Korean J Biol Sci 6: 167-170, 2002

Identification and Phylogeny of the Human Endogenous Retrovirus HERV-W LTR Family in Cancer Cells

Joo-Mi Yi, Hwan-Mook Kim[1], and Heui-Soo Kim*

Division of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Korea;
[1]Laboratory of Biopotency Evaluation, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-600, Korea.

Key Words:
Cancer cells
HERV-W
LTR elements
Phylogeny

The long terminal repeats (LTRs) of human endogenous retrovirus (HERV) have been found to be coexpressed with sequences of closely located genes. It has been suggested that the LTR elements have contributed to the structural change or genetic variation of human genome connected to various diseases and evolution. We examined the HERV-W LTR elements in various cancer cells (2F7, A431, A549, HepG2, MIA-PaCa-2, PC-3, RT4, SiHa, U-937, and UO-31). Using genomic DNA from the cancer cells, we performed PCR amplification and identified twelve new HERV-W LTR elements. Those LTR elements showed a high degree of sequence similarity (88-99%) with HERV-W LTR (AF072500). A phylogenetic tree obtained by the neighbor-joining method revealed that HERV-W LTR elements could be mainly divided into two groups through evolutionary divergence. Three HERV-W LTR elements (RT4-2, A431-1, and UO31-2) belonged to Group I, whereas nine LTR elements (2F7-2, A549-1, A549-3, HepG2-3, MP2-2, PC3-1, SiHa-8, SiHa-10, and U937-1) belonged to Group II. Taken together, our new sequence data of the HERV-W LTR elements may contribute to an understanding of tissue-specific cancer by genomic instability of LTR integration.

In humans and primates, endogenous retroviruses and related elements are an integral portion of the genome. Approximately 1% of the human genome is represented by human endogenous retroviruses (HERVs) that are footprints of ancient germ-cell retroviral infections (Sverdlov, 1998). Full length retroviral sequences may interact with cellular oncogenes (Varmus, 1982) and retroviral long terminal repeat (LTR) sequences have the capacity to exert a regulatory influence as promoters and enhancers of cellular genes. Most HERV families encompass a relatively low copy number per haploid genome (La Mantia et al. 1991), compared with others that are either high copy number or single copy retroviral elements (O'Connell et al., 1984). These different copy numbers could represent either multiple integration events or provirus amplified after the integration by retrotransposition. Comparative analysis of the HERV LTR elements in human genome could help us to understand the possible impact of HERVs on evolution and genome regulation.

Retroviral particles have been recovered from monocyte cultures from patients with multiple sclerosis (Perron et al. 1997) and virion-associated MSRV (multiple sclerosis associated retrovirus)-RNA has been reported in serum of patients with the disease (Garson et al. 1998). Expression of MSRV sequences in normal placenta allowed the reconstruction of a 7.6kb putative genomic retroviral RNA with RU5-gag-pol-env-U3R organization, with a polypurine binding site (PBS) showing similarity with avian retrovirus PBS used by tRNATrp (Blond et al. 1999). Southern blot hybridization using MSRV probes allowed characterization of a copy MSRV-related human endogenous retrovirus family named HERV-W (Blond et al. 1999). The expression, structure and promoter activity of HERV-W LTR elements were examined in human cell lines (Schon et al. 2001). Here we identified the HERV-W LTR family in the various cancer cells and phylogenetically analyzed them with those sequences derived from GenBank database.

* To whom correspondence should be addressed.
Tel: 82-51-510-2259, Fax: 82-51-581-2962
E-mail: khs307@pusan.ac.kr

Materials and Methods

Cell culture

Human cancer cells (2F7, A431, A549, HepG2, MIA-PaCa-2, PC-3, RT4, SiHa, U-937, and UO-31) were grown in Dulbeccos medium (DMEM) supplemented with 10% heat inactivated fetal calf serum, 2 mM glutamine, 1 mM non-essential amino acids, 1 mM sodium pyruvate, 100 U/mL penicillin, and 0.1 mg/mL strepto-

Table 1. HERV-W LTR elements identified from human cancer cells

Cancer Cell Lines	Clones
2F7 (B cell lymphoma)	2F7-2
A431 (epidermoid carcinoma)	A431-1
A549 (lung carcinoma)	A549-1, A549-3
HepG2 (liver hepatoblastoma)	HepG2-3
MIA-PaCa-2 (pancreas carcinoma)	MP2-2
PC-3 (prostate adenocarcinoma)	PC3-1
RT4 (urinary carcinoma)	RT4-2
SiHa (cervix squamous carcinoma)	SiHa-8, SiHa-10
U-937 (histiocytic carcinoma)	U937-1
UO-31 (kidney carcinoma)	UO31-2

mycin at 37℃ and 5% CO_2 incubator.

