Minireview

Genomic Impact, Chromosomal Distribution and Transcriptional Regulation of HERV Elements

Heui-Soo Kim*

Human endogenous retroviruses (HERVs) mediate structural variation and genomic instability based on their multiple copy number, inherent ability to mobilize via reverse transcriptase, and high sequence similarity. Moreover, they undergo multiple amplification and retrotransposition events, resulting in the widespread distribution of complete or partial retroviral sequences throughout the primate genome. As such, HERV elements have played important biological roles in genome evolution, and their long terminal repeat (LTR) elements contain numerous regulatory sequences, including effective promoters, enhancers, polyadenylation signals, and transcription factor-binding sites. Lastly, HERV elements are capable of influencing the expression of neighboring genes, a process that also contributed to primate evolution.

INTRODUCTION

Human endogenous retroviruses (HERVs) have been subjected to numerous amplification and transposition events, resulting in the widespread distribution of complete or partial retroviral sequences throughout the human genome. Approximately 8% of the human genome is composed of HERVs, which were transmitted in a Mendelian fashion after reverse transcriptase-mediated germ-line infection by ancient active exogenous retroviruses (International Human Genome Sequencing Consortium, 2001; Lower et al., 1996). Most HERV families inserted into the primate genome and underwent amplification on several occasions at the time of the divergence of hominoids from Old World monkeys approximately 30-45 million years ago (Sverdlov, 2000). HERVs are present within full-length or incomplete sequences and contain multiple stop codons, insertions, deletions, and frame shift mutations (de Parseval et al., 2001; Kim, 2001). However, the envelope proteins of some HERV families are expressed preferentially in the human placenta (Kamat et al., 2002; Mi et al., 2000; Venables et al., 1995) as well as several cancer cells (Armbruester et al., 2002; Yi et al., 2004; 2006).

Retroviruses are capable of entering host genomes at many sites, implying a capacity to inactivate genes by physical disruption (Varmus, 1982), and retroviral LTR sequences are able to exert regulatory effects as promoters and enhancers of cellular genes (Domansky et al., 2000; Kovalskaya et al., 2006). Most LTRs have no influence on gene function and are not relevant to disease pathology (Lower et al., 1996). However, a small minority of these sequences does participate in the regulation of gene expression by serving as alternative promoters (Landry et al., 2002; Medstrand et al., 2001; Sin et al., 2006). They also provide polyadenylation signals that contribute to the formation of alternative transcripts (Mager et al., 1999; Sin et al., 2007). Therefore, LTRs could change the regulatory patterns of neighboring genes in evolutionary processes.

HERVs have been proposed to act as etiological cofactors in several chronic diseases, including cancer, autoimmunity, and neurological diseases (Karlsson et al., 2001; Lower, 1999; Perron et al., 2005). HERV elements have also been implicated in other disorders such as male infertility (Kamp et al., 2000). It has also been suggested that they are involved in carcinogenesis based on the expression of HERV mRNA and functional proteins (Sauter et al., 1995). Further, HERVs may be associated with the generation of new promoters or the activation of proto-oncogenes (Huh et al., 2008; 2009; Schulte et al., 1996). Therefore, it is very likely that some HERVs influence gene regulation by expressing their retroviral genes at the transcriptional level, inducing rearrangement events, and providing regulatory sequences. Accordingly, the present review mainly focuses on the chromosomal distribution and transcriptional regulation of HERV elements in primates.

Chromosomal distribution and copy numbers of the HERV family in humans and great apes

HERVs represent footprints of previous retroviral infection and can be separated into three classes according to the level of *pol* sequence homology. Such a classification is based on complementary binding between tRNA and the primer-binding site (PBS) (Larsson et al., 1989). A one-letter amino acid code is used to denote the corresponding HERV families: HERV-K, HERV-R, HERV-W, and HERV-H. The basic genes of HERVs are *gag*, *pro*, *pol*, and *env*. The *gag* gene codes for the structural matrix (MA), capsid (CA), and nucleocapsid (NC) proteins, *pro* codes for protease (PR) protein, *pol* codes for the enzymes of reverse transcriptase (RT) and integrase (IN), and *env* codes for the envelope surface (SU) and transmembrane (TM) proteins (Fig. 1). The human genome has been estimated to contain about 30-50 HERV-K proviruses (Ono, 1986), and some

