hs elements, but six chromosomes (15, 16, 17, 21, 22, and Y) had not the transcriptionally active locations. Several locations were specific for normal cells and for cancer cells, and shared among them. For instance, locations on 7p15 and 10q11 were specific for the neuroblastoma SK-N-SH cells, and those on 1p13 and 5p13 did for the leukocyte KBG, respectively (Table 12). In addition, mutations of RNA transcripts might lead to changes of L1Hs p40 protein sequences (Appendix 3). Various amino acid substitutions has been shown by the putative translation. Substitutions in protein interaction motifs has also confirmed, supposing the differences of protein activation. For example, since transcript clone MO-1 has two substitutions in PKC phosphorylation motifs, activation of MO-1 L1Hs p40 by PKC phosphorylation will be different efficiency. In addition, four substitutions in HTH motif will change the nucleic acid binding affinity, resulting the altered activity in L1Hs retrotransposition (Pietrokovski and Henikoff, 1997).

In the phylogenetic tree based on L1Hs *p40* DNA sequences, these were divided into three main clusters (L1A, L1B, and L1C); L1A was subdivided into two, and L1B into three subclusters (Fig. 22). Protein sequence phylogeny showed that three main branches (p40A, p40B, and p40C) appeared. Both p40A and p40C are human-specific, and p40B shared those of mice, gorilla, chimpanzees (Fig. 23). Of the two human-specific, p40A is more ancestral than others, suggesting L1Hs introduction into mammalian genome might be occurred several times before and after the speciation. Mammalian shared p40B suggested that speciation of mammalian predecessors might

have occurred after introduction of almost L1Hs elements. In addition, putative retrotransposable active L1Hs groups scattered over three branches, suggesting that ancestral LINE-1 elements possess the retrotransposition activity even in extracellular environment for a long time.

Up to date, it has been estimated that from thirty to sixty active L1Hs elements reside in the average diploid genome (Sassaman *et al*., 1997). However, this study suggested that at least a hundred and forty L1Hs elements in the haploid human genome have hypothetical activity potential for RNP complex formation, and they distributed over whole chromosomes. As showed in Figure 22, the active L1Hs elements divided into two clusters, but almost all of them are the members of cluster L1C. This result showed that L1C is the major active form of the L1Hs elements. The non-LTR superfamily is very ancient and LINE-1-like elements are ubiquitous in nature, those were found in plants, fungi, invertebrates, and various vertebrate classes from fish to mammals. L1 elements have been replicating and evolving in mammals for at least the past 100 million years (Furano, 2000). Also, human-specific or primate-specific sequences suggested that multiple introductions of LINE-1 elements, mediated a certain viral genome, had occurred during the mammalian evolution. This result will be contributed to explain the evolution and the speciation of mammals, while more extensive sampling and detailed approaches will be needed to elucidate the relationships between the molecular evolution of LINE-1 and mammals.

3. Transcription and Transposition of Retroelements by Carcinogens

This study showed the differential gene expressions of retroelements and the transposition activity by carcinogens. In the transposition analysis, the altered amplification patterns were also observed, resulting in transposition by carcinogen exposure.

HERV-K gene expression increased in fibroblast exposed to carcinogens along with changes in the expression of regulatory genes. However, the expression patterns of HERV-K *env* gene and LTRs increased, although this result was not directly associated with changes in the expression patterns of regulatory genes. Since the effects of carcinogens were different, response of the fibroblasts varied. For instance, in cells treated with TPA and DMBA, the expression of *BAX*, *c-MYC*, *IL-6* mRNA increased, while *p53* mRNA decreased. Although the apoptosis and proliferation pathways were activated, the cells actually proliferated. Levels of *BAX* and *IL-6* in EtBr-treated fibroblasts were more expressed than did non-treated cells, while *c-MYC* and *p53* expression decreased. However, the only HERV-K gene expressed at a lower level was K10 LTR, suggesting that either HERV-K expression is independent of apoptosis or that HERV-K RNA transcripts have longer life span than cellular mRNAs. The expression of active HERV *env* has been described in various states; enveloping retrovirus-like particles, a transcript of cellular RNA, and a protein of normal human tissues and cell lines (Kitamura *et al*., 1996; Löwer *et al*., 1993; Sauter *et al*., 1995; Berkhout *et al*., 1999). Induction of *IL-6* expression by carcinogens may lead to change of the immune activity and differentiation of lung fibroblast (Miyagi *et al*., 1991; Samad *et al*., 1994; Moodley *et al*., 2003). Cytokines and carcinogens induced the gene expression of HERVs, considering roles in inflammation and carcinogenesis (Larsson *et al*., 1996; Katsumata *et al*., 1999).

Under the carcinogen treatments, the expression of L1Hs was the highest in the DMBA treatment, and the lowest in the EtBr treatment. The relative transcription rate of

L1Hs was increased, as compared to the saline control, in all but the EtBr treatment. Since the patterns of relative transcription were different in the EtBr and MNNG treatments, although both EtBr and MNNG eventually induced the apoptosis of CCD-25Lu fibroblasts, the transcription of L1Hs elements is unlikely to be directly related to the induction of cell apoptosis. This means that EtBr and MNNG are contributed to different cellular pathways in CCD-25Lu cells, although the outcome was still apoptosis. In this process, a signal triggered by MNNG might increase the transcription of L1Hs, while EtBr has no effect on this.

The associations between location of repetitive DNA and chromosomal fragile sites were described. Transposon mariner-like elements (*Hsmar*) were found in human genome, suggesting relationship with fragile site and with genetic disorders (Reiter *et al*., 1999; Liehr *et al*., 2001). *Hsmar* elements are DNA transposons flanking ITR sequences. Chromosomal locations of *Hsmar* were known as a 'hotspot' to homologous recombination (Koyosawa and Chance, 1996; Reiter *et al*., 1996; Lupski, 1998). Analysis on transposition using ITR primers of this study showed that the transposition activity was stimulated by exposure to carcinogens, suggesting transposition mediated-genomic instability by carcinogens. Various carcinogens can affect to chromosomal stability. Retroelements and their remainders are involved in homologous recombination, causing chromosome rearrangement, and their expression can facilitate to chromosomal instability (Mager and Goodchild, 1989; Duesberg *et al*., 2000; Paige *et al*., 2000; Bardelli *et al*., 2001; Honma *et al*., 2002). Since the carcinogenic initiation begins of genomic alteration, chromosomal aberration and the errors in DNA mismatch repair causing to the activation of oncogenes or the inactivation of tumor suppressor genes, it is very important to explain the possible roles of carcinogenic effects on resulting genomic alteration in the cells. Certain stress condition, such as exposure to UV light (Rolfe and Banks, 1986), chemical mutagens (Bradshaw and McEntee,

1989; Morawetz and Hagen, 1990), low temperatures (Paquin and Williamson, 1984) or nitrogen starvation (Santos *et al*., 1997) can cause an increase in transposition activity (Staleva and Venkov, 2001). Staleva and Venkov (2001) has described the induction of transcription and subsequent retrotransposition of *Ty1* retrotransposon by mutagens (methyl methanesulfonate and 4-nitro quinolineoxide) in cultures and cells, and postulated the activation of *Ty1* element by DNA damage. Therefore, this finding suggests that the activation of retroelements can lead to transposition activity causing to genomic instability by carcinogens, and can also accelerate the carcinogenesis of the cells.

4. Association between Active Retrolements and Human Genetic Disorders

Active retroelements HERV-K10 LTR and L1Hs were found in the human cells. DNA sequencing and subsequent analyses were used to determine the chromosomal locations of transcripts. Interesting informations were obtained by comparing this result to Human Genome Resources in GenBank and those previously reported.

Many loci of retroelements L1Hs (forty-three) and HERV-K10 LTR (seventeen) were found at fragile sites. Moreover, at several fragile sites, these two retroelements were aligned in a repetitive manner (Table 13 and Appendix 4). Many proto-oncogenes and tumor suppressor genes mapped to these fragile sites, suggesting association with cancer*s* and transformation of cells. For example, tumor suppressor genes *FGR, SRC*, and *TNFR2* mapped to 1p36 (FRA1A) associated with BC, CLL, myeloid cell leukemia (ML), and RC. 5q21 (FRA5F) contained an oncogene *APC* and associated with adenomatous polyposis of the colon (APC), BC, colon cancer (CC), esophageal cancer

(EC), and RC. Proto-oncogenes *JUNB*, *JUND*, *LPSA*, *MEL*, *MELL1*, *RAB3A*, and *VAV* mapped to cancer breakpoint 18q21 (FRA18B) related to BC, CLL, hepatocellular carcinoma (HCC), non-small lung cancer (NS), and RC (Table 14).

Fragile sites may be unstable regions of the human genome, which might play an important role in the genetic instability associated with cancer predisposition (Sutherland *et al*., 1998; Ribas *et al*., 1999). The most fragile sites were associated with cancer breakpoints. For example, the most common fragile sites involved in bladder cancer were FRA1D (1p22), FRA1F (1q21), FRA8C (8q24), FRA9D (9q22), FRA9E (9q32), and FRA11C (11p15) (Moriarty and Webster, 2003). Dhillon *et al*. (2003) described that the expression of the fragile sites induced by aphidicolin was evaluated from ovarian cancer (OC), breast (BR), and NS patients belonging to different stages of cancer development. The frequency of expression of the fragile sites in cancer patients and their first-degree relatives was found to be statistically significant than those of the controls. In addition, the distribution of fragile sites in BR patients was changed according to different stages of cancer. Fundia *et al*. (1998) identified the spontaneous breakage and fragile site expression in CLL and suggested that the most of cancer breakpoints were involved in structural abnormalities associated with CLL. FRA16D (16q23) is one of the regions of the long arm of chromosome 16 and has been reported to show loss of heterozygosity (LOH) in various cancers, have identified a over 600 kb homozygous deletion in some cases (Paige *et al*., 2000; Tunca et a., 2000; Bednarek *et al*., 2001; Dhillon *et al*., 2003), and involved in multiple myeloma translocations (Krummel *et al*., 2000). In addition, chromosomal fragility may be involved in neuro-behavioral disorders (Gericke, 1999), mental retardation (Tükün *et al*., 2000), senescence and schizophrenia (Morelli *et al*., 2002). In this study, transcriptional locations of active retroelements have also been found in fragile sites associated with BC (Moriarty and Webster, 2003), leukemia (Sbrana and Musio, 1995; Fundia *et al*., 1998), CC (Tunca *et al*., 2000), giant

cell tumors (Nilsson *et al*., 2002), and BR, OC, and NS (Dhillon *et al*., 2003) (Table 13). Therefore, this suggests that the location and activation of retroelements HERV-K and L1Hs may be involved in cancer development and genetic disorders.

In addition, a transposition event can result in the deletion, inversion, or translocation of the genes (Youssoufian *et al*., 1988; Sheen *et al*., 2000; Gilbert *et al*., 2002; Sankoff *et al*., 2002; Chi *et al*., 2003). Morelli *et al*. (2002) reported that the nucleotide sequence of FRA6F is rich in repetitive elements; deletion breakpoints within FRA6F are common in several human leukemias and solid tumors, and are associated with hereditary schizophrenia. The locations of the L1Hs elements supported the findings of Morelli *et al*. (2002), but not the HERV-K10 locations, suggesting that inactive or other types of HERVs are located on FRA6F. Florl and Schulz (2003) described that most deletion breakpoints in cancers were located in or close to L1Hs retrotransposon clusters, and therefore, deletions of *CDKN2A* in 9q21 (FRA9C) may be facilitated by the presence of L1Hs clusters flanking the locus. A possible mechanism of recombination and transposition of transposable elements might be mediated chromosomal deletion and rearrangement (Lupski, 1998; Kiyosawa and Chance 1996; Reiter *et al*., 1996; Kolomietz *et al*., 2002; Toffolatti *et al*., 2002; Goryshin *et al*., 2003; Bosch and Jobling, 2003). Kolomietz *et al*. (2002) suggested Alu repeat clusters as mediators of chromosomal aberrations in tumors. Bosch and Jobling (2003) described that the duplication of the *AZFa* region of the human Y chromosome is mediated by homologous recombination between HERVs. Double transposition of *Tn5* produced the chromosomal deletion (Goryshin *et al*., 2003). Huang *et al*. (1998) found HERV-H and small polydispersed circular DNA (spcDNA) molecules in the midst of FRA7G and FRA3B, and suggested an association with both spcDNA molecules and hotspots for viral integration. HERV-K10 and L1Hs loci are much more common at fragile sites, regardless of whether they are expressed. Fragile sites are hotspots for genomic

instability causing chromosome rearrangement and recombination (Bobe *et al*., 1993; Fundia *et al*., 1998; Sutherland *et al*., 1998; De Mesa *et al*., 2000; Tükün *et al*., 2000; Ostertag and Kazazian, 2001; Nowak *et al*., 2002; Tighe *et al*., 2002). This study showed the cancer specific transcriptional locations of active retroelements (Appendix 4).

Reports described genomic instability by the activation of interspersed elements such as transposons, by amplification of their flanking inverted terminal repeat sequences, between tumor and normal tissues (Maeda *et al*., 1999; Ayyadevara *et al*., 2000; Luceri *et al*., 2000; Jotwani *et al*., 2001; Singh and Roy, 2001; Misra *et al*., 2001). Misra *et al*. (2001) described the alteration of a sequence with homology to HERV-K in primary human glioma, and suggested the viral repeat mediated genomic rearrangement in tumorigenesis. Inverted repeats are as source of genomic instability in yeast (Gordenin *et al*., 1993). As far as human disease is concerned, recombination in a defective low density lipoprotein (LDL) receptor gene in cases of familial hypercholesterolemia (Horsthemke *et al*., 1987). Genomic changes by activation of HERVs has been detected in some cancer and cultured cells (Huang *et al*., 1998; Misra *et al*., 2001; Goryshin *et al*., 2003), and hypomethylation of HERV-rich regions is also reported (Florl *et al*., 1999).

In the cells, repression of retrotransposition activity is very important. Methylation is an important epi-genetic mechanery in the repression of the repetitive transposable elements (Jurgens *et al*., 1996; Hata and Sakaki, 1997; Woodcock *et al*., 1997; Yoder *et al*., 1997; Florl *et al*., 1999). Florl *et al*. (1999) found the tissue specific hypermethylation of LINE-1 and HERV-K in various human cells and speculated that decreased methylation of retroelements may contribute to genomic instability in specific human tumors such as urothelial carcinoma by rendering these normally repressed sequences component for transcription and recombination. In this regard, hypermethylation of L1Hs has been suggested to inactivate the gene and chromosome. It has been proposed that L1Hs repeats may be mediators for the spread of

X chromosome inactivation; in normal somatic cells, X-linked L1Hs elements are hypermethylated on the inactive X (Lyon, 1998; Bailey *et al*., 2000; Riggs, 2002; Allen *et al*., 2003; Hansen, 2003). In addition, hypomethylation of L1Hs has been observed in hepatocellular carcinomas, but not in surrounding liver cirrhosis. Differential methylation has also been found in various cancer cells, compared to normal controls. Hypomethylation of the promoters of retroelements has led to increased expression of retroelements and probably to retrotransposition (Alves *et al*., 1996; Feng *et al*., 1996; Takai *et al*., 2000; Lin *et al*., 2001). In addition, DNA hypomethylation had also reduced genomic instability and can lead to elevated mutation rates (Chen *et al*., 1998; Rizwana and Hahn, 1999). In this study, transcripts of L1Hs on chromosome X were found in several cell lines (CCD-25Lu, HCT-15, Jurkat, and MCF-7). However, it was impossible to determine the origin of transcripts whether the those derived from inactive or not.

Recent reports described the retrotransposition activity of SINE, LINE-1 and HERVs, resulting in the generation of pseudogenes and multiple genes (Misra *et al*., 2001; Pavlicek *et al*., 2002; McNaughton *et al*., 1998; Cardazzo *et al*., 2003). Identical sequences and its variants had also been found in this study, which concluded that widespread distributions of the retroelements should be resulted from the retrotransposition. In general, non-LTR element L1Hs and LTR-elements HERV-K also utilize the 'replicative transposition' mechanism of integration, which enables purifying selection to retain active copies in a lineage (Malik *et al*., 1999). In genomic evolution, intronic expansion and processed pseudogenes might be made by transposition activity. The enlargement of the genome size and the decrease in genome compactness increasing the number and size of introns is a general pattern during the evolution of eukaryotes. Among the possible mechanisms for modifying intron size, insertion of transposable elements might have important roles in driving intron evolution (Cardazzo *et al*., 2003;

Salem *et al*., 2003).