PCR amplification for the HERV-W LTR elements

New 416-bp LTR elements of HERV-W family were amplified by the primer pair HS47 (5′-TGGTCCATG-TTTCTTACGGCT-3′, bases 127-147) and DS16 (5′-AAGATGGTGGTGAACCACTTC-3′, bases 521-541) from the HERV-W (GenBank, accession no. AF072500). The PCR conditions followed those of Kim et al. (1996) with an annealing temperature of 56℃.

Molecular cloning of the PCR products

PCR products were separated on a 1.8% agarose gel, purified with the QIAEX II gel extraction kit (Qiagen) and cloned into the T-khs307 vector (Kim et al., 1998). The cloned DNA was isolated by the alkali lysis method using the High Pure plasmid isolation kit (Roche). Individual plasmid DNAs were screened for inserts by PCR.

Determination of DNA sequence and data analysis

Positive samples were subjected to sequence analyses on both strands with T7 and M13 reverse primers using an automated DNA sequencer (Model 373A) and the DyeDeoxy terminator kit (Applied Biosystem). Nucleotide sequence analyses were performed using GAP, PILEUP, and PRETTY from the GCG package (Genetics Computer Group, University of Wisconsin). Neighbor-joining phylogenetic analysis was performed with the MEGA program (Kumar et al., 1993). Nucleotide sequences of HERV-W LTR elements were retrieved from the GenBank database with the aid of BLAST network server (Altschul et al., 1997).

Nucleotide sequence accession numbers

The nucleotide sequence data reported in this paper will appear in the DDBJ/EMBL/GenBank nucleotide sequence database with the accession numbers AB055868 (2F7-1), AB055869 (A431-1), AB055870 (A549-1), AB055871 (A549-3), AB055872 (HepG2-3), AB055873 (MP2-2), AB055874 (PC3-1), AB055875 (RT4-2), AB055876 (SiHa-8), AB055877 (SiHa-10), AB055878 (U937-1), and AB055879 (UO31-2).

Results and Discussion

Using genomic DNA from the various cancer cells (2F7, A431, A549, HepG2, MIA-PaCa-2, PC-3, RT4, SiHa, U-937, and UO-31), we performed PCR amplification and identified twelve new HERV-W LTR elements (Table 1). Those LTR elements showed a high degree of sequence similarity (88.2-99.7%) with that of the HERV-W LTR (AF072500) (Table 2). To understand the evolutionary relationship among the HERV-W LTR elements, we retrieved the LTRs from the GenBank database and analyzed them with the new HERV-W LTR elements. A phylogenetic tree obtained by the neighbor-joining method revealed that HERV-W LTR elements could be mainly divided into two groups through evolutionary divergence. Three HERV-W LTR elements (RT4-2, A431-1, and UO31-2) isolated from RT4, A431, and UO-31 cells belonged to Group I, whereas nine LTR elements (A549-1, A549-3, U937-1, SiHa-8, SiHa-10, HepG2-3, 2F7-2, MP2-2, and PC3-1) isolated from A549, U-937, SiHa, HepG2, 2F7, MIA-PaCa-2, PC-3 belonged to Group II. Most HERV-W LTR elements retrieved from the GenBank belonged to Group I. Recently, several copy numbers of the HERV-W LTR elements were isolated from the human mammary carcinoma cell line T47D (Schon et al., 2001). They were also clearly divided into the two groups. Expression patterns of the HERV LTR elements varied in various cell lines (epidermal keratinocytes, liver cells, kidney cells, pancreatic cells, lymphocytes, and lung fibroblasts), in some cases showing strict cell

Table 2. Percentage similarity of nucleotide sequences of the HERV-W LTR family

HERV-W LTR family	1	2	3	4	5	6	7	8	9	10	11	12	13
1. HERV-W LTR	–												
2. 2F7-2	99.7	–											
3. A431-1	88.2	88.5	–										
4. A549-1	97.9	98.1	87.1	–									
5. A549-3	99.2	99.5	88.5	98.1	–								
6. HepG2-3	99.2	99.5	88.5	98.1	99.5	–							
7. MP2-2	98.9	99.2	88.7	97.9	99.2	99.2	–						
8. PC3-1	99.2	99.5	89.0	98.1	99.5	99.5	99.2	–					
9. RT4-2	88.2	88.5	92.0	87.7	88.5	88.5	88.7	89.0	–				
10. SiHa-8	99.5	99.7	88.7	98.4	99.7	99.7	99.5	99.7	88.7	–			
11. SiHa-10	98.9	99.2	88.2	97.9	99.2	99.2	98.9	99.2	88.2	99.5	–		
12. U937-1	98.4	98.7	87.9	96.3	98.7	98.7	98.4	98.7	88.2	98.9	98.4	–	
13. UO31-2	88.2	88.5	99.5	87.1	88.5	88.5	88.7	89.0	91.4	88.7	88.2	87.9	–

