Korean J Biol Sci 5: 243-246, 2001

Identification and Phylogenetic Analysis of the Human Endogenous Retrovirus HERV-W LTR Family in Placenta cDNA Library

Joo-Mi Yi, Ji-Won Lee, Kyung-Mi Shin, Jae-Won Huh, Won-Ho Lee, Kyung-Lib Jang, and Heui-Soo Kim*

Division of Biological Sciences, College of Natural Sciences, Pusan National University, Pusan 609-735, Korea

Key Words:
HERV-W
LTR elements
Placenta cDNA library
Phylogeny

Human endogenous retroviral long terminal repeats (LTRs) have been found to be coexpressed with sequences of genes closely located nearby. It has been suggested that the LTR elements have contributed to structural changes or genetic variations of human genome connected to various diseases and evolution. Using cDNA library derived from placenta tissue, we performed PCR amplification and identified five new HERV-W LTR elements. Those LTR elements showed a high degree of sequence similarity (98-99%) 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. Five new HERV-W LTR elements (pla-1, 4, 5, 6, 7) belonged to the group I with AX000960, AF072504, and AF072506 from GenBank database. The data suggest that several copy numbers of the HERV-W LTR elements are transcribed in placenta and may contribute to the understanding of biological function such as human placental morphogenesis.

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 regulatory influences 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 7.6 kb 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). We examined HERV-W *pol* and *env* gene sequences in human monochromosomes, and found multiple frameshift and termination codons by deletion/insertion or point mutation (Kim 2001a; Kim and Lee, 2001b). The HERV-W LTR elements were detected in hominoids, Old and New World monkeys, suggesting that they have inserted in the primate genome approximately 55 million years ago (Kim 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 placenta cDNA library and phylogenetically analyzed 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@hyowon.cc.pusan.ac.kr

Materials and Methods

PCR amplification for HERV-W LTR elements

The cDNA synthesized from mRNA of human placenta (Clontech) was used as a template for PCR amplification. New 416-bp LTR elements of HERV-W family

were amplified by the primer pair HS47 (5′-TGG-TCCATGTTTCTTACGGCT-3′, bases 127-147) and DS16 (5′-AAGATGGTGGTGAACCACTTC-3′, bases 521-541) from the HERV-W (GenBank, accession no. AF072500). The PCR condition was performed as follows. After the initial denaturation step at 94℃ for 3 min, DNA was amplified for 30 cycles at 94℃ for 1 min, 56℃ for 1 min and 72℃ for 1.5 min.

Molecular cloning of PCR products

PCR products were separated on 2% agarose gel, purified using QIAEX II gel extraction kit (Qiagen) and cloned into the T-khs307 vector (Kim et al., 1998). The cloned DNA was transformed and replicated in *Escherichia coli* K-12 strain, JM109. Plasmid DNA was isolated by the alkali lysis method using High Pure plasmid isolation kit (Roche). Individual plasmid DNAs were screened for inserts by PCR using the original primers designed for the locus.

DNA sequencing 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 DyeDeoxy terminator kit (Applied Biosystem). Nucleotide sequence analysis was performed using GAP and PILEUP programs of GCG software (Genetics Computer Group, University of Wisconsin). The neighbor-joining phylogenetic analysis was performed with MEGA program (Kumar et al., 1993). Nucleotide sequences of HERV-W LTR elements were retrieved from GenBank database with the aid of BLAST network server (Altschul et al., 1997).

Results and Discussion

Using cDNA library derived from the human placenta, we performed PCR amplification (Fig. 1) and identified five new HERV-W LTR elements. Those LTR elements showed a high degree of sequence similarity (98-99%) with that of HERV-W LTR (AF072500) (Table 1). In order to understand the evolutionary relationship among HERV-W LTR elements, we retrieved the LTRs from the GenBank database and analyzed them with new HERV-W LTR elements. A phylogenetic tree obtained by the neighbor-joining method revealed that the HERV-W LTR elements could be mainly divided into two groups by sequence variation through evolutionary divergence (Fig. 2). Five new HERV-W LTR elements (pla-1, 4, 5, 6, 7) belonged to the group I including GenBank accession no. AX000960, AF072504, and AF072506, and they were also aligned with the HERV-W LTR element (Fig. 3). One or two bp deletions were noted in AC007244-2 sequences. Most HERV-W LTR elements retrieved from the GenBank belonged to the group II. Recently, several copy numbers of the HERV-W

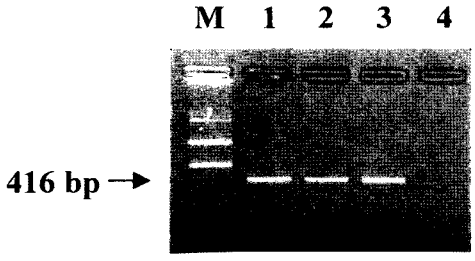

Fig. 1. PCR analysis of placenta cDNA library for the presence of the HERV-W LTR elements. Lane M, marker (pUC18/*Taq I*); 1, human male genomic DNA (positive control); 2, human female genomic DNA (positive control); 3, placenta cDNA; 4, ddH_2O (negative control).