The present study showed transcription and chromosomal location of active retroelements HERV-K and L1Hs in the human cells. These retroelements are multicopy and widespread on chromosomes, and also potentially active. Its activation of transposition and recombination events may be caused to human genetic disorders. Table 15 shows the human genetic disorders related to the transcriptionally active chromosomal locations of these retroelements. Gene expression, germ line and somatic mutations, and insertions of endogenous retroviruses have been found, suggesting a possible role for HERV-K elements in human diseases. Roles in insulin dependent diabetes mellitus (Conard *et al*., 1997), hypertension (Sirokman *et al*., 1997), cardiovascular disease (Bing *et al*., 1998), autoimmune disease (Krieg *et al*., 1992; Wu *et al*., 1993; Andersson *et al*., 1998), multiple sclerosis (Clausen, 2003), neoplastic disease (Suzuki *et al*., 1986; Bussemakers *et al*., 1992; Zsiros *et al*., 1998) and mouse embryonic lethal mutation in α_1 (I) collagen gene (Schnieke *et al*., 1983) have been postulated (Misra *et al*., 2001; Florl *et al*., 1999). Germ line mutations having the retrotransposition of L1Hs elements which abrogate the functions of important genes are known in several diseases, such as the coagulation factor VIII gene in hemophilia A (Kazazian *et al*., 1988), the *dystrophin* gene in Duchenne muscular dystrophy (Narita *et al*., 1993; McNaughton et a., 1998), the *fukutin* gene in Fukuyama-type congenital muscular dystrophy (Kobayashi *et al*., 1998) and the cytochrome *b*558 heavy-chain (*CYBB*) gene in chronic granulomatous disease (Kumatori *et al*, 1998; Meischl *et al*., 2000). In addition, somatic mutations caused by retrotransposition of an L1 element had been known in the *APC* gene in a colon cancer (Miki *et al*., 1992; Miki, 1998) and in the oncogene c-*MYC* in a breast cancer (Morse *et al*., 1988) and in a colon cancer (Naas *et al*., 1988).

Consequently, it is believed that the locations and expression of retroelements are

Table 15. Chromosomal location of detected retroelement transcripts and the map of human genetic disorders

Locus	Retroelement		Genetic disorders identified
	K10 LTR	L1Hs	
1p11		+	Adrenal hyperplasia II
1p13		++++	Hypothyroidism; Adrenal hyperplasia
1p21		+	Stargardt MD; Glycogen storage disease III
1p22	+	++	Zellweger S-2; Congenital disorder of glycosylation
1p31		++	*Maple-syrup urine disease type III*; Corneal dystrophy, gelatinous drop-like; *Multiple sclerosis*
1p35	++		Hypercholesterolemia, autosomal recessive
1p36	+++		Erythroblastosis fetalis
1q21	+++++++++		Gaucher's disease type 1
1q22	+		Urate oxidase deficiency
1q24		+	Deafness, nonsyndromic
1q25		+	Nephrotic syndrome, steroid-resistant, autosomal recessive; Short stature, pituitary and cerebellar defects
1q31		++	Cardiomyophathy familial 2; Branchiootic S2; RP 12
1q44	+		Fanconi's anemia; Familial cold autoinflammatory S; *Lupus erythematosus, systemic*
2p11		+	Smith-Magenis S; Methionine adenosyltransferase deficiency
2p13		+	Pakinson disease 3, autosomal dominant; *Lower motor neuron disease, progressive without sensory symptoms*
2p22		++	*Cáfe-au-lait spots, multiple*
2p24		+	NIDDM
2q24		+	Abetalipoproteinemia
2q31		++	Fibromuscular dystrophy of arteries; *Syndactyly, type II*
2q32		+	Winkle skin S
2q33	+	+	Gracile S; *IDDM*
2q34		+	Ehlers-Dalnos S, type X
2q35		+	Brachydactyly, type A1; Dermatitis, atopic
3p11		+	*Dementia, familial*
3p12		+	Spinocerebellar ataxia 7
3p24		++	Thyroid hormone unresponsiveness; Asthma; Marfan S; Moyamoya disease
3q11		++	*SCZD 1*
3q13		+	Hypoparathyroidism, Familial isolated; Hyperparathyroidism, neonatal severe primary
3q21	++	+	Alkaptonuria, Allergic rhinitis; *SCZD2*; Dermatitis atopic
3q22		+	*Night blindness, congenital stationary*
3q24		+	Thyroid hormone resistant, *Xenoderma pigmentosa*
3q26	++		Sucrose intolerance

Table 15 (Continued)

Locus	Retroelement		Genetic disorders identified
	K10 LTR	L1Hs	
4p15		+++	Phenylketonuria II
4p16	+	+	Huntington's disease; Ellis van Creveld S; *Hypochondroplasia*
4q12		++	Limb-girdle MD; Graves disease
4q13	+	++	Dentonogenesis imferpecta-1
4q22		+++	Abetalipoproteinemia
4q25		+	longevity 1; Mental retardation autosomal recessive
4q28		+	Afibrinogenemia, congenital
4q32		+++	Aspartylglucosaminuria
5p13		+	C6, C7, C9 deficiency; SCID
5q11	+		Mucopolyssacharidosis type IV; Polycystic ovary S; *SCZD 1*
5q12	+	+	Stroke (cerebrovascular accident); *Spinal muscular atrophy 1*
5q13		+	Sandhoff disease; Laron dwarfism
5q21		+++	*5q- S; Desmoid disease, hereditary*
5q22		+++	Corneal dystrophy; *MD, limb-girdle, autosomal dominant*
5q31		+	Cortisol resistance; *Asthma*; Corneal dystrophy; Osteoporosis; Graves disease; Carnitine deficiency;
5q32		+	Achondrogenesis type IB; Neonatal osseous dysplasia; IgE responsiveness; Dermatitis, atopic 6
6p11		++++	*Autoimmune disease* (*MHC related*)*
6p12	++		NIDDM
6p21	+	+	C2, *C4 deficiency*, Bare lymphocyte S; *IDDM*; *SCZD*; Polycystic kidney disease, autosomal recessive, *Autoimmune disease*; *Asthma*; Laterality defects, autosomal dominant
6p22		++	*Fanconi's anemia, complementation group E*; *Maple syrup urine disease 3, SCZD*
6q16	+		Chorioretinal atrophy, progressive; MD, retinal, California type I
6q21		+++	*Senescence; SCZD; IDDM*
6q22		+	Hepatic fibrosis; MD, congenital merosin deficient
6q24		+	Diabetes mellitus, transient neonatal
6q25	+	+	Pakinson disease, juvenile, autosomal recessive
7p11		+	Nystagmus, congenital, autosomal dominant
7p15	+	+++	Spinal muscular atrophy; Hyperinsulinism, autosomal dominant; RP 9

Table 15 (Continued)

Locus	Retroelement		Genetic disorders identified
	K10 LTR	L1Hs	
7q11		++	Williams S
7q21		++	Erythremia
7q22		+	*Chloride diarrhea, familial; Osteogenesis imperfecta, congenital*
7q31		+	Cystic fibrosis
7q32	++		RP10
7q35		+	Pancreatitis, hereditary; Glaucoma I
8p11	+		Werner S; Pfeiffer S; SCID1
8q11	+++		SCID1; RP1
8q21		+++	Cockane S; Aldesterone deficiency; Hyperaldosteronism; Giant cell hepatitis, *SCZD*
8q22		+++	Cohen S; *Hyperlipoproteinemia*
8q23		+	Hashimoto thyroiditis; *Autoimmune thyroid disease*
8q24	++++		Langer-Giedion S; Epilepsy; *Myeolcytomatosis*
9q31		+	*Fukuyama-type congenital MD*
9q34	+++++++++		C5 deficiency
10p15		+	*SCZD*
10q11	+	+++	Chronic infections
10q21	+		Moebius S; Multiple endocrine neoplasia
10q24	++	+	Congenital adrenal hyperplasia B
10q25		+	*IDDM*
11p11		+	Dysprothrombinemia
11p15	+		Niemann-Pick disease; Sickle-cell anemia
11q12	+		Atopy
11q13	++++	+	multiple endocrine neoplasia I
11q14		+	Systemic lupus erythematosus Papillon-Lefervre S
11q21		+	*SCZD*
11q22	+		Ataxia, episodic
12p11	++		Pakinson disease; Alzheimer disease type V
12q12		+	Rickets, vitamin D-resistant
12q13	+	+	Achondrogenesis type II; Stickler S
12q14		++	Tyrosinemia, type III
12q21		++	*MD*
12q24		+	Alcohol intolerance, acute; Fetal alcohol S, *SCZD*

Table 15 (Continued)

Locus	Retroelement		Genetic disorders identified
	K10 LTR	L1Hs	
13q14	+++	+	*Wilson's disease; Retinoblastoma*
13q22		+	Megacolon
13q32		+	*SCZD*
13q34		++	Factor Ⅶ, Ⅹ deficiency; *SCZD*
14q11	++	+	Ichthyosis; RP; Inflammatory Bowel disease
14q12		+	Cardiomyopathy, familial hypertrophic
14q23		++	Cataract, anterior polar
15q15		+	*MD, Limb-girdle, autosomal recessive*; *SCZD*
15q24		++	Mental retardation, severe, with spasticity; Cerebellar ataxia with mental retardation
15q25	+++	+++	Tay-Sachs disease
15q26	+	++	Bloom S
16p12	++		Batten disease
16p13	+		Cataract, congenital *Familial Mediterranean fever*
16q12		+	Inflammatory Bowel disease Ⅰ; Amyotrophic lateral sclerosis
16q21		+++	Hypodontia; Cataract lamellar
16q23	+	++	Pseudohyperkamelia
17q21	++		Glanzmann thromboasthemia; Hypertension essential
17q25	+		Rusell-Silver S
18p11		+	Glucocorticoid deficiency
18q11	+++		Amyloid neuropathy, familial
18q21		+	Cholestasis; *IDDM*
19p13	+		*IDDM*; *Hypercholesterolemia, familial*; *Cerebellar ataxia, Cayman type*
19q12	++	+	Nephrosis 1, congenital, Finnish type; Pseudoachodroplastic dysplasia; Migraine, familial; Alzheimer disease 2
20p12		+++	Huntington disease-like 1; Kindler S
20q12	+		NIDDM; Immunodeficiency with hyper-IgM, *SCZD*
22q11		+	*Cat-eye S*; *DeGeorge S*; *SCZD*
Xp11		+	Brunner S; *Night blindness, type I*
Xp21		+	*Becker MD*[4]; *MD*; *Retinitis pigmentosa*; *Cone dystrophy X-linked*; *Duchenne MD*
Xp22		+++	*Glycogen storage disease Ⅷ*
Xq11		+	Testicular feminization S
Xq13		+	Cleft plate; Turner S
Xq21		++	Agammaglobulinemia type 1; *Duchenne MD*; *Granulomatous disease, chronic*

Table 15 (Continued)

Locus	Retroelement		Genetic disorders identified
	K10 LTR	L1Hs	
Xq22		+	Agammaglobulinemia type 2; Alagille S; Amelogenesis imperfecta; *Alport S and associated diffuse leiomyomatosis*
Xq23		+	Hyperglycerolemia; Cerebellar ataxia 2; Arthrogrypopsis X-linked; Mental retardation X-linked
Xq24		+	Glycogen storage disease Ⅱb; Albinism-deafness S; Mental retardation X-linked
Xq26	++	+++	Albinism-deafness S; Immunodeficiency, X-linked, with hyper-IgM
Yp11		+	Gonadal dysgenesis; *Azoospermia factor 1, AZF1*

Abbreviations: IDDM, insulin-dependent diabetes mellitus; MD, muscular dystrophy; NIDDM, noninsulin-dependent diabetes mellitus, SCZD, schizophrenia; SCID, severe combined immunodeficiency; RP, Retinitis pigmentosa; S, syndrome.

Italic-lettered names are retroelement-related diseases previously reported.

closely related to the chromosomal instability responsible for genetic stability. This study showed that retroelements HERV-K and L1Hs are transcriptionally active in the human cells, as the cell-type specific transcription in HERV-K, and as the cell-type independent in L1Hs, respectively. Many retroelements are still transposing, suggesting the associations between the retrotransposition activity and genomic instability, cancer developments, genetic or epigenetic diseases. In addition, carcinogens may induce the transposition activity of transposable elements in the cells. In the activation of retroelements, various cellular factors may involve in transcription of retroelements and post-translational processes of their proteins. Since the molecular mechanisms that causes the hypomethylation and the activation of L1Hs, HERV-K provirus, and other repetitive sequences are virtually unknown, investigations of this mechanism will be very important. Also, further studies on biological function of retroelements and on the associations between the activation of retroelements and the human genetic disorders will be necessary.

LITERATURE CITED

Allen E, Horvath S, Tong F, Kraft P, Spiteri E, Riggs AD and Marahrens Y (2003) High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes. Proc Natl Acad Sci USA 100: 9940-9945.

Alves G, Tatro A and Fanning T (1996) Differential methylation of human LINE-1 retrotransposons in malignant cells. Gene 176: 39-44.

Andersson G, Svensson AC, Setterblad N and Rask L (1998) Retroelements in the human MHC class II region. Trends Genet 14: 109-114.

Andersson A, Venables PJW, Tönjes RR, Scherer J, Eriksso L and Larsson E (2002) Developmental expression of HERV-R (ERV3) and HERV-K in Human Tissue. Virology 297: 220-225.

Asch HL, Eliacin E, Fanning TG, Conolly JL, Bratthauer G and Asch BB (1996) Comparative expression of the LINE-1 p40 protein in human breast carcinomas and normal breast tissues. Oncol Res 8: 239-247.

Augé-Gouillou C, Hamelin MH and Demattei MV (2001) The wild-type conformation of the Mos-1 inverted terminal repeats in suboptional for transposition in bacteria. Mol Genet Genomics 265: 51-57.

Ayyadevara S, Thaden JJ and Shmookler Reis RJ (2000) Anchor polymerase chain reaction display: a high-throughput method to resolve, score, and isolate dimorphic genetic markers based on interspersed repetitive DNA elements. Anal Biochem 284: 19-28.

Bailey JA, Carrel L, Chakravarti A and Eichler EE (2000) Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proc Natl Acad Sci USA 97: 6634–6639.

Bairoch A, Bucher P and Hofmann K (1997) The PROSITE database, its status in 1997. Nucleic Acids Res 25: 217-221.

Bardelli A, Cahill DP, Lederer G, Speicher MR, Kinzler KW, Vogelstein B and Lengauer C (2001) Carcinogen-specific induction of genetic instability. Proc Natl Acad Sci USA 98: 5770-5775.

Baust C, Seifarth W, Germaier H, Hehlmann R and Leib-Mösch C (2000)

HERV-K-T47D-related long terminal repeats mediate polyadenylation of cellular transcripts. Genomics 66: 98-103.

Baust C, Seifarth W, Söchn U, Hehlmann R and Leib-Mösch C (2001) Functional activity of HERV-K-T47D-related long terminal repeats. Virology 283: 262-272.

Becker KG, Swergold GD, Ozato K and Thayer RE (1993) Binding of the ubiquitous nuclear transcription factor YY1 to a *cis* regulatory sequence in the human LINE-1 transposable element. Hum Mol Genet 2: 1697−1702.

Bednarek AK, Keck-Waggoner CL, Daniel RL, Bergsagel PL, Kiguchi K, Brenner AJ and Aldaz CM (2001) *WWOX*, the FRA16D gene, behaves as a suppressor of tumor growth. Cancer Res 61: 8068-8073.

Benit L, Lallemand JB, Casella JF, Philippe H and Heidmann T (1999) ERV-L elements: a family of endogenous retrovirus-like elements active throughout the evolution of mammals. J Virol 73: 3301-3308.

Berkhout B, Jebbink M and Zsiros J (1999) Identification of an active reverse transcriptase enzyme encoded by a human endogenous HERV-K retrovirus. J Virol 73: 2365-2375.

Bing OHL, Sirokman G and Humphries DE (1998) Hypothesis: link between endogenous retroviruses and cardiovascular disease. J Mol Cell Cardiol 30: 1257-1262.

Birren B., Green ED, Klapholz S, Myers RM and Roskams J (1997) *Genome analysis: A laboratory manual*. Cold Spring Harbor Laboratory Press, USA.

Bobe P, Benihoud K and Kiger N (1993) Allogenic MHC class II determinant(s) in MRL/Ipr autoimmune disease-prone mice. Unusual expression of an L1 transposable element creates molecular mimicry. J Immunol 151: 2813-2819.

Boller K, Kong H, Sauter M, Müeller-Lantzsch N, Löwer R, Löwer J and Kurth R (1993) Evidence of HERV-K is the endogenous retrovirus sequence that codes for the human teratocarcinoma-derived retrovirus HTDV. Virology 196: 349-353.

Bosch E and Jobling MA (2003) Duplications of the *AZFa* region of the human Y chromosome are mediated by homologous recombination between HERVs and are compatible with male fertility. Hum Mol Genet 12: 341-347.

Bradshaw V and McEntee K (1989) DNA damage activates transcription and transposition of yeast *Ty* retrotransposons. Mol Gen Genet **218:** 465−474.

Bratthauer GL, Cardiff RD and Fanning TG (1994) Expression of LINE-1

retrotransposons I human breast cancer. Cancer 73: 2333-2336.

Bratthauer GL and Fanning TG (1992) Active LINE-1 retrotransposons in human testicular cancer. Oncogene 7: 507-510.

Bratthauer GL and Fanning TG (1993) LINE-1 retrotransposon expression in pediatric germ cell tumors. Cancer 71: 2383-2386.

Brodsky I, Foley B, Haines D, Johnston J, Cubby K and Gillespie D (1993) Expression of HERV-K proviruses in human leukocytes. Blood 81: 2369-2374.

Brouha B, Meischl C, Ostertag E, De Boer M, Zhang Y, Neijens H, Roos D and Kazazian HHJr (2002) Evidence consistent with human L1 retrotransposition in maternal meiosis I. Am J Hum Genet 2002 71: 327-336.

Bussemakers MJG, Van Bokhoven A, Van Groningen JJM, Debruyne FMJ and Schalken JA (1992) Increased expression of retroviral sequences in progressionally advanced prostatic tumors. Biochem Biophys Res Comm 182: 318-324.

Cardazzo B, Bargelloni L, Toffolatti L and Patarnello T (2003) Intervening sequences in paralogous genes: a comparative genomic approach to study the evolution of X chromosome introns. Mol Biol Evol (in press).