Department of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Korea
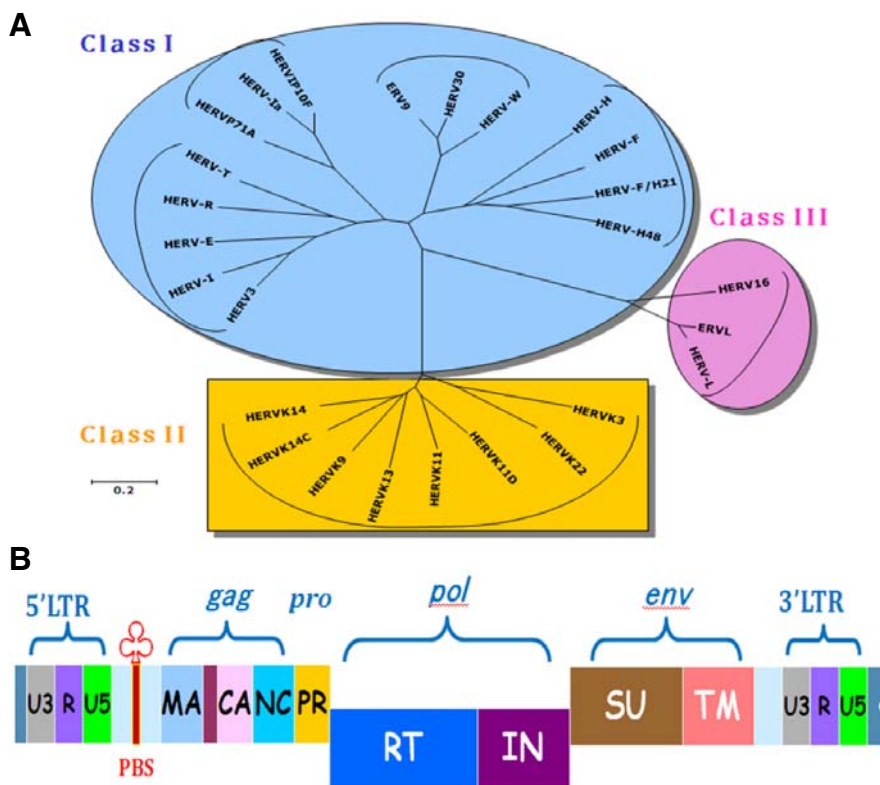*Correspondence: khs307@pusan.ac.kr

Fig. 1. Phylogenetic tree obtained by the neighboring-joining method for *pol* sequence homology in humans. Branch lengths are proportional to the distances between the taxa (A). Genomic structure of HERV possessing the genes *gag*, *pro*, *pol*, and *env* (B). The *gag* gene codes for the structural matrix (MA), capsid (CA), and nucleocapsid (NC) proteins, *pro* codes for protease (PR) protein, *pol* codes for the enzymes of reverse transcriptase (RT) and integrase (IN), and *env* codes for the envelope surface (SU) and transmembrane (TM) proteins. LTR consists of U3, R, and U5 elements. PBS represents primer-binding sites.

[HERV-K101, K102, K103, K104, K106, K107 (K10), K108 (HLM-2), K109] were previously found to be unique to humans, showing full-length open reading frames for retroviral protein precursors (Barbulescu et al., 1999). As an estimate, there are about 1,000 copies of HERV-K, including truncated forms, in total (Ha et al., 2011). The genomes of chimpanzee and orangutan contain 862 and 796 copies of HERV-K, respectively. In addition, in comparison with other HERV families (HERV-H, HERV-R, and HERV-W), the HERV-K family including solitary LTR elements was found to be abundantly present in both humans and great apes using the genome-wide browser for retroelements (GEBRET) database (Table 1). In a previous study, BLAST search of the DDBJ/EMBL/GenBank databases identified 140 sequences of HERV-W family viruses representing 39 HERV-W proviruses, 40 full-length HERV-W viruses, and 61 truncated HERV-W elements (Costas, 2000). Using the UCSC genome browser and Repbase, the copy number of HERV-W family viruses has been calculated, and 153 HERV-W elements and 209 solitary sequences were reported. According to this study, these are distributed in all chromosomes except for chromosomes 16 and 22, which is consistent with the Southern blot and PCR data (Kim 2001; Voisset et al., 2000). Similar copy numbers have also been detected in the chimpanzee genome by database and fluorescence in situ hybridization (FISH) analyses (Table 1) (Kim et al., 2008a). However, the copy numbers of Alu elements significantly differ between humans and great apes (Locke et al., 2011). This difference could be explained by a retrotransposition event during primate evolution. The decreased copy numbers of Alu elements could indicate a reduction in the reverse transcriptase activity of L1 elements, since both Alu and L1 elements are dependent on the same enzymatic machinery for propagation.

Sequence divergence between 5′LTR and 3′LTR sequences constitutes a viable molecular clock in terms of evolution (Fig. 2). According to the mechanism of reverse transcription, the two LTRs must be identical. However, with time, they could undergo divergence due to various mutations (substitutions, insertions, deletions, rearrangement, and inversion) acquired during cellular DNA replication, thus indicating that divergence between the two LTRs could be used as a molecular clock to understand the evolutionary radiation of the HERV family (Mager and Freeman, 1995). The 5′LTR and 3′LTR sequences of HERV-K in long C4 genes of human, orangutan, and African green monkey display similar sequence divergence values of 9.1-10.5% (Dangel et al., 1995). These data indicate that the HERV family is stabilized in the primate genome during evolutionary radiation, and could serve as a reliable reference point for studies on gene duplication and evolution.