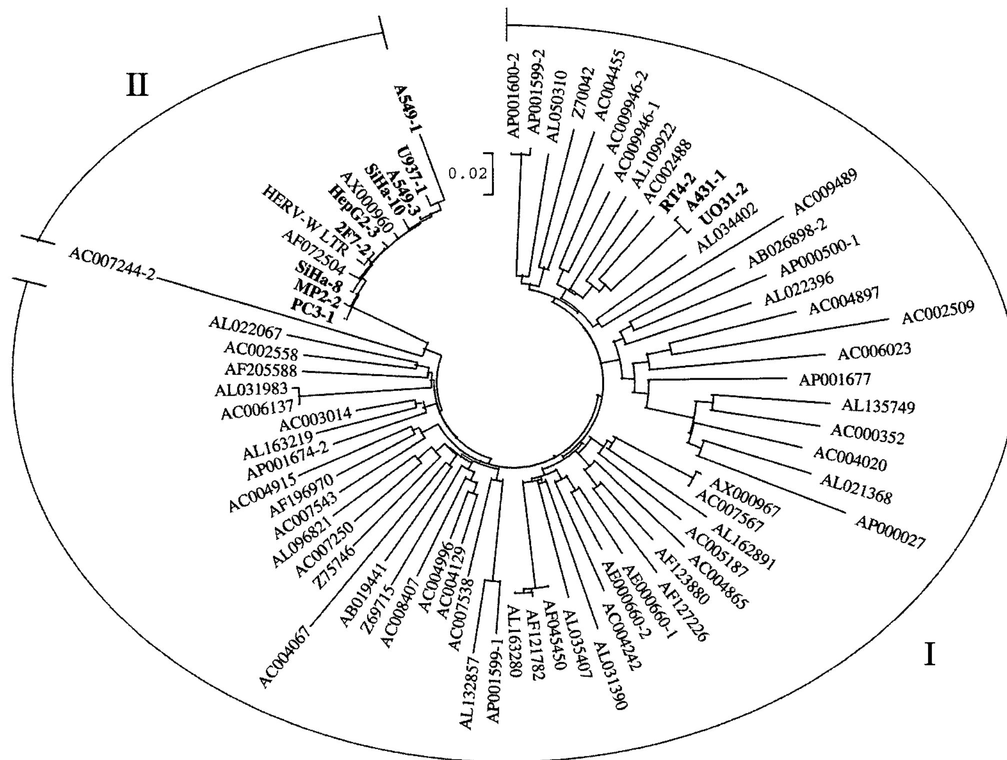

Fig. 1. Phylogenetic tree obtained by neighbor-joining method for the LTR elements of the HERV-W family in cancer cells. Branch lengths are proportional to the distances between the taxa.

type specificity (Schon et al., 2001). The HERV LTR elements could be useful for obtaining tissue-specific promoters. Akopov et al. (1998) have noted that such sequences have the capacity to modify the expression of neighboring genes, and suggested that such modifications may have been acquired in the course of human evolution. The HERV-K-T47D-related LTR element has been shown to mediate polyadenylation of cellular transcripts (Baust et al., 2000). Such phenomenon was demonstrated recently in nucleosomal binding protein NSBP1 in Xq13.3 (King and Francomano, 2001). In the case of another retro-element (the HERV-F LTR element), a similar phenomenon was observed in relation to the Krüppel-related zinc finger gene ZNF195 (Kjellman et al. 1999). A solitary HERV-K LTR element in the HLA DQ region (DQ- LTR3) resulted in suffering from the type I diabetes mellitus in 246 German and Belgian families (Donner et al. 1999b). The retroviral LTR element (DQLTR3) was human-specific insertion (Donner et al. 1999a). This type of the retroviral elements also induced alternative splicing in the human leptin receptor (Kapitonov and Jurka, 1999). The solitary HERV LTR elements showed that they retained detectable activity in human carcinoma cells, and could direct the transcription in both orientations relative to the reporter gene (Domansky et al. 2000). Medstrand et al. (2001) reported that LTR elements were used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans. In our previous study, the HERV-K LTR elements were expressed in human brain tissues, suggesting that they could play a pathogenic role in brain diseases such as psychosis (Kim and Lee 2001). Taken together, various copy numbers of the HERV-W LTR elements are detectable in cancer cells, although this phenomenon may not directly lead to human cancers. Our new sequence data of the HERV-W LTR elements in cancer cells may contribute to an understanding of tissue-specific cancer by promoter activity and genomic instability of LTR integration.