Caricasole A, Ward-van Oostwaard D, Mummery C and van den Eijnden-van Raaij A (2000) Bone morphogenetic proteins and retinoic acid induce human endogenous retrovirus HERV-K expression in NT2D1 human embryonal carcinoma cells. Dev Growth Differ 42: 407-411.

Chen RZ, Pettersson U, Beard C, Jackson-Grusby L and Jaenisch R (1998) DNA hypomethylation leads to elevated mutation rates. Nature 395: 89-93.

Chi CS, Tsai CR, Chen LH, Lee HF, Mak BS, Yang SH, Wang TY, Shu SG and Chen CH (2003) Maple syrup urine disease in the Austronesian aboriginal tribe Paiwan of Taiwan: a novel DBT (E2) gene 4.7 kb founder deletion caused by a nonhomologous recombination between LINE-1 and Alu and the carrier-frequency determination. Eur J Hum Genet 11: 931-936.

Chu X, Rittner C and Schneider PM (1995) Length polymorphism of the human complement C4 gene is due to an ancient retroviral integration. Exp Clin Immunogenet 12: 74-81.

Clausen J (2003) Endogenous retroviruses and MS: using ERVs as disease markers. Int MS J 10: 22-28.

Clements AP and Singer MF (1998) The human LINE-1 reverse transcriptase: effect of

deletions outside the common reverse transcriptase domain. Nucleic Acid Res 26: 3528-3535.

Coffin JM, Hughes SH and Heidmann T (1997) *Retroviruses*, Cold Spring Harbor Laboratory Press, Cold spring Harbor, NY.

Conard B, Weissmahr RN, Böni J, Arcari R, Schüpbach J and Mach BA (1997) A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes. Cell 90: 303-313.

Cost GJ and Boeke JD (1998) Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. Biochemistry 37: 18081-18093.

Dematteі MV, Auge-Gouillou C, Pollet N, Hamelin MH, Meunier-Rotival M and Bigot Y (2000) Features of the mammal *mar1* transposons in the human, sheep, cow, and mouse genomes and implications for their evolution. Mamm Genome 11: 1111-1116.

De Mesa RL, Sierrasesúmaga L, Calasanz MJ, de Cerain AL and Patiño-García A (2000) Nonclonal chromosomal aberrations induced by anti-tumoral regimens in childhood cancer: relationship with cancer-related genes and fragile sites. Cancer Genet Cytogenet 121: 78-85.

Depil S, Roche C, Dussart P and Prin L (2002) Expression of a human endogenous retrovirus, HERV-K, in the blood cells of leukemia patients. Leukemia 16: 254-259.

Dhellin O, Maestre J and Heidmann T (1997) Functional differences between the human LINE retrotransposon and retroviral reverse transcriptase for *in vivo* mRNA reverse transcription. EMBO J 16: 6590-6602.

Dhillon VS, Huain SA and Ray GN (2003) Expression of aphidicolin-induced fragile sites and their relationship between genetic susceptibility in breast cancer, ovarian cancer, and non-small-cell lung cancer patients. Teratog Carcinog Mutagen Suppl 1: 35-45.

Dickson D, Eisenmann R, Fan H, Hunter E and Teich N (1982) Protein biosynthesis and assembly. In RNA Tumor Viruses, pp 135-146. Edited by Weiss, R. Reich N, Varmus H, and Coffin J, Cold spring Harbor, NY.

Dodd IB and Egan JB (1990) Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. Nucleic Acids Res 18: 5019-5026.

Domansky AN, KoDpantzev EP, Snezhkov EV, Lebedev YB, Leib-Mösch C and

Sverdlov ED (2000) Solitary HERV-K LTRs posses bi-directional promoter activity and contain a negative regulatory element in the U5 region. FEBS Lett 472: 191-195.

Dombroski BA, Feng Q, Mathias SL, Sassaman DM, Scott AF, Kazazian HHJr and Boeke JD (1994) An *in vivo* assay for the reverse transcriptase of human retrotransposon L1 in *Saccharomyces cereviciae*. Mol Cell Biol 14: 4485-449.

Drechsler M and Royer-Pokora B (1996) A LINE element is present at the site of a 300-kb deletion starting in intron 10 of the *PAX6* gene in a case of familial aniridia. Hum Genet 98: 297-303.

Duesberg P, Li R, Rasnick D, Rausch C, Willer A, Kraemer A, Yerganian G and Hehlmann R (2000) Aneuploidy precedes and segregates with chemical carcinogenesis. Cancer Genet Cytogenet 119: 83-93.

Fanning TG and Singer MF (1987) LINE-1: a mammalian transposable element. Biochem Biophys Acta 910: 203-212.

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package), ver. 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.

Feng Q, Moran JV, Kazazian HHJr and Boeke JD (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87: 905-916.

Feng WJ, Frost AF, Jian B, Azuero R, Lu DW, Chen D and Johanning GL (2003) Detecting the expression of Human endogenous retrovirus E envelope transcripts in human prostate adenocarcinoma. Cancer 98: 187-197.

Feng WJ, Frost AF, Johanning GL, Khazaeli MB, LoBuglio AF, Shaw DR and Strong TV (2001) Expression of human endogenous retrovirue K envelope transcripts in human breast cancer. Clin Cancer Res 7: 1553-1560.

Feuchter-Murthy AE, Freeman JD and Mager DL (1993) Splicing of a human endogenous retrovirus to a novel phospholipase A2 related gene. Nucleic Acids Res 21: 135-143.

Florl AR, Löwer R, Schmitz-Drager BJ and Schulz WA (1999) DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas. Br J Cancer 80: 1312-1321.

Florl AR and Schulz WA (2003) Peculiar structure and location of 9p21 homozygous deletion breakpoints in human cancer cells. Genes Chromosomes Cancer 37: 141-148.

Fundia A, Giere I, Larripa I and Slavutsky I (1998) Spontaneous breakage and fragile site expression in chronic lymphocytic leukemia. Cancer Genet Cytogenet 103: 144-148.

Furano AV (2000) The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. Prog Nucleic Acid Res Mol Biol 64: 255-294.

George JA, Burke WD and Eickbush TH (1996) Analysis of the 5' junctions of R2 insertions with the 28S gene: implications for non-LTR retrotransposition. Genetics 142: 853-863.

Gericke GS (1999) Chromosomal fragility may be indicative of altered higher-order DNA organization as the underlying genetic diathesis in complex neurobehavioural disorders. Med Hypotheses 52: 210-208.

Gilbert N, Lutz-Prigge S and Moran JV (2002) Genomic deletions created upon LINE-1 retrotransposition. Cell 110: 315−325.

Gordenin DA, Lobachev KS, Degtyareva NP, Malkova AL, Perkins E and Resnik MA (1993) Inverted DNA repeats: a source of eukaryotic genomic instability. Mol Cell Biol 13: 5315-5322.

Goryshin IY, Naumann TA, Apodaca J and Reznikoff WS (2003) Chromosomal deletion formation system based on Tn5 double transposition: use for making minimal genomes and essential gene analysis. Genome Res 13: 644-653.

Gotzinger N, Sauter M, Roemer K and N Müeller-Lantzsch (1996) Regulation of human endogenous retrovirus-K Gag expression in teratocarcinoma cell lines and human tumors. J Gen Virol 77: 2983-2990.

Grand RJ (1989) Acylation of viral and eukaryotic proteins. Biochem J 258: 625-638.

Gunning P, Ponte P, Okayama H, Engel J, Blau H and Kedes L (1983) Isolation and characterization of full-length cDNA clones for human α-, β-, and gamma-actin mRNAs: skeletal but not cytoplasmic actins have an amino-terminal cysteine that is subsequently removed. Mol Cell Biol 3: 787-795.

Hansen RS (2003) X inactivation-specific methylation of LINE-1 elements by DNMT3B: implications for the Lyon repeat hypothesis. Hum Mol Genet 12: 2559-2567.

Hata K and Sakaki Y (1997) Identification of critical CpG sites for repression of

L1 transcription by DNA methylation. Gene 189: 227-234.

Herbst H, Sauter M and Müeller-Lantzsch N (1996) Expression of human endogenous retrovirus K elements in germ cell and trophoblastic tumors. Am J Pathol 149: 1727-1735.

Hohjoh H and Singer MF (1997) Sequence-specific single-strand RNA binding protein encoded by the human LINE-1 retrotransposon. EMBO J 16: 6034-6043.

Holmes SE, Dombroski BA, Krebs CM, Boehm CD and Kazazian HHJr (1994) A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. Nat Genet 7: 143-148.

Honma M, Tadokoro S, Sakamoto H, Tanabe H, Sugimoto M, Furuichi Y, Satoh T, Sofuni T, Goto M and Hayashi M (2002) Chromosomal instability in B-lymphoblastoid cell lines from Werner and Bloom syndrome patients. Mutat Res 520: 15-24.

Horsthemke B, Beisiegel U, Dunning A, Havinga JR, Williamson R and Humphries S (1987) Unequal crossing-over between two alu-rpetitive DNA sequences in the low-density-lipoprotein-receptor gene; a possible mechanism for the defect in a patient with familial hypercholesterolemia, Eur J Biochem 164: 77-81.

Huang H, Qian J, Proffit J, Wilber K, Jenkins R and Smith DI (1998) FRA7G extends over a broad region: coincidence of human endogenous retroviral sequences (HERV-H) and small polydispersed circular DNAs (spcDNA) and fragile sites. Oncogene 16: 2311-2319.

Hugh MR (1996) Members of the pogo superfamily of DNA-mediated transposons in the human genome. Mol Gen Genet 252: 761-766.

Hutchison CA, Hardies SC, Loeb DD, Shehee WR and Edgell MH (1989) LINES and related retroposons: long interspersed sequences in the eukaryotic genome. In *Mobile DNA*, Berg DE and Howe MM eds, pp. 593-617. ASM press, Washington, DC.

Hwu HR, Roberts JW, Davidson EH and Britten RJ (1986) Insertion and/or deletion of many repeated DNA sequences in human and higher ape evolution. Proc Natl Acad Sci USA 83: 3875-3879.

Johnston JB, Silva C, Holden J, Warren KG, Clark AW and Power C (2001) Monocyte activation and differentiation augment human endogenous retrovirus expression: implications for inflammatory brain diseases. Ann Neurol 50: 434-442.

Jotwani G, Misra A, Chattopadhyay P, Sarkar C, Mahapatra AK and Sinha S (2001) Genetic heterogeneity and alterations in chromosome 9 loci in a localized region of a functional pituitary adenoma. Cancer Genet Cytogenet 125: 41-45.

Jurgens B, Schmitz-Drager BJ and Schulz WA (1996) Hypomethylation of L1 LINE sequences prevailing in human urothelial carcinoma. Cancer Res 56: 5698-5703.

Kalter SS, Helmke RJ, Heberling RL, Panigel M, Fowler AK, Strickland JE and Hellman A (1973) C-type particles in normal human placentas. J Natl Cancer Inst 50: 1081-1084.

Kapitonov VV and Jurka J (1999) The long terminal repeat of an endogenous retrovirus induces alternative splicing and encodes and additional carboxyl-terminal sequences in the human leptin receptor. J Mol Evol 48: 248-251.

Katsumata K, Ikeda H, Sato M, Ishizu A, Kawarada Y, Kato H, Wakisaka A, Koike T and Yoshiki T (1999) Cytokine regulation of env gene expression of human endogenous retrovirus-R in human endothelial cells. Clin Immunol 93: 75-80.

Kazazian HHJr (2000) Genetics. L1 retrotransposons shape the mammalian genome. Science 289: 1152-1153.

Kazazian HHJr and Goodier JL (2002) LINE drive. retrotransposition and genome instability. Cell 110: 277-280.

Kazazian HHJr, Wong C, Youssoufian H, Scott AF, Phillips DG and Antonarakis SE (1988) Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332: 164-166.

Kimberland ML, Divoky V, Prchal J, Schwahn U, Berger W and Kazazian HHJr (1999) Full-length human L1 insertions retain the capacity for high frequency retrotransposition in cultured cells. Hum Mol Genet 8: 1557-1560.

Kim HS, Yi JM and Jeon SH (2001) Isolation and phylogenic analysis of HERV-K long terminal repeat cDNA in cancer cell. AIDS Res Hum Retroviruses 17: 987-990.

Kimura M (1980) The neutral theory of molecular evolution, Cambridge University press, Cambridge.

Kitamura Y, Ayukawa T, Ishikawa T, Kanda T and Yoiiske K (1996) Human endogenous retrovirus K10 encodes a functional integrase. J Virol 70: 3302-3306.

Kiyosawa H and Chance PF (1996) Primate origin of the CMT1A-REP repeat and

analysis of a putative transposon-associated recombination hotspot. Hum Mol Genet 5: 745-753.

Kjellman C, Sjgren H-O, Salford LG and Widergen B (1999) HERV-F(XA34) is a full-length human endogenous retrovirus expressed in placental and fetal tissues. Gene 239: 99-107.

Knössl M, Löwer R and Löwer J (1999) Expression of the human endogenous retrovirus HTDV/HERV-K is enhanced by cellular transcription factor YY1. J Virol 73: 1254-1261.

Kobayashi K, Nakahori Y, Miyake M, Matsumura K, Kondo-Iida E, and Nomura Y, Segawa M, Yoshioka M, Saito K, Osawa M, Hamano K, Sakakihara Y, Nonaka I, Nakagome Y, Kanazawa I, Nakamura Y, Tokunaga K and Toda T (1998) An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy. Nature 394: 388-392.

Kolomietz E, Meyn MS, Pandita A and Squire JA (2002) The role of Alu clusters as mediators of recurrent chromosomal aberrations in tumors. Genes Chromosomes Cancer 35: 97-112.

Kowalski PE, Freeman JD and Mager DL (1999) Intergenic splicing between a HERV-H endogenous retrovirus and two adjacent human genes. Genomics 57: 371-379.

Krieg AM, Gourley MF and Perl A (1992) Endogenous retroviruses: potential etiologic agents in autoimmunity. FASEB J 6: 2537-2544.

Krummel KA, Roberts LR, Kawakami M, Glover TW and Smith DI (2000) The characterization of the common fragile site FRA16D and its involvement in multiple myeloma translocations. Genomics 69: 37-46.

Kumatori A, Faizunnessa NN, Suzuki S, Moriuchi T, Kurozumi H and Nakamura M (1998) Nonhomologous recombination between the cytochrome b_{558} heavy chain gene (*CYBB*) and LINE-1 causes an X-linked chronic granulomatous disease. Genomics 53: 123-128.

Kuo KW, Sheu HM, Huang YS and Leung WC (1998) Expression of transposon LINE-1 is relatively human specific and function of the transcripts may be proliferation-essential. Biochem Biophys Res Comm 253: 566-570.

Larsson E, Venables PJ, Andersson AC, Fan W, Rigby S, Botling J, berg F, Cohen M and Nilsson K (1996) Expression of the endogenous retrovirus ERV3 (HERV-R) during induced monocyte differentiation in the U-937 cell line. Int J Cancer 67: 451-456.

Lee WJ, Kim HJ and Jang KL (2003) Analysis of transcriptional regulatory sequences in the human endogenous retrovirus W long terminal repeat. J Gen Virol 84: 2229-2235.

Leib-Mösch C, Halmeier M, Werner Y, Geigl EM, Brack-Werner R, Francke U, Erfle V and Hehlmann R (1993) Genomic distribution and transcription of solitary HERV-K LTRs. Genomics 18: 261-269.

Liehr T, Reiter LT, Lupski JR, Murakami T, Claussen U and Rautenstrauss B (2001) Regional localization of 10 mariner transposon-like ESTs by means of FISH-evidence for a correlation with fragile sites. Mamm Genome 12: 326-328.

Lin CH, Hsieh SY, Sheen IS, Lee WC, Chen TC, Shyu WC and Liaw YF (2001) Genome-wide hypomethylation in hepatocellular carcinogenesis. Cancer Res 61: 4238-4243.

Lin L, Xu B and Rote NS (1999) Expression of endogenous retrovirus ERV-3 induces differentiation in BeWo, a choriocarcinoma model of human placenta trophoblast. Placenta 20: 109-118.

Lin L, Xu B and Rote NS (2000) The cellular mechanism by which the human endogenous retrovirus ERV-3 env gene affects proliferation and differentiation in human placental trophoblast model, BeWo. Placenta 21: 73-78.

Löwer R, Boller K, Hasenmaier B, Korbmacher C, Müeller-Lantzsch N, Löwer J and Kurth R (1993) Identification of human endogenous retroviruses with complex mRNA expression and particle formation. Proc Natl Acad Sci USA 90: 4480-4484.

Löwer R, Löwer J and Kurth R (1996) The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. Proc Natl Acad Sci USA 93: 5177-5184.

Luan DD, Korman MH, Jakubczak JL and Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72: 595-605.

Luceri C, De Filippo C, Caderni G, Gambacciani L, Salvadori M, Giannini A and Dolara P (2000) Detection of somatic DNA alterations in azoxymethane-induced F344 rat colon tumors by random amplified polymorphic DA analysis. Carcinogenesis 21: 1753-1756.

Lupski JR (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. Trends Genet 14: 417-422.

Lyon MF (1998) X-chromosome inactivation: a repeat hypothesis. Cytogenet Cell

Genet 80: 133−137.

Maeda T, Jikko A, Hiranuma H and Fuchihata H (1999) Analysis of genomic instability in squamous cell carcinoma of the head and neck using random amplified polymorphic DNA method. Cancer Lett 138: 183-188.