Table 1. Percentage similarity of nucleotide sequences of HERV-W LTRs

	1	2	3	4	5	6	7	8	9	10
1. W-LTR	–									
2. AX000960	98.4	–								
3. AF072504	98.9	98.7	–							
4. AF072506	99.5	98.9	99.5	–						
5. AC007244-2	83.9	83.6	84.1	84.4	–					
6. **pla-1**	98.7	98.1	98.7	99.2	83.6	–				
7. **pla-4**	98.4	97.9	98.4	98.9	83.6	98.1	–			
8. **pla-5**	98.7	98.1	98.7	99.2	83.6	98.4	98.7	–		
9. **pla-6**	99.2	98.7	99.2	99.7	84.1	98.9	99.2	99.5	–	
10. **pla-7**	99.2	98.7	99.2	99.7	84.1	98.9	98.7	98.9	98.5	–

LTR elements were isolated from the human mammary carcinoma cell line T47D (Schon et al., 2001). They also clearly divided into the two groups. In an analysis of promoter activity, the W8 - LTR element showed highest promoter activity in LC5 cells, while the W23 - LTR element did not show the activity in any cell lines. 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 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 mediated polyadenylation of cellular transcripts (Baust et al., 2000). Such phenomenon was very recently demonstrated in nucleosomal binding protein NSBP1 in Xq13.3 (King and Francomano, 2001). In case of another retro-element (the HERV-F LTR element), a similar phenomenon was observed in relation to the Krppel-related zinc finger gene ZNF195 (Kjellman et al., 1999). A solitary HERV-K LTR element in the HLA DQ region (DQ-LTR3) resulted in type I diabetes mellitus in 246 German and Belgian families (Donner et al., 1999a). The retroviral LTR element (DQLTR3) showed a human-specific insertion (Donner et al., 1999b). This type of retroviral elements also induced alternative splicing in human leptin receptor

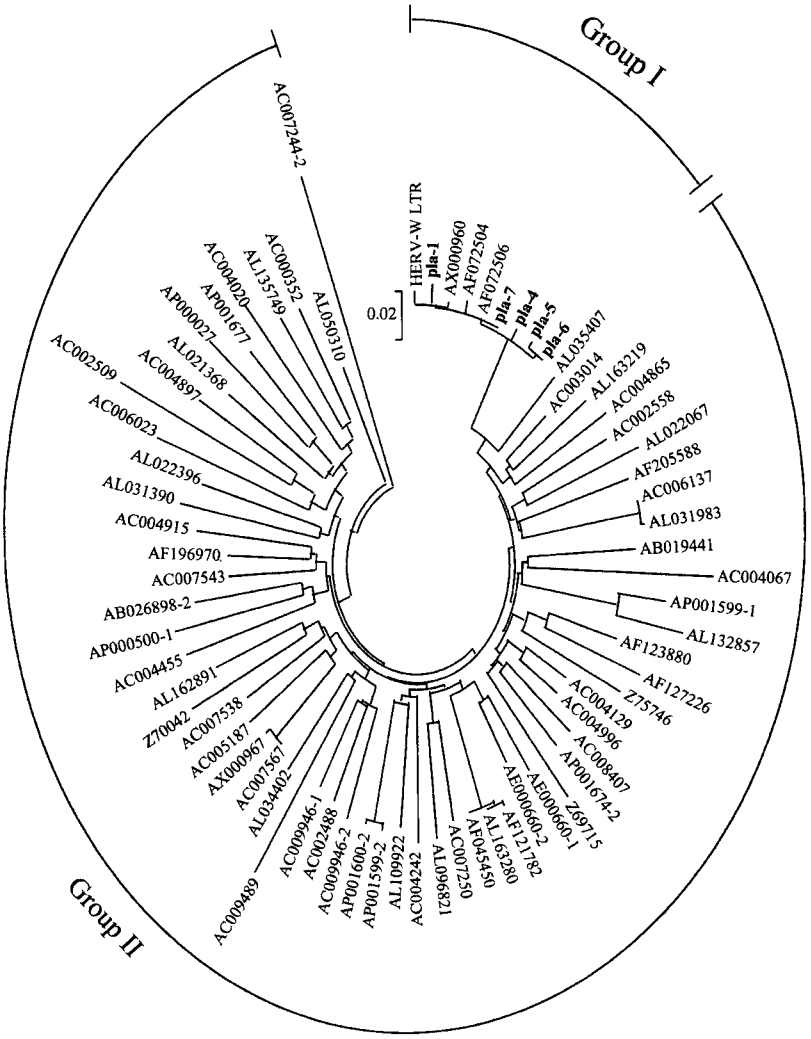

Fig. 2. Phylogenetic tree obtained by neighbor-joining method for the LTR elements of the HERV-W family in placenta cDNA library. Branch lengths are proportional to the distances between the taxa. The cDNA sequence data reported in this paper will appear in the DDBJ/EMBL/ GenBank nucleotide sequence databases with the accession numbers AB066201 (pla-1), AB066202 (pla-4), AB066203 (pla-5), AB066204 (pla-6), and AB066205 (pla-7).