References

Akopov SB, Nikolaev LG, Khil PP, Lebedev YB, and Sverdlov ED (1998) Long terminal repeats of human endogenous retrovirus K family (HERV-K) specifically bind host cell nuclear proteins. *FEBS Lett* 421: 229-233.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, and Lipman J (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.

Baust C, Seifarth W, Germaier H, Hehlmann R, and Leib-Mosch C (2000) HERV-K-T47D-related long terminal repeats mediate polyadenylation of cellular transcripts. *Genomics* 66: 98-103.

Blond JL, Beseme F, Duret L, Bouton O, Bedin F, Perron H, Mandran B, and Mallet F (1999) Molecular characterization

and placental expression of HERV-W, a new human endogenous retrovirus family. *J Virol* 73: 1175-1185.
Domansky AN, Kopantzev EP, Snezhkov EV, Lebedev YB, Leib-Mosch C, and Sverdlov ED (2000) Solitary HERV-K LTR possess bi-directional promoter activity and contain a negative regulatory element in the U5 region. *FEBS Lett* 472: 191-195.
Donner H, Tonjes RR, Bontrop RE, Kurth R, Usadel KH, and Badenhoop K (1999a) Intronic sequence motifs of HLA-DQB1 are shared between humans, apes and Old World monkeys, but a retroviral LTR element (DQLTR3) is human specific. *Tissue Antigens* 53: 551-558.
Donner H, Tonjes RR, Van der Auwera B, Siegmund T, Braun J, Weets I, Herwig J, Kurth R, Usadel KH, and Badenhoop K (1999b) The presence or absence of a retroviral long terminal repeat influences the genetic risk for type I diabetes conferred by human leukocyte antigen DQ haplotypes. *J Clin Endocrinol Metab* 84: 1404-1408.
Garson JA, Tuke PW, Giraud P, Paranhos-Baccaia G, and Perron P (1998) Detection of virione-associated MSRV-RNA in serum of patients with multiple sclerosis. *Lancet* 351: 33-33.
Kapitonov VV and Jurka J (1999) The long terminal repeat of an endogenous retrovirus induces alternative splicing and encodes an additional carboxy-terminal sequence in the human leptin receptor. *J Mol Evol* 48: 248-251.
Kim HS, Chen Y, and Lonai P (1998) Complex regulation of multiple cytohesin like genes in murine tissues and cells. *FEBS Lett* 433: 312-316.
Kim HS, Hirai H, and Takenaka O (1996) Molecular features of the TSPY gene of gibbons and Old World monkeys. *Chrom Res* 4: 500-506.
Kim HS and Lee YC (2001) Identification and phylogenetic analysis of long terminal repeat elements of the human endogenous retrovirus K family (HERV-K) from a human brain cDNA library. *Korean J Biol Sci* 5: 133-137.
King LM and Francomano CA (2001) Characterization of a human gene encoding nucleosomal binding protein NSBP1. *Genomics* 71: 163-173.
Kjellman C, Sjogren H-O, Salford L, and Widegren B (1999) HERV-F (XA34) is a full-length human endogenous retrovirus expressed in placental and fetal tissues. *Gene* 239: 99-107.
Kumar S, Tamura K, and Nei N (1993) MEGA: Molecular Evolutionary Genetics Analysis, Version 1.01. The Pennsylvania State University, University Park.
La Mantia G, Maglione D, Pengue G, Di Cristofano A, Simeone A, Lanfrancone L, and Lania L (1991) Identification and characterization of a novel human endogenous retroviral sequences preferentially expressed in undifferential embryonal carcinoma cells. *Nucleic Acids Res* 19: 1513-1520.
Medstrand P, Landry JR, and Mager DL (2001) Long terminal repeats are used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans. *J Biochem* 276: 1896-1903.
O'Connell C, O'Brien S, Nash WG, and Cohen M (1984) ERV3, a full-length human endogenous provirus: chromosomal localization and evolutionary relationships. *Virology* 138: 225-235.
Perron H, Garson JA, Bedin F, Beseme F, Paranhos-Baccala G, Komurian-Pradel F, Mallet F, Tuke PW, Voisset C, Blond JL, Lalande B, Seigneurin JM, Mandrand B, and the collaborative research group on MS (1997) Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. *Proc Natl Acad Sci USA* 94: 7583-7588.
Schon U, Seifarth W, Baust C, Hohenadl C, Erfle V, and Leib-Mosch C (2001) Cell type-specific expression and promoter activity of human endogenous retroviral long terminal repeats. *Virology* 279: 280-291.
Sverdlov ED (1998) Perpetually mobile footprints of ancient infections in the human genome. *FEBS Lett* 428: 1-6.
Varmus HE (1982) Form and function of retroviral proviruses. *Science* 216: 812-820.

[Received April 8, 2002; accepted May 6, 2002]