Mager DL and Goodchild NL (1989) Homologous recombination between the LTRs of a human retrovirus-like element causes a 5-kb deletion in two siblings. Am J Hum Genet 45: 848-854.

Mager DL, Hunter DG, Schertzer M and Freeman JD (1999) Endogenous retroviruses provide the primary polyadenylation signal for two new human genes (*HHLA2* and *HHLA3*). Genomics 59: 255-263.

Majors J (1990) The structure and function of retroviral long terminal repeats. Curr Top Microbiol Immunol 157: 49-92.

Malik HS, Burke WD and Eickbush TH (1999) The age and evolution of non-LTR retrotransposable elements. Mol Biol Evol 16: 793-805.

Martin MA, Bryan T, McCutchan TF and Chan HW (1981) Detection and cloning of murine leukemia virus-related sequences from African green monkey liver DNA. J Virol 39: 835-844.

Mayer J and Meese EU (2002) The human endogenous retrovirus family HERV-K (HML-3). Genomics 80: 331-343.

Mayer J, Meese EU and Müeller-Lantzsch N (1998) Human endogenous retrovirus K homologous sequences and their coding capacity in old world primates. J Virol 72: 1870-1875.

Mayer J, Meese EU and Müeller-Lantzsch N (1997) Multiple human endogenous retrovirus (HERV-K) loci with gag open reading frames in the human genome. Cytogenet Cell Genet 78: 1-5.

McNaughton JC, Cockburn DJ, Hughes G, Jones WA, Laing NG, Ray PN, Stockwell PA and Petersen GB (1998) Is gene deletion in eukaryotes sequence-dependent? A study of nine deletion junctions and nineteen other deletion breakpoints in intron 7 of the human dystrophin gene. Gene 222: 41-51.

Medstrand P and Blomberg J (1993) Characterization of novel reverse transcriptase encoding human endogenous retroviral sequences similar to type A and type B retroviruses: Differential transcription in normal human tissues. J Virol 67: 6778-6787.

Medstrand P, Lindeskog M and Blomberg J (1992) Expression of human endogenous retroviral sequences in peripheral blood mononuclear cells of healthy individuals. J Gen Virol 73: 2463-2466.

Meischl C, Boer M, Ahlin A and Roos D (2000) A new exon created by intronic insertion of a rearranged LINE-1 element as the cause of chronic granulomatous disease. Eur J Hum Genet 8: 697-703.

Miki Y, Nishisho I, Horii A, Miyoshi Y, Utsunomiya J, Kinzler KW, Vogelstein B and Nakamura Y (1992) Disruption of the *APC* gene by a retrotransposal insertion of L1 sequence in a colon cancer. Cancer Res 52: 643-645.

Miki Y (1998) Retrotransposal integration of mobile genetic elements in human diseases. J Hum Genet 43: 77-84.

Minakami R, Kurose K, Etoh K, Furuhata Y, Hattori M and Sakaki Y (1992) Identification of an internal *cis*-element essential for the human L1 transcription and a nuclear factor(s) binding to the element. Nucleic Acids Res 20: 3139-3145.

Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, Lavallie L, Tang XY, Edouard P, Howes S, Keith JCJr and McCoy JM (2000) Syncytin as a captive retroviral envelope protein involved in human placental morphogenesis. Nature 403: 785-789.

Misra A, Chosdol K, Sarkar C, Mahapatra AK and Sinha S (2001) Alteration of a sequence with homology to human endogenous retrovirus (HERV-K) in primary human glioma: implications for viral repeat mediated rearrangement. Mutat Res 484: 53-59.

Miyagi T, Akachi M, Yamato K, Miyoshi I and Koeffler HP (1991) D-factor: modulation of expression in fibroblasts. Leuk Res 15: 441-451.

Mold DE, Wu TC, Askin F and Huang RC (1997) Four classes of HERV-K long terminal repeats and their relative promoter strengths for transcription. Biomed Sci 4: 78-82.

Moodley YP, Misso NL, Scaffidi AK, Fogel-Petrovic M, McAnulty RJ, Laurent GJ, Thompson PJ and Knight DA (2003) Inverse effects of interleukin-6 on apoptosis of fibroblasts from pulmonary fibrosis and normal lungs. Am J Respir Cell Mol Biol 29: 490-498.

Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD and Kazazian HHJr (1996) High frequency retrotransposition in cultured mammalian cells. Cell 87: 917-927.

Morawetz I and Hagen U (1990) Effect of irradiation and mutagenic chemicals on the generation of *ADH*2- and *ADH*4-constitutive mutants in yeast: the inducibility of Ty transposition by UV and ethyl methanesulfonate. Mutat Res 229: 69−77.

Morelli C, Karayianni E, Magnanini C, Mungall AJ, Thorland E, Negrini M, Smith DI and Barbant-Brodano G (2002) Cloning and characterization of the common fragile site FRA6F harboring a replicative senescence gene and frequently deleted in human tumors. Oncogene 21: 7266-7276.

Moriarty HT and Webster LR (2003) Fragile sites and bladder cancer. Cancer Genet Cytogenet 140: 89-98.

Morse B, Rotherg PG, South VJ, Spandorfer JM and Astrin SM (1988) Insertional mutagenesis of the *myc* locus by a LINE-1 sequence in a human breast carcinoma. Nature 333: 87-90.

Müeller-Lantzsch N, Sauter M, Weiskircher A, Weiskircher K, Best B, Buck M and Graesser F (1993) Human endogenous retroviral element K10 (HERV-K10) encodes a full-length gag homologous 73-kDa protein and a functional protease. AIDS Res Hum Retroviruses 9: 343-350.

Naas TP, DeBerardinis JR, Moran JV, Ostertag EM, Kingsmore SF, Seldin MF, Hayashizaki Y, Martin SL and Kazazian HHJr (1988) Disruption of the *APC* gene by a retrotransposal insertion of L1 sequence in a colon cancer. EMBO J 17: 590−597

Narita N, Nishio H, Kitoh Y, Ishikawa Y, Minami R, Nakamura H and Matsuo M (1993) Insertion of a 5' truncated L1 element into the 3' end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. J Clin Invest 91: 1862-1867.

Nilsson M, Hoglund M, Panagopoulos I, Sciot R, Dal Cin P, Debiec-Rychter M, Mertens F and Mandahl N (2002) Molecular cytogenetic mapping of recurrent chromosomal breakpoints in tenosynovial giant cell tumors. Virchows Arch 441: 475-480.

Nowak MA, Komarova NL, Sengupta A, Jallepalli PV, Shih leM and Vogelstein B (2002) The role of chromosomal instability of tumor initiation. Proc Natl Acad Sci USA 99: 16226-16231.

Ono M, Yaunaga Y, Miyata T and Ushikobo H (1986) Nucleotide sequence of human endogenous retrovirus genome related to the mouse mammary tumor virus genome. J Virol 60: 589-598.

Ostertag EM and Kazazian HHJr (2001) Biology of mammalian L1 retrotransposons. Annu Rev Genet 35: 501-538.

Ovchinnikov I, Rubin A and Swergold GD (2002) Tracing the LINEs of human evolution. Proc Natl Acad Sci USA 99: 10522-10527.

Ovchinnikov I, Troxel AB and Swergold GD (2001) Genomic characterization of recent human LINE-1 insertions: evidence supporting random insertion. Genome Res 11: 2050-2058.

Paige AJ, Taylor KJ, Stewart A, Sgouros JG, Gabra H, Sellar GC, Smyth JF, Porteous DJ and Watson JE (2000) A 700-kb physical map of a region of 16q23.2 homozygously deleted in multiple cancers and spanning the common fragile site FRA16D. Cancer Res 60: 1690-1697.

Paquin CE and Williamson VM (1984) Temperature effects on the rate of Ty transposition. Science 226: 53−55.

Patience C, Simpson GR, Colletta AA, Welch HM, Weiss RA and Boyd MT (1996) Human endogenous retrovirus expression and reverse transcriptase activity in the T47D mammary carcinoma cell line. J Virol 70: 2654-2657.

Pavlicek A, Paces J, Ellder D and Hejnar J (2002) Processed pseudogene of human endogenous retroviruses generated by LINEs: their integrations, stability, and distribution. Genome Res 12: 391-399.

Pepinsky RB and Vogt VM (1984) Fine-structure analyses of lipid-protein and protein-protein interactions of gag protein p19 of the avian sarcoma and leukemia viruses by cyanogen bromide mapping. J Virol 52: 145-153.

Pietrokovski S and Henikoff S (1997) A helix-turn-helix DNA-binding motif predicted for transposases of DNA transposons. Mol Gen Genet 254: 689-695.

Prestridge DS (1991) SIGNAL SCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements. CABIOS 7: 203-206.

Reiter LT, Liehr T, Rautenstrauss B, Robertson HM and Lupski JR (1999) Localization of *mariner* DNA transposons in the human genome by PRINS. Genome Res 9: 839-843.

Reiter LT, Murakami T, Koeuth T, Pentano L, Muzny DM, Gibbs RA and Lupski JR (1996) A recombination hotspot responsible for two inherited peripheral neuropathies is located near a mariner transposon-like element. Nat Genet 12: 288-297.

Ribas M. Miró R, Gelagert A and Egozcue J (1999) Chromosome instability in lymphocytes from two patients affected three sequential primary cancers: the role of fragile sites. Cancer Genet Cytogenet. 110: 133-135.

Riggs AD (2002) X chromosome inactivation, differentiation, and DNA methylation revisited, with a tribute to Susumu Ohno. Cytogenet Genome Res 99: 17-24.

Rizwana R and Hahn PJ (1999) CpG methylation reduces genomic instability. J Cell Sci 112: 4513-4519.

Rolfe M and Banks G (1986) Induction of yeast Ty element transcription by ultraviolet light. Nature 319: 339−340.

Ruvolo M, Disotell TR, Allard MW, Brown WM and Honeycutt RL (1991) Resolution of the African hominid trichotomy by use of a mitochondrial gene sequence. Proc Natl Acad Sci USA 88: 1570-1574.

Salem AH, Myers JS, Otieno AC, Watkins WS, Jorde LB and Batzer MA (2003) LINE-1 preTa elements in the human genome. J Mol Biol 326: 1127-1146.

Samad F, Bergtrom G and Amrani DL (1994) Regulation of plasminogen activation by interleukin-6 in human lung fibroblasts. Biochem Biophys Acta 28: 307-314.

Sankoff D, Deneault M and Turbis P (2002) Chromosomal distributions of breakpoints in cancer, infertility, and evolution. Theor Popul Biol 61: 497-501.

Santos R, Schenberg ASG, Gardner DSG and SG Oliver (1997) Enhancement of Ty transposition at the *ADH*4 and *ADH*2 loci in meiotic yeast cells. Mol Gen Genet 254: 555−561.

Sassaman DM, Dombroski BA, Moran JV, Kimberland ML, Naas TP, DeBerardinis RJ, Gabriel A, Swergold GD and Kazazian HHJr (1997) Many human L1 elements are capable of retrotransposition. Nat Genet 16: 37-43.

Sauter M, Roemer K, Best B, Afting M, Schommer S, Seitz G, Hartmann M and Müeller-Lantzsch N (1996) Specificity of antibodies directed against Env protein of human endogenous retroviruses in patients with germ cell tumors. Cancer Res 56: 4362-4365.

Sauter M, Schommer S, Kremmer E, Remberger K, Dolken G, Lemm I, Buck M, Best B, Neumann-Haefelin D and Müeller-Lantzsch N (1995) Human endogenous retrovirus K10: expression of gag protein and detection of antibodies in patients with seminomas. J Virol 69: 414-421.

Sbrana I and Musio A (1995) Enhanced expression of common fragile site with occupational exposure to pesticides. Cancer Genet Cytogenet 82: 123-127.

Schneider PM, Witzel-Schlmp K and Rittner C (2001) The endogenous retroviral insertion in the human complement C4 gene modulates the expression of homologous genes by antisense inhibition. Immunogenetics 53: 1-9.

Schnieke A, Harbers K and Jaenisch R (1983) Embryonic lethal mutation in mice induced by retrovirus insertion intron the alpha$_1$ (I) collagen gene. Nature 304: 315-320.

Schön U, Seifarth W, Baust C, Hohenadl C, Erfle V and Leib-Mösch C (2001) Cell type-specific expression and promoter activity of human endogenous retroviral long terminal repeats. Virology 279: 280-291.

Seifarth W, Baust C, Murr A, Skladny H, Krieg-Schneider F, Blusch J, Werner T, Hehlman R and Leib-Mösch C (1998) Proviral structure, chromosomal location, and expression of HERV-K-T47D, a novel human endogenous retrovirus derived from T47D particles. J Viol 72: 8384-8391.

Sheen FM, Sherry ST, Risch GM, Robichaux M, Nasidze I, Stoneking M, Batzer MA and Swergold GD (2000) Reading between the LINEs: human genomic variation induced by LINE-1 retrotransposition. Genome Res 10: 1496-1508.

Sichangi MW, Langat DK and Mwenda JM (2002) Endogenous retrovirus sequences expressed in male mammalian reproductive tissues: a review. East Afr Med J 79: 38-41.

Simon M, Haltmeier M, Papakonstantinou G, Werner T Hehlmann R and Leib-Mösch C (1994) Transcription of HERV-K-related LTRs in human placenta and leukemic cells. Leukemia 8 Suppl 1: S12-S17.

Singer MF, Krek V, McMillan JP, Swergold GD and Thayer RE (1993) *LINE-1*: a human transposable element. Gene 135: 183-188.

Singh KP and Roy D (2001) Identification of novel breast tumor-specific mutation(s) in the q11.2 region of chromosome 17 by RAPD/AP-PCR fingerprinting. Gene 269: 33-43.

Sirokman GF, Humphries DE and Bing OHL (1997) Endogenous retroviral transcripts in myocytes from spontaneously hypertensive rats. Hypertension 30: 88-93.

Smit AF (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr Opin Genet Dev 9: 657-663.

Staleva LS and Venkov P (2001) Activation of Ty transposition by mutagens. Mutat Res 474: 93-103.

Strickland JE, Fowler AK, Kind PP, Hellman A, Kalter SS, Heberling RL and Helmke RJ (1973) Group specific antigen and RNA-directed DNA polymerase activity in normal baboon placenta. Proc Soc Exp Biol Med 144: 256-258.

Sugimoto J, Matsuura N, Kinjo Y, Takasu N, Oda T and Jinno Y (2001) Transcriptionally active HERV-K genes: Identification, isolation and chromosomal mapping. Genomics 72: 137-144.

Sutherland GR, Baker E and Richards RI (1998) Fragile sites still breaking. Trends Genet 14: 501-506.

Suzuki N, Fujiyoshi T, Maehara Y, Takahashi M, Yamamoto M and Endo H (1986) A new family of LTR-like sequences abundantly expressed in rat tumors. Nucleic Acids Res 14: 9271-9289.

Sverdlov DL (1998) Perpetually mobile footprints of ancient infections in human genome. FEBS Lett 428: 1-6.

Takai D, Yagi Y, Habib N, Sugimura T and Ushijima T (2000) Hypomethylation of LINE1 retrotransposon in human hepatocellular carcinomas, but not in surrounding liver cirrhosis. Jpn J Clin Oncol 30: 306-309.

Thompson JD, Higgins DG and Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.

Tighe PJ, Stevens SE, Dempsey S, Le Deist F, Rieux-Laucat F and Edgar JD (2002) Inactivation of the Fas gene by Alu insertion: retrotransposition in an intron causing splicing variation and autoimmune lymphoproliferative syndrome. Genes Immun 3 Suppl 1: S66-70.

Ting CN, Rosenberg MP, Snow CM, Samuelson LC and Meisler MH (1992) Endogenous retroviral sequences are required for tissue-specific expression of human salivary amylase gene. Genes Dev 6: 1457-1465.

Toffolatti L, Cardazzo B, Nobile C, Danieli GA, Gualandi F, Muntoni F, Abbs S, Zanetti P, Angelini C, Ferlini A, Fanin M and Patarnello T (2002) Investigating the mechanism of chromosomal deletion: characterization of 39 deletion breakpoints in introns 47 and 48 of the human dystrophin gene. Genomics 80: 523-530.

Tönjes RR, Boller K, Limbach C, Lugert R and Kurth R (1997) Characterization of human endogenous retrovirus type K virus-like particles generated from recombinant baculoviruses. Virology 233: 280-291.

Towler DA, Adams SP, Eubanks SR, Towery DS, Jackson-Machelski E, Glaser L and Gordon JI (1988) Myristoyl CoA: protein N-myristoyltransferase activities from rat liver and yeast possess overlapping yet distinct peptide substrate specificities. J Biol Chem 263: 1784-1790.

Tükün A, Renda Y, Topçu M, Tuncali T and Bökesoy I (2000) Mental retardation with rare fragile site expressed at 2q11. Brain Devel 22: 498-500.

Tunca B, Egeli U, Zorluoglu A, Yilmazlar T, Yerci O, Kizil A (2000) The expression of fragile sites in lymphocytes of patients with rectum cancer and their first-degree relatives. Cancer Lett 152: 201-209.

Varmus H (1988) Retroviruses. Science 240: 1427-1435.

Varmus H (1987) Reverse transcription. Sci Am 257: 56-59.

Vinogradova TV, Leppik LP, Nikolaev LG, Akopov SB, Kleiman AM, Senyuta NB and Sverdlov ED (2001) Solitary human endogenous retroviruses-K LTRs retain transcriptional activity *in vivo*, the mode of which is different in different cell types. Virology 290: 83-90.