(Kapitonov and Jurka, 1999). The solitary HERV LTR elements showed that they retained detectable activity in human carcinoma cells, and could direct 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 human. In this report, our new sequence data of the HERV-W LTR elements may contribute to the understanding of biological function (placental morphogenesis) connected to human diseases in placenta.

Acknowledgement

This work was supported by Korea Research Foundation Grant (KRF-2000-015-DP0342).

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, Schffer 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.

Donner H, Tonjes RR, Auwera B, Siegmund T, Braun J, Weets IS, Registry BD, Herwig J, Kurth R, Usadel KH, and Baden-

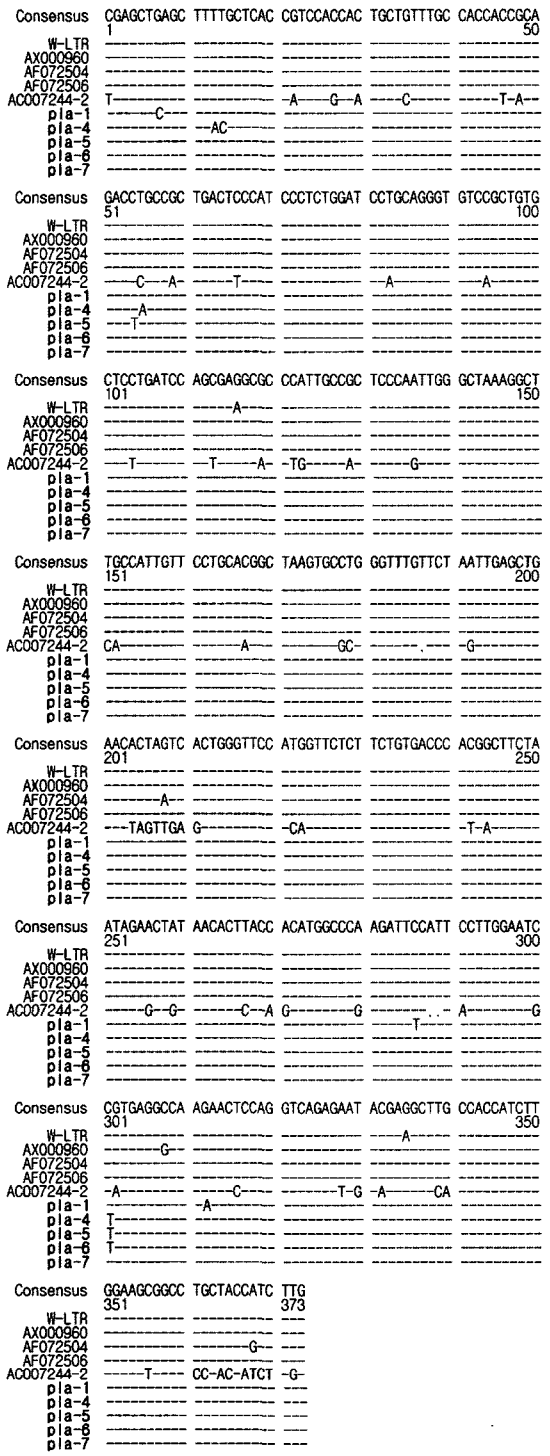

Fig. 3. Sequence alignments of the HERV-W LTR elements belonged to the Group I in phylogenetic tree. Consensus sequences are shown on the top row. Dashes indicate no change to the consensus sequences. AC007244-2 was used as outgroup.

hoop K (1999a) 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 Endocr Metab* 84: 1404-1408.

Donner H, Tonjes RR, Bontrop RE, Kurth R, Usadel KH, and Badenhoop K (1999b) 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.

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.

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.

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 (2001a) Isolation and phylogeny of HERV-W *pol* fragments. *AIDS Res Hum Retroviruses*: in press.

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 WH (2001b) Human endogenous retrovirus HERV-W family: chromosomal localization, identification, and phylogeny. *AIDS Res Hum Retroviruses* 17: 643-648.

Kim HS, Takenaka O, and Crow TJ (1999) Isolation and phylogeny of endogenous retrovirus sequences belonging to the HERV-W family in primates. *J Gen Virol* 80: 2613-2619.

King LM and Francomano CA (2001) Characterization of a human gene encoding nucleosomal binding protein NSBP1. *Genomics* 71: 163-173.

Kjellman C, Sjogren HO, Salford LC, 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 July 18, 2001; accepted August 20, 2001]