Wilkinson DA, Mager DL and Leong JA (1994) *Endogenous human retroviruses*, pp. 465-535. In The Retroviridae, J.A. Levy ed., Plenum Press, NY.

Wilson GA and Young FE (1975) Isolation if a sequence specific endonuclease (*Bam*I) from *Bacillus amyloliquefaciens* H. J Mol Biol 97: 123-125.

Wingender E (1994) Recognition of regulatory regions in genomic sequences. J Biotechnol 35: 273-280.

Woodcock DM, Lawler CB, Linsenmeyer ME, Doherty JP and Warren WD (1997) Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. J Biol Chem 272: 7810-7816.

Woods-Samuels P, Wong C, Mathias SL, Scott AF, Kazazian HHJr and Antonarakis SE (1989) Characterization of a nondeleterious L1 insertion in an intron of the human factor VIII gene and further evidence of open reading frames in functional L1 elements. Genomics 4: 290-296.

Wu J, Zhou T, He J and Mountz JD (1993) Autoimmune disease in mice due to integration of an endogenous retrovirus in an apoptosis gene. J Exp Med 178:

461-468.

Yi JM, Kim HM and Kim HS (2001) Molecular cloning and phylogenetic analysis of the human endogenous retrovirus HERV-K long terminal repeat elements in various cancer cells. Mol Cells 12: 137-141.

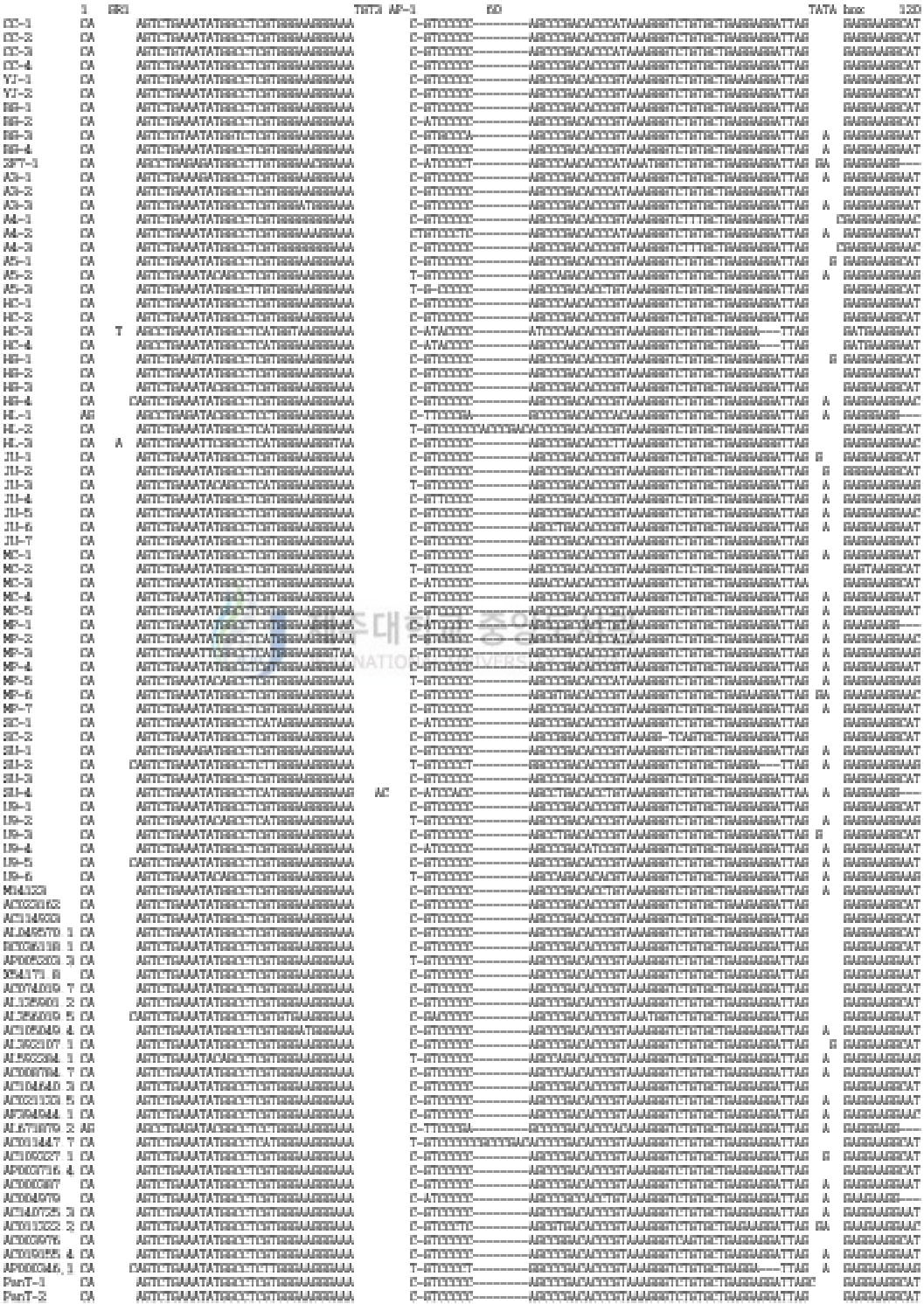
Yoder JA, Walsh CP and Bestor TH (1997) Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 13: 335−340.

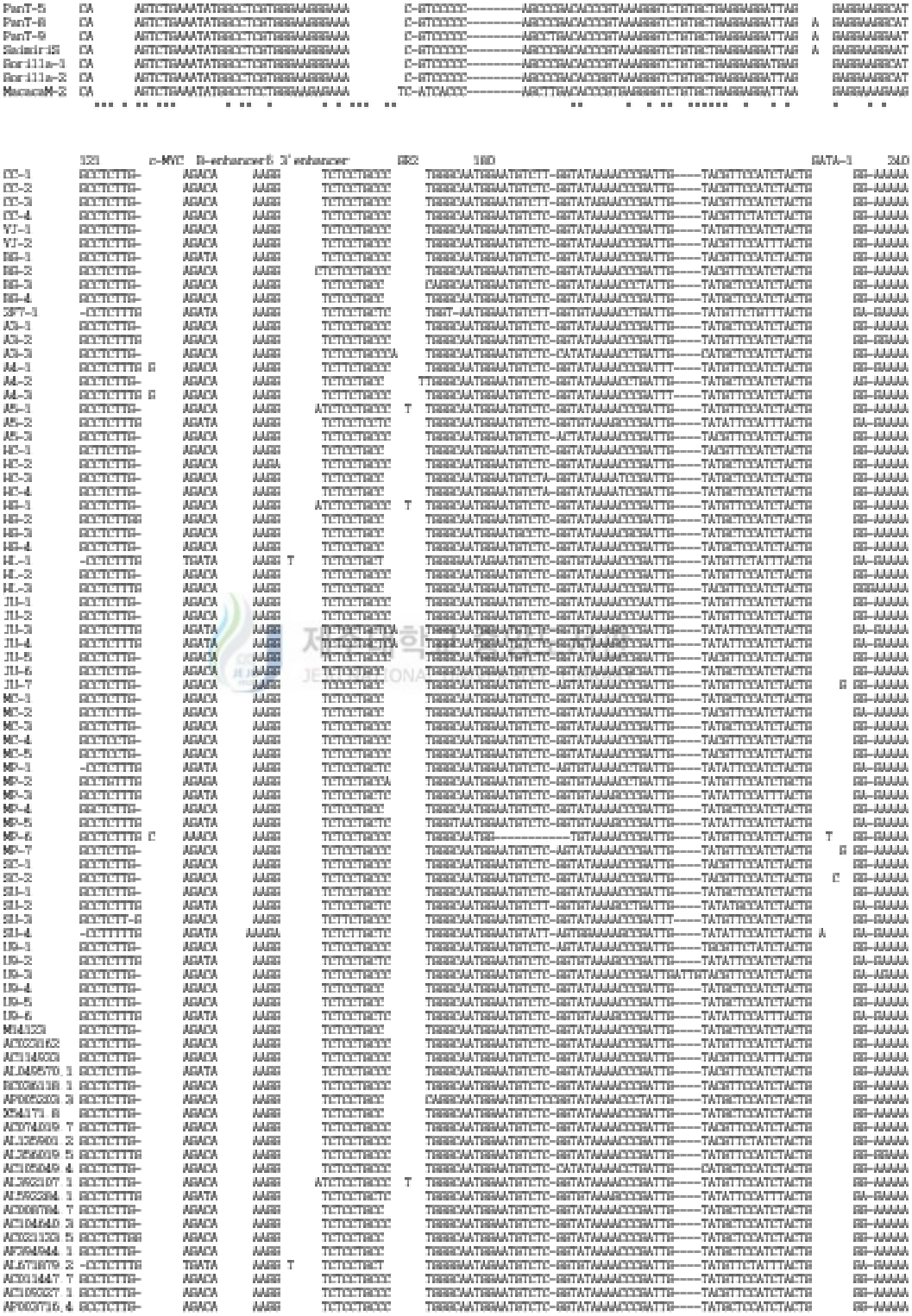
Youssoufian H, Kasper CK, Phillips DG, Kazazian HHJr and Antonarakis SE (1988) Restriction endonuclease mapping of six novel deletions of the factor VIII gene in hemophilia A. Hum Genet 80: 143-148.

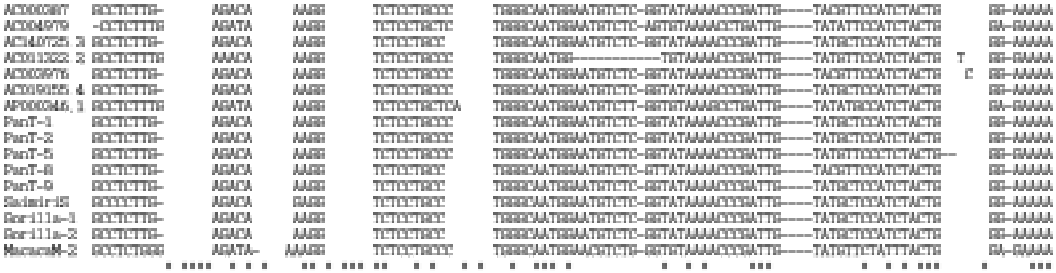
Zsiros J, Jebbink MF, Lukashov VV, Voute PA and Berkhout B (1998) Evolutionary relationships within a subgroup of HERV-K-related human endogenous retroviruses. J Gen Virol 79: 61-70.

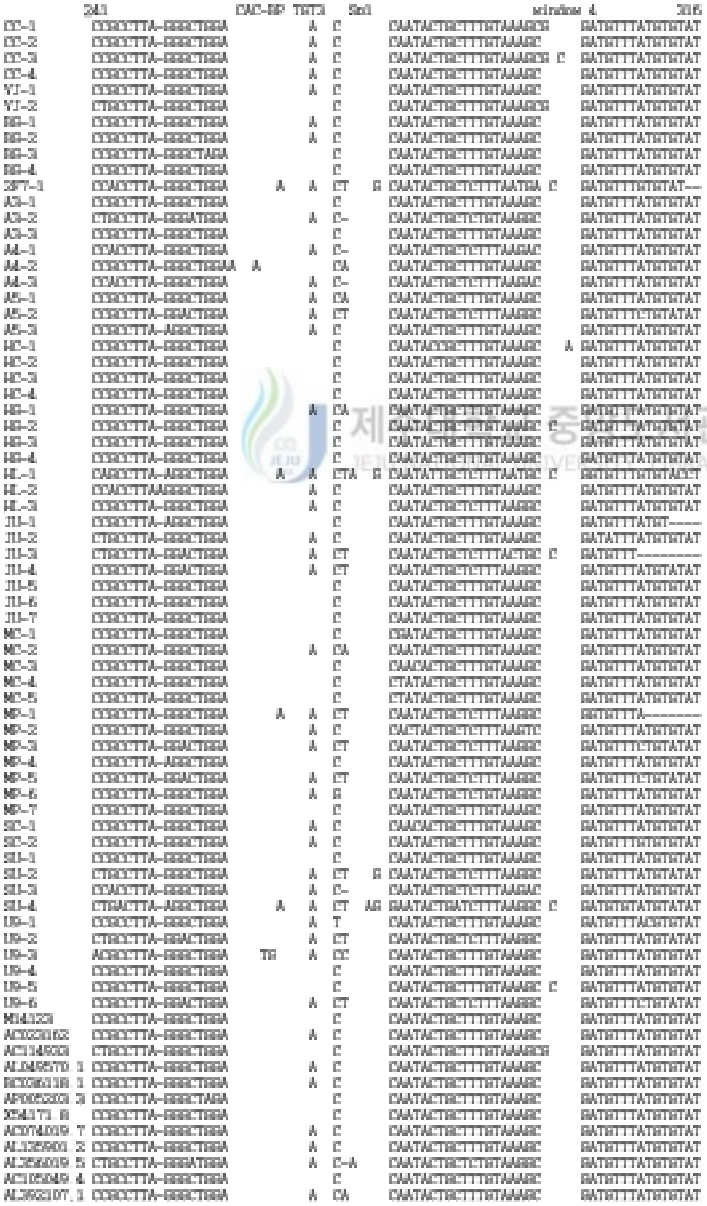
APPENDIX I

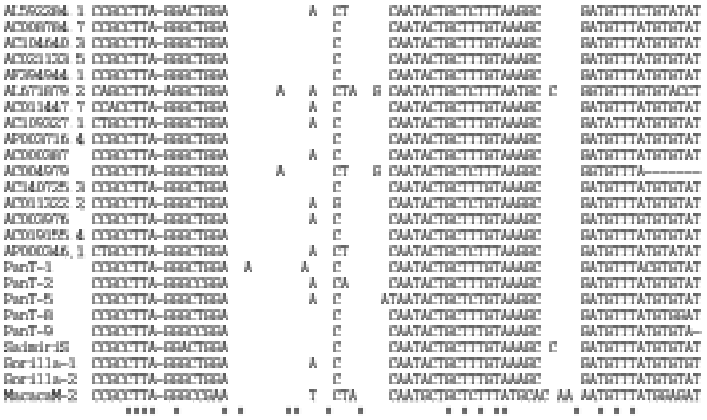
Multiple aligned sequences of the HERV-K10 LTR from human cells







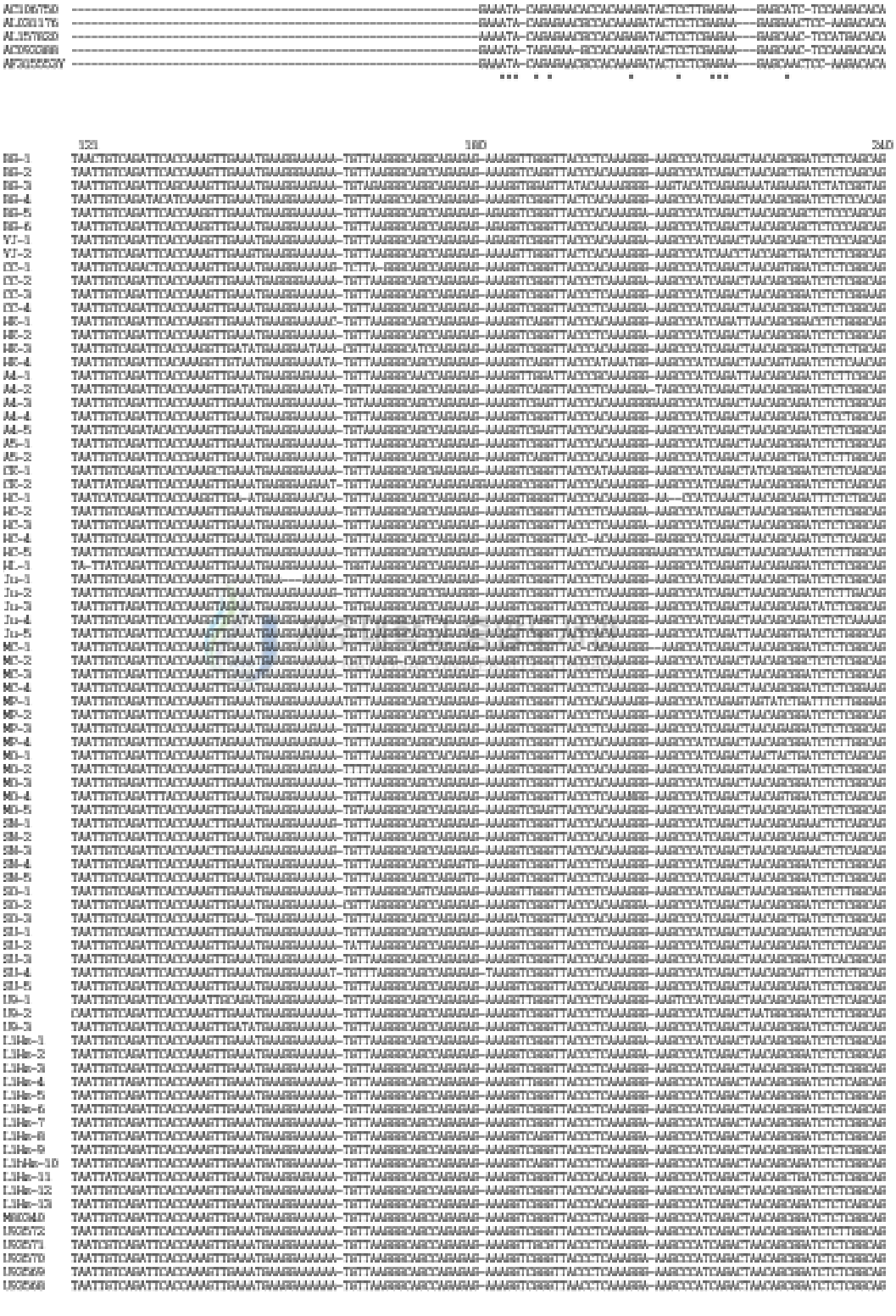


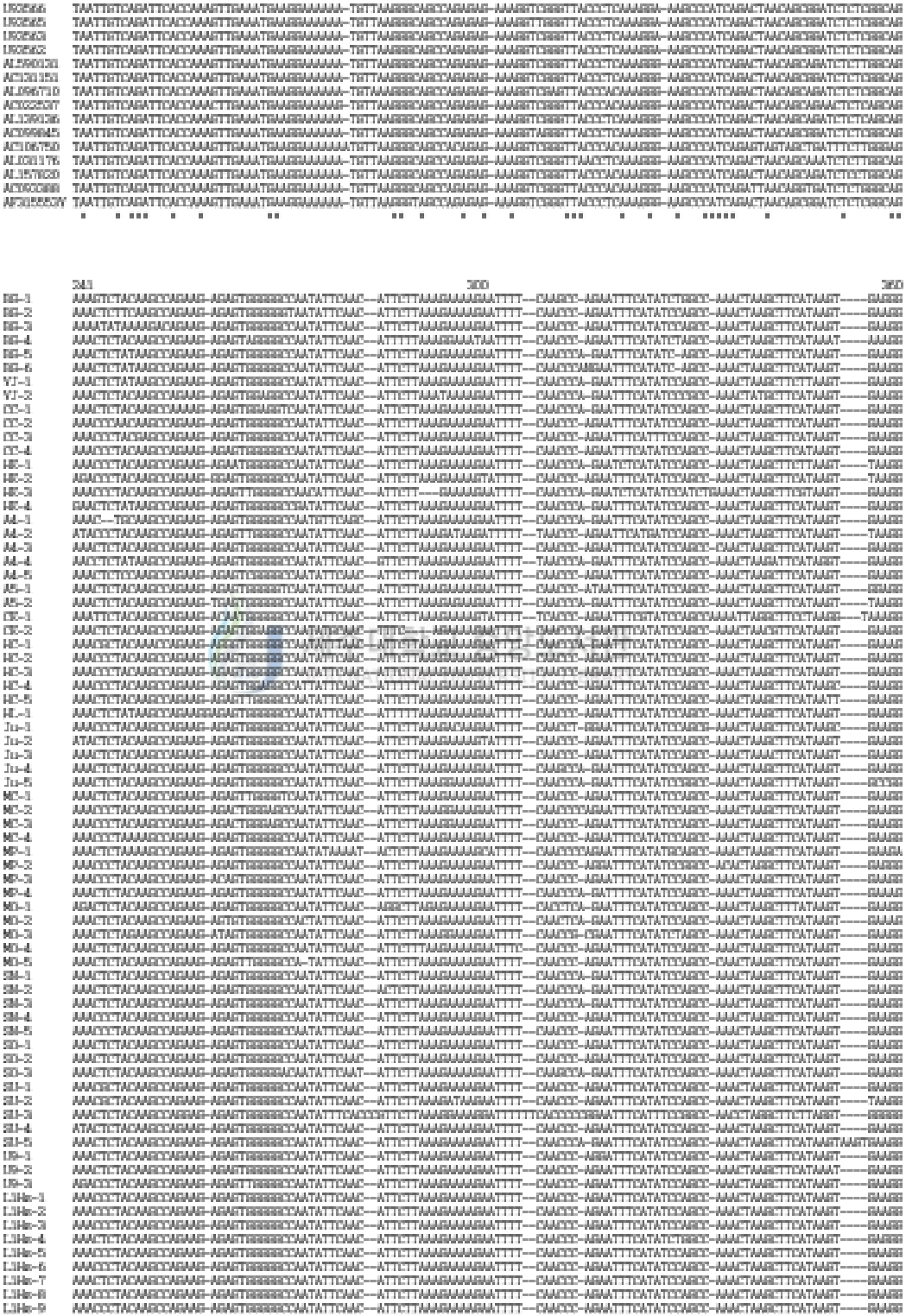


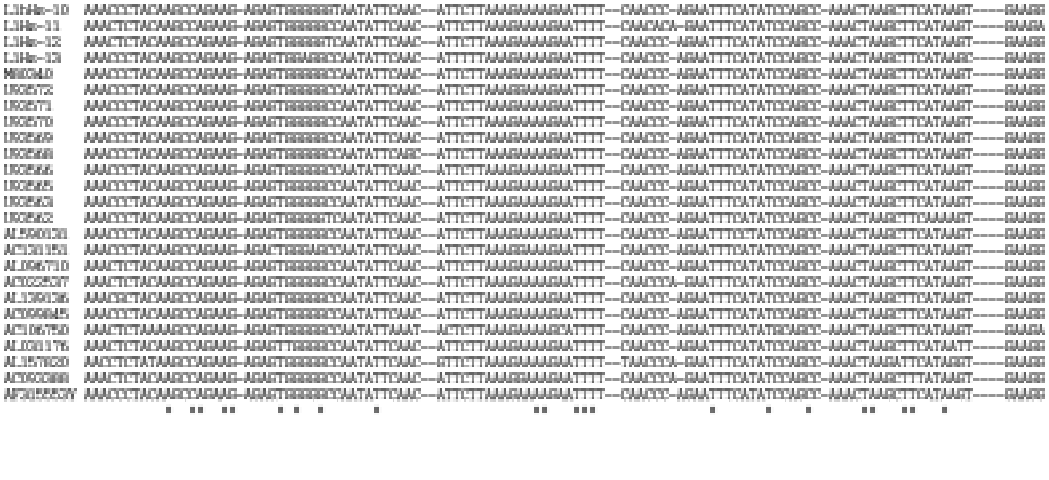
Abbreviations of each K10 LTR clone were represented in Table 4. Bars in the sequences indicate the gaps, and asterisks under the sequences do the identical sequences. Shadowed sequences are the conserved and corresponding to each signal at the top of the sequences. Gorilla, *Gorilla gorilla*; MacacaM, *Macaca mulatta*; Pant, *Pan troglodytes*; SaimiriS, *Saimiri sciureus*. AP-1, AP-1 binding site; CAC-BP, CAC-binding protein binding site; c-MYC, c-MYC binding site; GATA-1, GATA-1 binding site; GR1, glucocorticoid hormone response element 1, GR2, delayed secondary glucocorticoid response element; Sp1, Sp1 binding site; TGT3, TGT3 binding site; window 4 and window 10, human epidermal growth factor response elements.

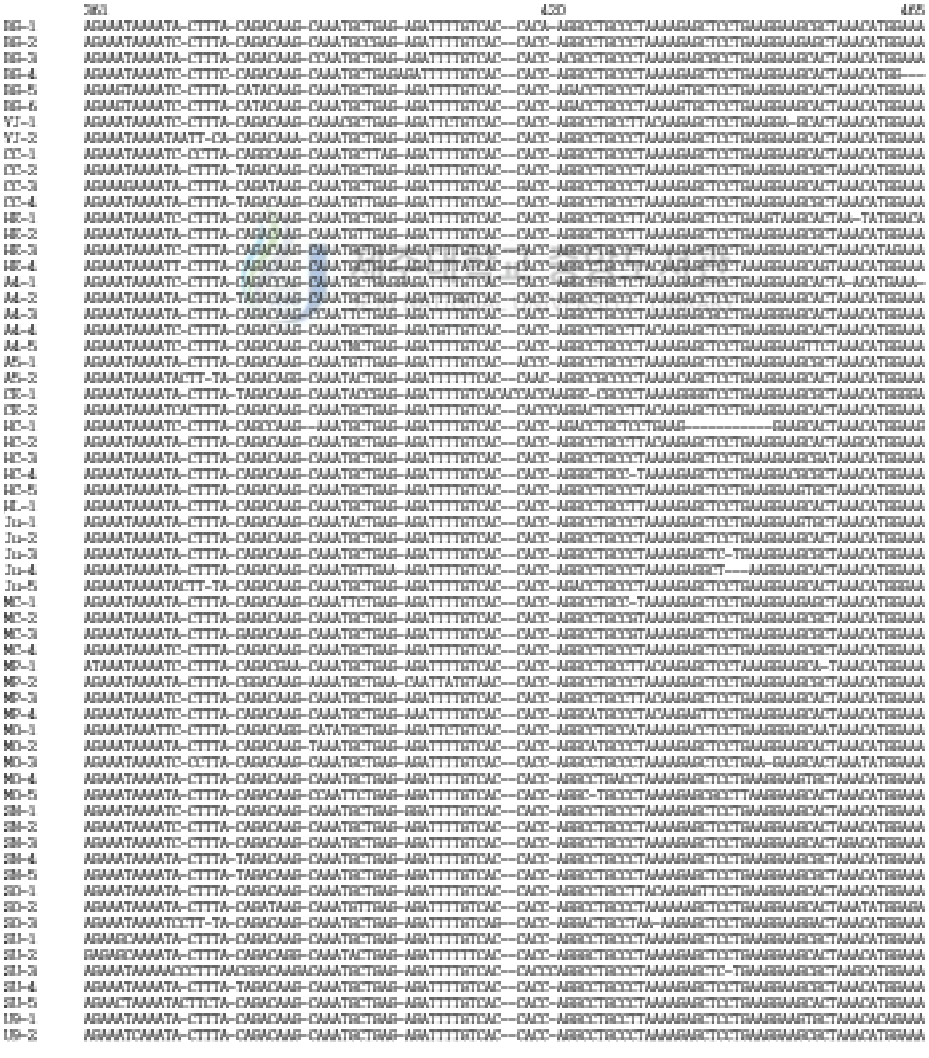
APPENDIX Ⅱ

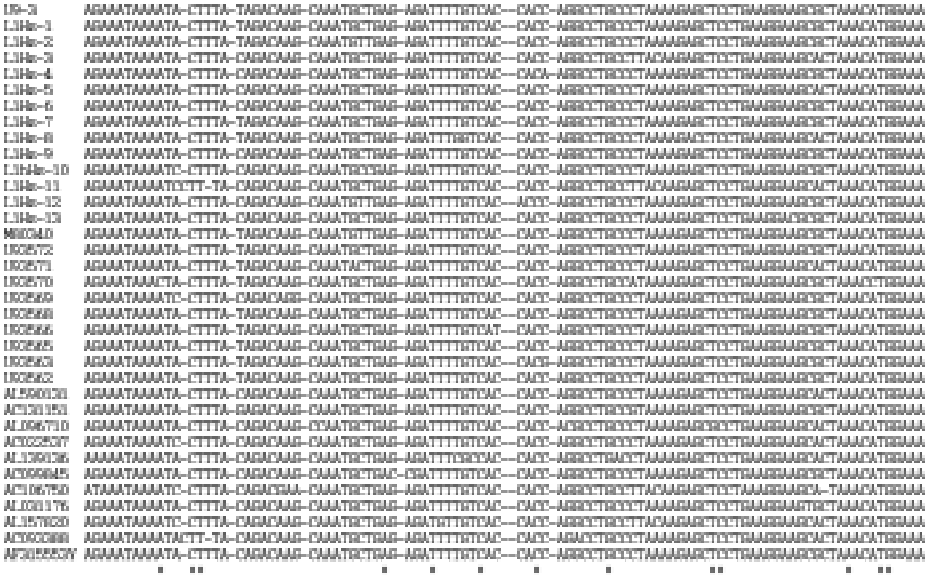
Multiple aligned sequences of the L1Hs *p40* gene from human cells











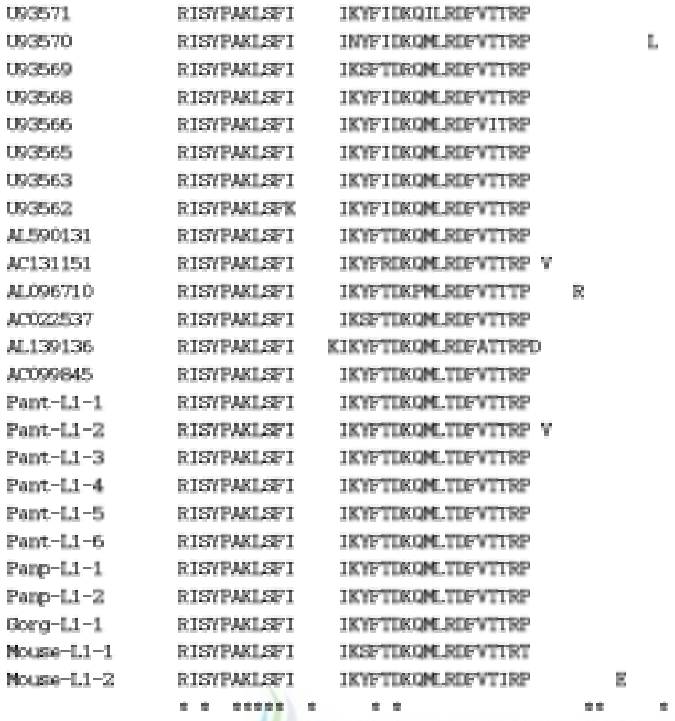
Abbreviations of each L1Hs *p40* clone and each symbol of L1Hs groups were represented in Table 8. Bars in the sequences indicate the gaps, and asterisks under the sequences do the identical sequences.

APPENDIX Ⅲ

Multiple aligned sequences of the L1Hs p40 protein from human cells

```
                1                  20                  40                  60                  80
                     PKC(1, 2, and 3)  CX11                PKC4
BG-1            EIQRTPQRY         HITVRF    MKEKMLRAGREKDGWV    GKPIRLTADLSAESLQARREMGPIFNILKEKNPQA
CC-2            EIQRMPQRY         HIIVRF    MRDKMLRAAREKDRV     GKPIRLTADLSAETQQARREMGPIFNILKEKNPQP
CC-3            EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSRETLRARREMGPIFNILKEKNPQP
CC-4            GIQRMPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
A5-2            EIQRMPQRC         HIIVRF E  MKEKMLRAAREKDQV H  GKPIRLTADLLAETLQARSEMGPIFNILKEKNPQP
HC-2            EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
HC-3            EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
Ju-2            EIQRTQQRS L       HIIVRF    MKEKVLRAAEKDRV      GKPIRLTADLLTDTLQARREMGPIFNILKEKYPQP
MC-3            EIQRTRQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARRDMEPIFNILKDKNPQP
MC-4            EIQRTPQRY   G     HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSRETLKARREMGPIFNILKEKNPQP
MP-3            EIQRMPQRY         HIIVRF    MKEEMLRAAREKDRV     GKPIRLTEDLSAETLQARRQMGPIFNILKEKNPQP
MP-4            EIQRTPQRS         HIIVRF    MKEEMLRAGREKDRV H   GKPIRLTADLLAETLQARREMGPIFNILKEKNPQP
MO-1            EIQRTPQRY   T     NIIVRF    MKEKMLRAATEKDRV H   GKPIRLTTDLSAETLQARREMGPIFNMLREKNPHL
MO-2            KIQRTPQRY L TTP   HIILRF    MKEEILRAAREKDRV H   GKPIRVTADLSAETLQARRYMGPLFNILKEKNPQL
SK-1            EIQRTPQRY         HIIVRF    MKEKVLRAAREKDRV H   GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
SK-2            EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV H   GKPIRLTADLSAETLQARREMGPIFNTLKEKNPQP
SK-3            EIQRAPQRY        IHIIVRF    KKEKVLRAAREKDRV H   GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
SK-4            EIQRMPQRY         HIIVRF    MKEKMLRAARYKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
SK-5            EIQRMPQRY         HIIVRF    MKEKMLRAARYKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
SO-1            GIQRMPQRY         HIIVRF    MKEKMLRAVREKDWV     GKPIRLTADLLAETLQARREMGPIFNILKEKNPQP
SO-2            EIQRMPQRY         HIIVRF    MKEKTLQAAREKDRV H   GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
SU-1            EIQRTPQRYA        RIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
U9-1            EIQRMPQRY M       HIIVRF  LQMKEKMLRAAREKDWV     GKSIRLTADLSAETLQARREMGPIFNILKEKNPQP
U9-2            EIQRMPQRY L  V R  HTIVRF    MKEKMLRAAREKDRV     GKPIRLMADLSAETLQARREMGPIFNILKEKNPQP
U9-3            EIQRTPQRY         HIIVRF  IMKEKMLRAAREKDRV     GKPIRLTADLSAETLQARRELGPIFNILKEKNPQP
L1Hs-1          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
L1Hs-2          GIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
L1Hs-3          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARRDMGPIFNILKEKNPQP
L1Hs-4          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDWV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
L1Hs-5          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
L1Hs-6          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
L1Hs-7          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
L1Hs-8          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDQV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
L1Hs-9          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
L1HS-10         EIQRTPQRY         HIIVRF    MMEKMLRAAREKDQV     GKPIRLTADLSAETLQARREMGVIFNILKEKNPQP
L1Hs-11         EIQRMPQRY    V    HIIIRF    MKEKMLRAAREKDRV H   GKPIRLTADLSAETLQARREMGPIFNILKEKNPQH
L1Hs-13         EIQRTPRRY         HIIVRF    MKEKMLRAAREKDRV H   GKPIRLTADLSAETLQARREMGPIFNIFKEKNPQP
MB0340          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
UW3572          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLLAETLQARREMGPIFNILKDKNPQP
UW3571          EIQRMPQRY         HIIVRF    MKEKMLRAAREKDCV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
UW3570          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
UW3569          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
UW3568          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRVN    GKPIRLTADLSAETLQARREMGPIFSILKEKNPQP
UW3566          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
UW3565          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDWV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
UW3563          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
UW3562          EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGSIFNILKEKNPQP
AL590131        EIQRTQQRS L       HIIVRF    MKEKMLRAAGKDRV      GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
AC131151        EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMEPIFNILKDKNPQP
AL096710        EIQRMPQRY   G     HIIVRF    MKEKMLRAAREKDRV H   GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
AC022537        EIQRTPQRY         HIIVRF    MKEKVLRAAREKDRV H   GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
AL139136        EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
AC099845        EIQRTPQRY         HIIVRF    MKEKMLRAAREKDRV     GKPIRLTADLSAETLQARREMGPIFNILKEKNPQP
```

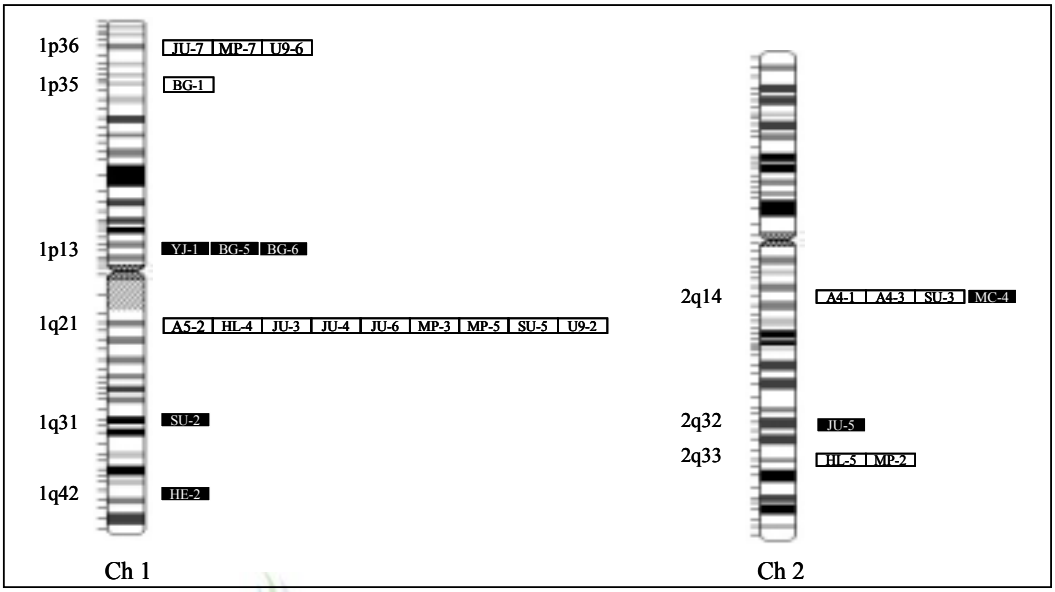
```
Pant-L1-1       EIQRTPQRY        HIIVRF    MKEKMLRAARQKGRV   GKPIRLTADLSAETLQARREWGPIFNILKEKNFQP
Pant-L1-2       EIQRTPQRY        HIIVRF    MKEKMLRAARQKGRV   GKPIRLTADLSAETLQARREWGPIFNILKEKNFQP
Pant-L1-3       EIQRTPQRY        HIIVRF    MKEKMLRAARQKGRV   GKPIRLTADLSAETLQARREWGPIFNILKEKNFQP
Pant-L1-4       EIQRTPQRY        HIIVRF    MKEKMLRAARQKGRV   GKPIRLTADLSAETLQARREWGPIFNILKEKNFQP
Pant-L1-5       EIQRTPQRY        HIIVRF    MKEKMLRAARQKGRV   GKPIRLTADLSAETLQARREWGPIFNILKEKNFQP
Pant-L1-6       EIQRTPQRY        HIIVRF    MKEKILRAARQKGRV   GKPIRLTADLSAETLQARREWGPIFNILKEKNFQP
Panp-L1-1       EIQRTPQRY        HIIVRF    MKEKMLRAARQKGRV   GKPIRLTADLSAETLQARREWGPIFNILKEKNFQP
Panp-L1-2       ERQRTPQRY        HIIVRF    MKEKMLRAARQKGQV   GKPIRLTADLSAETLQARREWGPIFNILKEKNFQP
Gorg-L1-1       EIQRTPQRY        HIIVRF    MKEKMLRAAREKGRV   GKPIRLTADLSAETLQARREWGPIFNILKEKNFQP
Mouse-L1-1      EIQRTPQRY        HIIVRF    MQEKMLRAAKEKGQV   GKPIRLTADPLAETLQARREWGPISNLLKEKNFQP
Mouse-L1-2      EIQRTPQRY        HIIVRF    MKEKMLRTAREKGRV   GKPIRLTVDLSAETLQARREWGPIFNILKGKNFQP
                  **        *         ***         *     ** *  *** **           **             * *

                81               100                 120
                           CKII                      HTH
BG-1            RISYLAKLSFI    IKYFTDKQMLRDFVTTRP
CC-2            RISYPAKLSFI    IKYFIDKQMLRDFVTTRP
CC-3            RISFPAKLSFI    RKYFTDKQMLRDFVTTRP
CC-4            RISYPAKLSFI    IKYFIDKQMLRDFVTTRP
A5-2            RISYPAKLSFI    IKYFTDRQILRDFFTNRPP Q
HC-2            RISYPAKLSFI    IKYFTDKQMLRDFVTTRP  Q       S
HC-3            RISYPAKLSFI    IKYFTDKQMLRDFVTTRP         I
Ju-2            RISYPAKLSFI    IKYFTDKQMLRDFVTTRP
MC-3            RISYPAKLSFI    IKYFRDKQMLRDFVTTRP V
MC-4            RISYPAKLSFI    IKSFTDKQMLRDFVTTRP
MP-3            RISYPAKLSFI    IKSFTDKQMLRDFVTTRP  Q
MP-4            RFSYPAKLSFI  R IKSFTDKQMLRNFVTTRH  Q F
MO-1            RISYPAKLSFI    INSFTDRHMLRDSVTTRP I D   G I
MO-2            RISYPAKLSFI  R IKYFTDKHMLRDFVTTRH
SN-1            RISYPAKLSFI    IKSFTDKQMLRDFVTTRP
SN-2            RISYPAKLSFI    IKSFTDKQMLRDFVTTRP
SN-3            RISYPAKLSFI    IKSFTDKQMLRDFVTTRP          D
SN-4            RISYPAKLSFI    IKYFIDKQMLRDFVTTRP
SN-5            RISYPAKLSFI    IKYFIDKQMLRDFVTTRP
SD-1            RISYPAKLSFI    IKYFTDKQMLRDFVTTRP  Q F
SD-2            RISYPAKLSFI    IKYFTDKQMLRDFVTTRP   K
SU-1            RISYPAKLSFI    AKYFTDKQMLRDFVTTRP
U9-1            RISYPAKLSFI    IKYFTDKQMLRDFVTTRP       V  T
U9-2            RISYPAKLSFIN   IKYFTDKQMLRDFVTTRP
U9-3            RISYPAKLSFI    IKYFIDKQMLRDFVTTRP
L1Hs-1          RISYPAKLSFI    IKYFIDKQMLRDFVTTRP
L1Hs-2          RISYPAKLSFI    IKYFIDKQMLRDFVTTRP
L1Hs-3          RISYPAKLSFI    IKYFTDKQMLRDFVTTRP  Q
L1Hs-4          RISYLAKLSFI    IKYFTDKQMLRDFVTTRP
L1Hs-5          RISYPAKLSFI    IKYFTDKQMLRDFVTTRP
L1Hs-6          RISYPAKLSFI    IKYFTDKQMLRDFVTTRP
L1Hs-7          RISYPAKLSFI    IKYFIDKQMLRDFVTTRP
L1Hs-8          RISYPAKLSFI    IKYFIDKQMLRDLVTTRP   D
L1Hs-9          RISYPAKLSFI    IKYFTDKQMLRDFVTTRP
L1HS-10         RISYPAKLSFI    IKSFTDKQMPRDFVTTRP
L1Hs-11         RISYPAKLSFI  E IKSFTDKQMLRDFVTTRP  Q
L1Hs-13         RISYPAKLSFI    IKYFTDKQMLRDFVTTRP       D
M80340          RISYPAKLSFI    IKYFIDKQMLRDFVTTRP
U93572          RISYPAKLSFI    IKYFIDKQMLRDFVTTRP
```

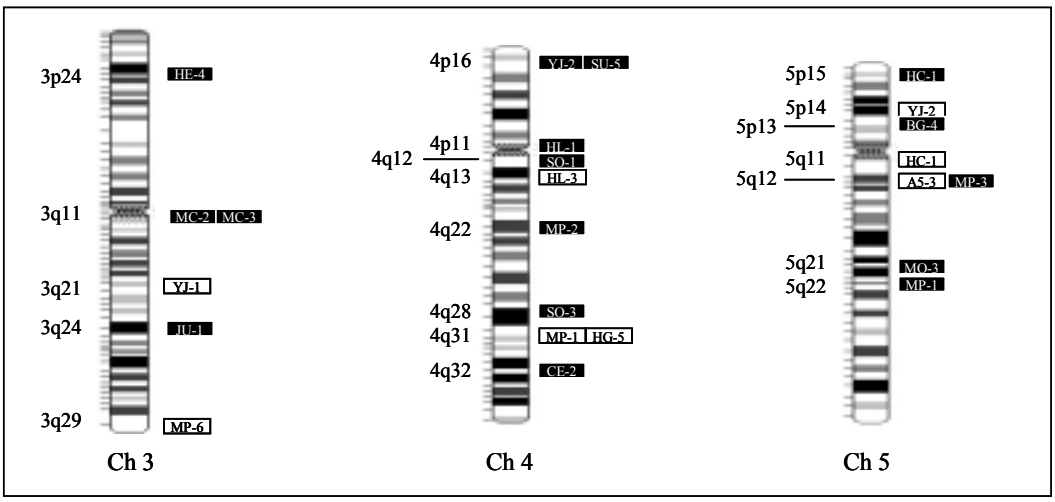


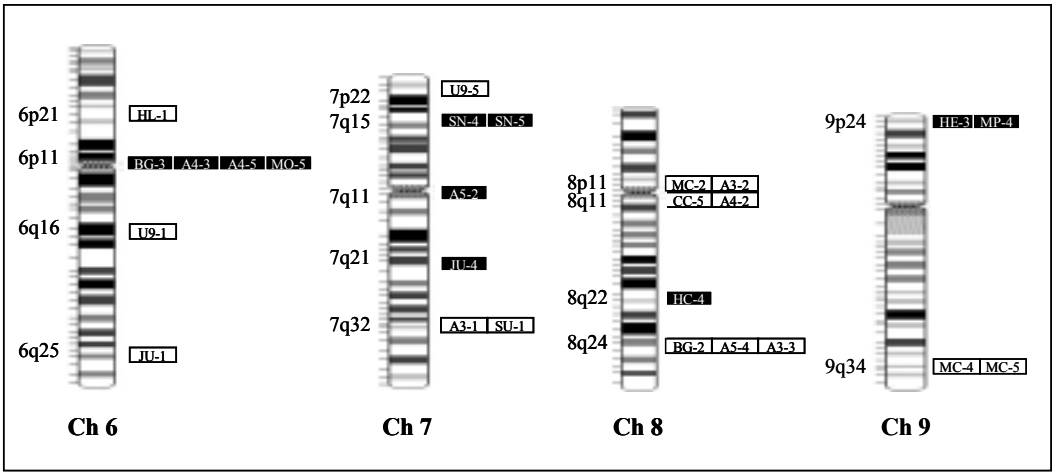
Abbreviations of each L1Hs *p40* clone and each symbol of L1Hs groups were represented in Table 8. Gorg, *Gorilla gorilla*; Mouse, *Mus musculus*; Panp, *Pan paniscus*; Pant, *Pan troglodytes*. HTH, helix-turn-helix motif; PKC, protein kinase C phosphorylation site; CKII, casein kinase II phosphorylation site. Asterisks indicate the identical sequences. Shadowed sequences are the conserved sequences corresponding to each protein interaction motif at the top of the sequences.

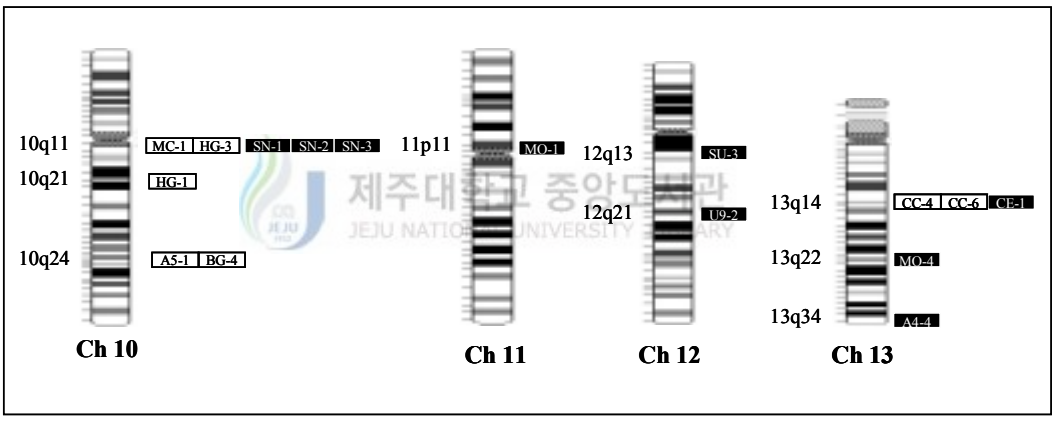
APPENDIX Ⅳ

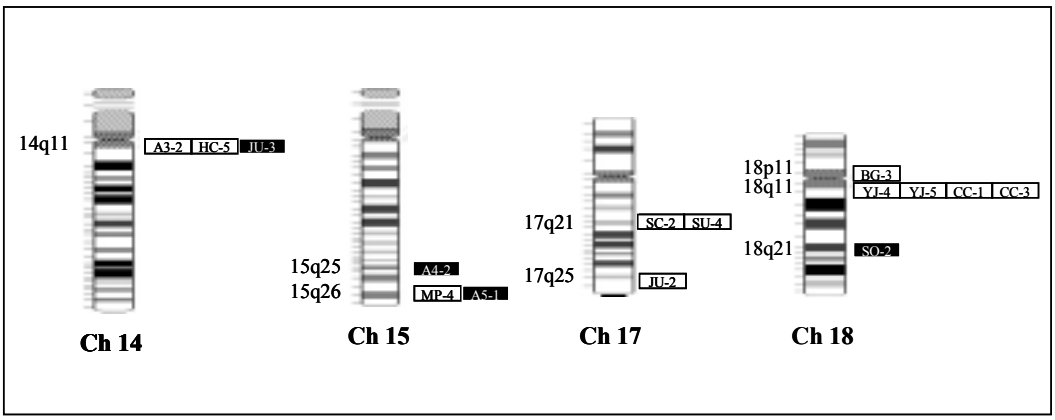
Chorompsomal locations of each transcript of the HERV-K10 LTR and the L1Hs *p40*

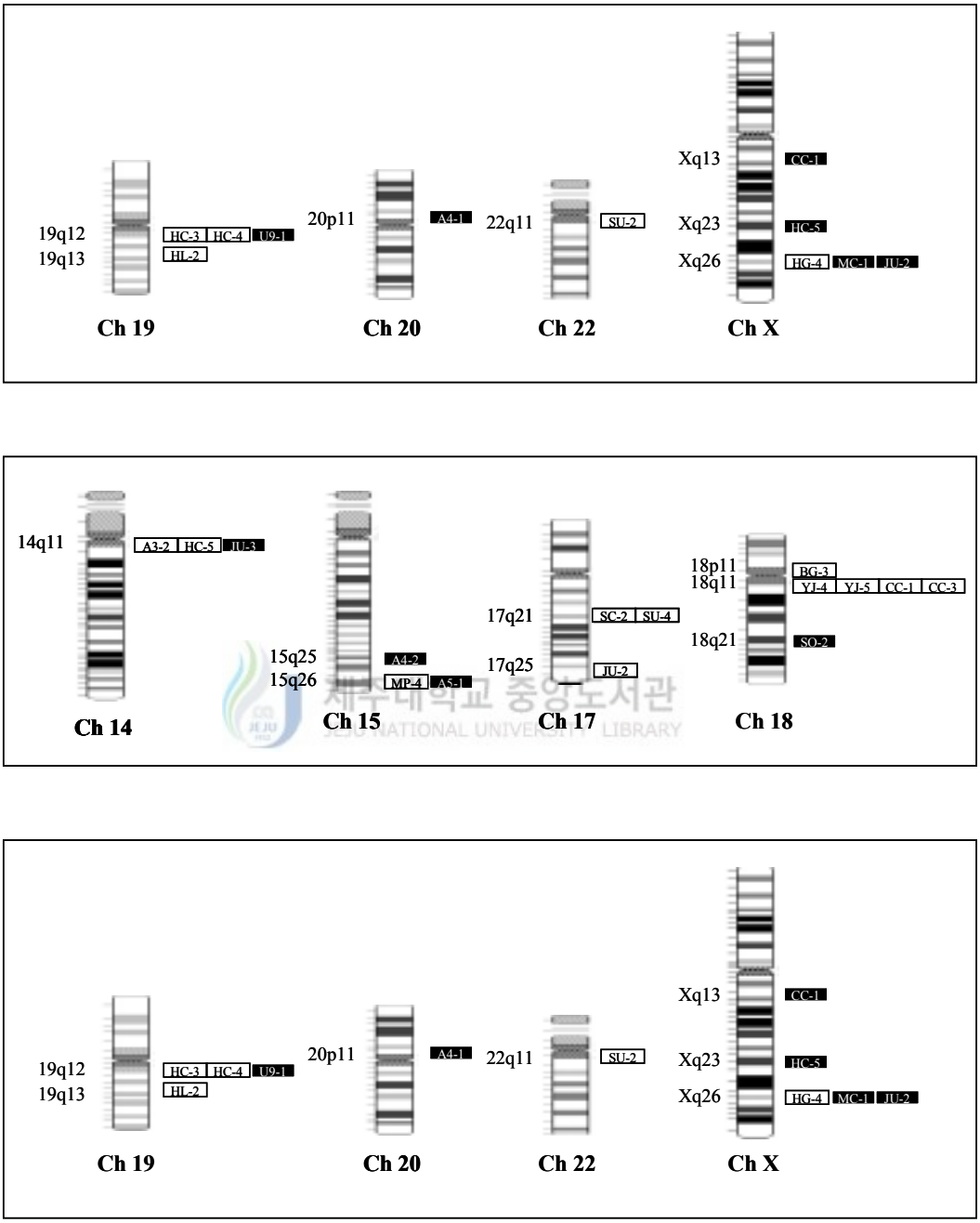












Idiograms of the human chromosomes were obtained from Human Genome Resources in GenBank. Information of each clone was given in Table 5 to HERV-K10 LTR and Table 9 to L1Hs *p40*. Open rectangles indicate the transcript of HERV-K10 LTR and filled rectangles do that of L1Hs *p40*.

APPENDIX Ⅴ

Examples of Default Display of L1Hs Sequences Deposited in GenBank Database and the Search Result of the Part of Submissions

```
LOCUS       AY392754              379 bp    DNA     linear   PRI 15-OCT-2003
DEFINITION  Homo sapiens map 7q cell-line CCD-25Lu LINE-1 element, partial
            sequence.
ACCESSION   AY392754
VERSION     AY392754.1  GI:37624749
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Han,S.-H., Oh,J.-H., Oh,Y.-S., Kim,J.-H., Song,J.-H. and Oh,M.-Y.
  TITLE     LINE-1 sequence variations in human lung fibroblasts
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 379)
  AUTHORS   Han,S.-H., Oh,J.-H., Oh,Y.-S., Kim,J.-H., Song,J.-H. and Oh,M.-Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-2003) Department of Life Science, College of
            Natural Sciences, Cheju National University, 1 Ara 1-Dong, Jeju,
            Jeju 690-756, Korea
FEATURES             Location/Qualifiers
     source          1..379
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q"
                     /cell_line="CCD-25Lu"
                     /cell_type="normal fibroblast"
                     /tissue_type="lung"
     repeat_region   <1..>379
                     /note="p40"
                     /rpt_family="LINE-1"
                     /rpt_type=dispersed
ORIGIN
        1 ggaatacaga gaatgccaca aagatactcc ccgagaagag caactccaag acacataatt
       61 gtcagattca ccaaagttga aatgaaggaa aaaatgttaa gggcagccag agagaaaggt
      121 cgggttaccc tcaaaggaaa gcccatcaga ctaacagcgg atctctcggc agaaacccta
      181 caagccagaa gagagtgggg gccaatattc aacattctta aagaaaagaa ttttcaaccc
      241 agaatttcat atccagccaa actaagcttc ataagtgaag gagaaataaa atactttata
      301 gacaagcaaa tgttgagaga ttttgtcacc accaggcctg ccctaaaaga gctcctgaag
      361 gaagcgctaa acatggaaa
```

1: AY392754 Links
Homo sapiens map 7q cell-line CCD-25Lu LINE-1 element, partial sequence
gi|37624749|gb|AY392754.1|[37624749]

2: AY392753 Links
Homo sapiens map 10q cell-line CCD-25Lu LINE-1 element, partial sequence
gi|37624748|gb|AY392753.1|[37624748]

3: AY392752 Links
Homo sapiens map 7p cell-line CCD-25Lu LINE-1 element, partial sequence
gi|37624747|gb|AY392752.1|[37624747]

4: AY392751 Links
Homo sapiens map Xq13.1 cell-line CCD-25Lu LINE-1 element, partial sequence
gi|37624746|gb|AY392751.1|[37624746]

5: AY392750 Links
Homo sapiens map Xq23 cell-line HCT-15 LINE-1 element, partial sequence
gi|37624745|gb|AY392750.1|[37624745]

6: AY392749 Links
Homo sapiens map 8q cell-line HCT-15 LINE-1 element, partial sequence
gi|37624744|gb|AY392749.1|[37624744]

7: AY392748 Links
Homo sapiens map 7q cell-line HCT-15 LINE-1 element, partial sequence
gi|37624743|gb|AY392748.1|[37624743]

8: AY392747 Links
Homo sapiens map 5q cell-line HCT-15 LINE-1 element, partial sequence
gi|37624742|gb|AY392747.1|[37624742]

9: AY392746 Links
Homo sapiens map 5p cell-line HCT-15 LINE-1 element, partial sequence
gi|37624741|gb|AY392746.1|[37624741]

초 록

전위인자 (transposable element)는 유전체 기능에 중요한 역할을 담당하며 전위와 재조합을 통해서 염색체의 다른 위치로 전위할 수 있다. 그 결과로 생성된 반복서열들은 많은 유전적 재조합의 기회를 제공하게 된다. 뿐만 아니라, 염색체상에서 전위인자의 위치 변화는 유전체 불안정성, 세포의 형질전환과 분화를 초래하기도 한다. 인간 유전체에서 몇 종류의 역전위인자 (retro-transposable elements)들이 알려져 있다. 대표적 역전위인자인 인간내생성레트로바이러스 (HERVs)와 긴 산재형 고반복 인자 (LINEs)들은 기능적인 전위활성 보유하고 있다. HERV는 인간 유전체의 0.5%-3% 정도로 추정된다; 전형적인 레트로바이러스 단백질인 gag, pol, env를 암호화하고 있으며, 양 끝에 유사 프로모터로 작용하는 장만달반복배열 (LTR)들이 위치한다. 인간 유전체에서 발견된 HERV LTR과 독자적인 LTR들은 세포생물학적으로 중요한 역할을 수행하고 있다. LTR들은 HERV 유전자 및 인접한 숙주 유전자의 발현을 조절하는 다양한 요소들을 보유하고 있다. 인간 LINE-1 (L1Hs)은 두 개의 열린 해독틀 (ORF)이 내재하고 있다: 그 중 하나는 역전사를 담당하는 역전사효소를 암호화하고, 다른 하나는 RNA-단백질 복합체 형성을 담당한다. 본 연구는 다양한 인간 세포주들 (2종의 정상 림프구와 2종의 정상 섬유아세포, 16 종의 암세포주)을 대상으로 HERV-K와 L1Hs의 전사 및 전사체의 염색체 위치를 조사하였다. 또한, 이들 역전위인자들의 유전자 발현에서 몇 가지 화학적 발암물질의 효과를 시험하였다. 역전위인자들의 RNA 전사체는 사용된 모든 세포주에서 발견되었으나, 확인된 전사체의 상대적인 양은 세포 유형에 따라 서로 다르게 나타났다. 예를 들어, 섬유아세포들과 백혈구 세포들에서 전사 수준이 세포의 발생단계에 따라 상이하게 나타났다. 또한, HERV-K 유전자들 (HERV-K10 LTR, HERV-K(T47D) LTR, HERV-K(C4) *env*, and HERV-K10 *env*)과 L1Hs *p40* 유전자, 조절 단백질 유전자들 (*BAX*, *c-MYC*, *IL-6*, and *p53*)의 발현 양상은 처리된 발암물질에 따라 변화하였

다. 다양한 세포에서 HERV-K 유전자의 활성이 세포유형 특이적으로 보고 되었기에, HERV-K 유전자들이 세포의 증식, 사멸, 분화, 암 발생 및 외생성레트로바이러스의 유입에 대항하는 항바이러스 방어 전략 등에 관여함을 암시한다. L1Hs의 전사는 세포유형적 특이성을 보이지는 않았으나, 스트레스나 암 발생 인자에 대한 반응에서는 다양하게 나타났다. 비록, 역전위인자 RNA 전사체가 연구에 사용한 모든 세포주에서 발견되었지만, 역전위인자의 발현과 인간 암 사이에서 직접적인 관련성에 대한 증거를 찾기에는 불충분하였다. 이상의 결과는 발암물질에 대한 반응에서 HERV-K와 L1Hs *p40* 유전자 발현의 변화를 근거하여, 역전위인자의 활성 수준의 변화가 세포의 형질전환에 기여함을 암시하고 있다. 몇몇 발암물질은 전위활성을 유도하였다. 트랜스포존 특이적 ITR과 인위적인 무작위 ITR을 증폭 시발체로 사용하여 유전체 수준의 증폭을 수행한 결과, 증폭산물은 변화된 양상을 보였다. 평균 5세대 배양 후, 증폭산물 중에서 구별되는 절편들이 관찰되었다. 염기서열 분석에 수반된 BLAST 검색, 계통유전학적 분석을 통해 역전위인자의 전사체들이 거의 모든 염색체의 다양한 좌위로부터 전사됨을 보여주었다. HERV-K10 LTR 인자의 전사체들은 21번 염색체와 Y 염색체에서는 나타나지 않았으며, 전사좌위는 염색체의 동원체 부근부터 말단소립에 이르기까지 전 영역에서 걸쳐 나타났다. HERV-K10 LTR은 포유류 종분화 이전과 이후에 수차례에 걸쳐 인간 유전체로 유입된 것으로 보인다. L1Hs 전사체들은 21번 염색체를 제외한 모든 염색체에서 확인되었다. 유사도 검색과 DNA 서열에 대한 가상적인 단백질 서열 해독은 인간 반수체 유전체 내에 적어도 140 개의 전사활성 좌위가 확인되었다. 또한 절대성 원칙, 단순통합분지도 (majority rule, strict consensus tree)에서 활성 L1Hs 인자들이 두 집단 (cluster)을 형성하였다. 염색체상의 위치를 살펴보면, 역전위인자 전사체들이 다수의 염색체취약부위 (fragile site)로부터 전사되며, 이와 같은 사실은 역전위인자의 염색체상 위치와 염색체 불안정성 (chromosomal instability)과 암관련 절단부위 (cancer-related breakpoint), 유전병관련 좌위 사이의 관련을 암시하고 있다. 역전위인자들이 인간유전체로 유입된 이후, 유전체의 인트론을 확장시키고, 많은 유사

유전자 (pseudogene)들을 생성하였다. 결론적으로, 본 연구는 많은 역전위인자들이 숙주 유전체 내에서 전사, 전위되고 있을 뿐만 아니라, 세포내외 환경으로부터 다양한 자극에 대해서도 반응할 수 있음을 제안하고 있다. 또한, 활성형 HERV-K와 L1Hs 역전위인자들은 인간 유전체에서 자신의 전위 활성이 잠재성을 유지한 채 광범위하게 분포하고 있다.

주요어: 돌연변이, 레트로포손, 레트로바이러스, 발암물질, 암, 역전위, 역전위인자, 염색체, 염색체 불안정성, 염색체 재배열, 유전병, 유전자 발현, 유전체, 유전체 불안정성, 전사, 전위인자, 전위, 재조합, 취약부위, 트랜스포존, env, HERV-K, HERV-K10 LTR, LINE-1, L1Hs p40, LTR

감사의 글

하루하루 가던 시간속에 어느덧 홀쩍 오늘 이 자리까지 오게 되었습니다. 그리고 지난 4년 동안의 박사학위연구가 이 논문으로 결실을 맺게 되었습니다. 어찌보면 가장 중요하면서도 힘들었던 시기에 곁에서 또는 멀리서 격려와 도움을 주신 모든 분들께 짤막하나마 감사 인사를 드립니다.

학문 연구와 논문 작성에 이르기까지 모든 면에서 부족하기만 했던 저를 위해 끊이 없는 관심과 사랑으로 보살펴주셨고, 학문에 임하는 마음가짐을 흩어지지 않게 인도해주신 오문유 교수님께 진심으로 감사의 말씀을 드립니다. 지난 6년, 대학원 석사, 박사 과정을 거치는 동안 항상 학문에 정진할 수 있게 가르쳐주셨고, 언제나 강직하고 올곧은 학자의 상을 보여주셨기에, 아직도 많이 모자라지만 오늘 이 자리에 설 수 있게 되었습니다. 감사하다는 말로는 이루 다 보답할 수 없는 은혜를 입고 떠나는 제자를 아껴주시니, 은혜에 감사하고 항상 올바른 학문탐구에 매진할 것을 다짐하며, 다시 한 번 머리숙여 감사드리며 항상 건강하시길 기원합니다.

부족한 제 논문에 대한 심사를 맡아 바쁘신 와중에도 세세히 다듬어 주시고, 더불어 더 나은 연구자의 길을 보여주신 제주대 생명과학과 김세재 교수님, 의학과 강희경 교수님, 제주교대 홍승호 교수님, 제주하이테크산업진흥원 김기옥 박사님께도 깊은 감사의 말씀을 드립니다. 그리고 생물학에 관한 많은 가르침을 주신 김원택 교수님, 이용필 교수님, 오덕철 교수님, 이화자 교수님, 고석찬 교수님께도 감사의 말씀을 드립니다. 석사논문 연구때부터 많은 가르침을 주셨던 김세재 교수님께 재삼 감사드립니다.

절부지갈던 시절 대학원 생활의 두려움을 없애 주었고, 학문이라는 험한 길을 걸어갈 힘을 더해 준 고미희 박사님, 정용환 박사님께 감사드립니다. 실험과정에서 같이 밤새며 부푼 라면으로 허기를 달래도 불평 한 번 않던 유전학실험실 가족들. 오유성 선생님과 오주형 선생님 고맙습니다. 김재환, 송지훈 대학원 후배들, 이제 대학원생이 될강민철, 오대주 후배들 너무 고맙고, 황준호, 강성일, 김윤정, 김동철 후배들도 너무너무 고맙다. 더불어 그동안 유전학실험실을 거쳐간 선후배들에게도 감사의 뜻을 전합니다. 특히, 대학원 입학에 대한 용기를 일깨워주었고, 때로 형보다도 더 친근한 모습으로 다독이며 앞서간 어느 자리에서도 후배에게 격려와 함께 많은 도움을 준 정용환 박사님께 진심으로 고마움을 전합니다.

지난 6년간의 대학원 생활에 많은 도움을 준 분자생물학실험실 박지권 박사님, 강

신해 박사님, 정형복, 진영준, 최진영 대학원생들, 미생물학실험실 이동헌 박사님, 이정배 선배님, 생태학실험실 송관필선배님, 문명옥 선배님과 생물학과 모든 대학원생들에게 감사드립니다.
의과대학 해부학교실 이봉희 교수님, 김명주 교수님, 조직학교실 이영기 교수님, 박덕배 교수님, 병리학교실 강현욱 교수님, 생화학교실 조문제 교수님, 기생충학교실 정영배 교수님께도 감사드립니다. 특히, 실험을 위해 많은 도움을 주신 약리학교실 강희경 교수님, 유은숙 교수님, 박수영 박사님, 이혜자, 김상철, 현재희, 윤원종 대학원생, 미생물학교실 고영상 교수님과 윤지현 후배에게도 진심으로 감사드립니다. 임희경, 김영미, 강윤석, 이정희 선생에게도 감사의 뜻을 전합니다.
생물학에 대한 제 관심과 연구에 대한 의욕을 키워낼 수 있었던 사범대학 과학교육과 생물교육 박행신 교수님, 정충덕 교수님, 오홍식 교수님, 이순동 교수님께도 감사드립니다. 제주도 자연사박물관 정세호 박사님, 김완병 선배님께도 감사드립니다. 농업시험장 김성철 박사님께도 감사드립니다. 그리고 쉽지 않은 연구 소재에 애타고 있을 때, 멀리서 힘을 실어준 부산대학교 김희수 교수님, 이주미 대학원생에게도 진심으로 감사드리며, 격려해주신 해양과학대학 이제희 교수님께도 감사드립니다.
고등학교를 졸업하고 지금껏 못난 친구 둔 덕에 마음고생 많으면서도, 어느날 불쑥 나타나도 반겨주던 친구 최종철에게 감사드립니다. 짧지만은 않았던 1년여의 시간 동안 내 곁을 지켜주며 때로 나보다 더 몸과 마음을 걱정해 주며, 웃음과 기쁨을 건네준 고윤미에게도 진심으로 감사의 뜻을 전합니다.
바쁜 생활 속에서도 언제나 동생에 대한 관심을 늦추지 않았던 하나뿐인 누님과 넉넉한 매형, 웃음이 예쁜 귀염덩어리 오하나양에게도 감사드립니다.

언젠가 어린시절, "대학원 가겠다"는 말을 들으시고는 많이 놀라시고도 그러마 하셨고, 없는 형편에 홀로 한숨과 눈물로 살아오시면서도 너무나 당당한 모습으로, 오늘의 저가 있기까지 밤낮으로 비바람 가리지 않고 어려움 막아주시고 보살펴 주신 사랑하는 나의 어머님. 지난 30년간 드릴 것 하나없던 제가 오늘 비로소 어머님께 감사하다는 말씀과 함께 미약하나마 이 논문을 드릴 수 있어 더없이 기쁩니다.
부디 오래오래 건강하십시오.