Intraelement homologous recombination between the 5′LTR and 3′LTR of HERVs results in the excision of structural genes (*gag-pol-env*), leaving a solitary LTR element (Mager and Goodchild, 1989). Multiple copy numbers of solitary LTR elements of the HERV-K family (GenBank accession no. AC002350, AC002400, AC002508, L47334, U47924, Z80898, AL034407) have been identified as being unique to humans (Buzdin et al., 2003; 2005; Medstrand and Mager, 1998). Short interspersed element, SINE-R.C2 retroposon derived from the solitary HERV-K LTR is also a human-specific element. Even more, two similar retroposons, HS307 and HS408, were previously identified in the Xq21.3 region, which is linked to schizophrenia and schizoaffective disorder (Kim et al., 1999). These specific insertion sites for HERV elements with relevance to human disease could be useful in resolving phylogeny and evolution among closely related species. In fact, they exhibit insertional polymorphisms in the human population (Hughes and Coffin, 2004; Macfarlane and Simmonds, 2004), indicating that they inserted

Table 1. Chromosomal distribution and copy numbers of HERV family in humans and great apes

| Chr	HERV-H / solitary LTR			HERV-R / solitary LTR			HERV-W / solitary LTR			HERV-K / solitary LTR		
	Hu	Ch	Or	Hu	Ch	Or	Hu	Ch	Or	Hu	Ch	Or
1	25 / 54	22 / 33	21 / 40	6 / 8	9 / 4	6 / 15	13 / 11	10 / 10	12 / 7	88 / 79	72 / 78	60 / 63
2	28 / 46	16 / 27	13 / 26	3 / 10	7 / 11	0 / 8	13 / 24	20 / 12	18 / 12	51 / 67	26 / 52	28 / 44
3	24 / 64	23 / 39	25 / 41	1 / 9	6 / 5	1 / 8	19 / 15	14 / 24	14 / 19	57 / 47	71 / 72	70 / 69
4	20 / 58	8 / 25	12 / 18	0 / 7	2 / 2	3 / 2	15 / 21	5 / 14	4 / 6	81 / 70	29 / 56	35 / 48
5	13 / 29	13 / 33	15 / 26	0 / 2	2 / 4	0 / 4	6 / 15	18 / 9	16 / 13	41 / 54	49 / 56	57 / 55
6	24 / 49	10 / 18	11 / 19	0 / 4	7 / 6	1 / 9	15 / 9	5 / 22	5 / 20	69 / 63	25 / 48	41 / 46
7	14 / 27	8 / 43	8 / 30	2 / 7	0 / 3	0 / 1	7 / 15	6 / 7	7 / 3	37 / 46	61 / 40	51 / 43
8	8 / 44	12 / 11	11 / 9	0 / 1	0 / 2	0 / 1	6 / 5	3 / 8	7 / 5	79 / 44	17 / 41	28 / 39
9	8 / 14	10 / 13	8 / 9	0 / 1	3 / 2	0 / 1	3 / 7	4 / 12	4 / 10	19 / 40	21 / 32	16 / 25
10	16 / 22	9 / 17	12 / 16	1 / 2	3 / 6	1 / 10	5 / 12	6 / 9	6 / 8	33 / 35	39 / 28	46 / 21
11	16 / 23	13 / 19	10 / 18	5 / 7	11 / 3	3 / 2	6 / 6	12 / 10	9 / 16	56 / 32	38 / 35	36 / 31
12	12 / 25	11 / 17	8 / 14	3 / 4	1 / 4	0 / 3	12 / 13	7 / 8	5 / 9	46 / 36	19 / 16	25 / 18
13	9 / 25	14 / 9	9 / 16	0 / 7	4 / 4	0 / 4	4 / 12	5 / 6	6 / 4	16 / 23	23 / 44	32 / 46
14	7 / 21	9 / 19	6 / 13	0 / 6	0 / 4	0 / 5	3 / 9	5 / 6	4 / 11	20 / 34	12 / 23	20 / 26
15	6 / 7	5 / 14	6 / 13	0 / 1	0 / 5	1 / 6	4 / 6	5 / 8	4 / 7	20 / 14	21 / 34	19 / 32
16	9 / 19	3 / 8	3 / 6	1 / 3	0 / 0	0 / 0	0 / 0	4 / 5	2 / 4	19 / 27	17 / 13	23 / 17
17	2 / 7	6 / 9	3 / 3	0 / 4	1 / 5	1 / 6	2 / 3	0 / 0	0 / 0	17 / 20	15 / 26	11 / 16
18	2 / 16	2 / 4	4 / 7	0 / 3	0 / 5	0 / 4	2 / 5	2 / 4	1 / 0	5 / 12	13 / 21	13 / 15
19	10 / 12	5 / 13	4 / 11	5 / 6	1 / 3	0 / 3	2 / 1	2 / 8	4 / 3	84 / 55	6 / 6	10 / 8
20	1 / 14	7 / 6	7 / 5	0 / 0	1 / 4	1 / 3	2 / 1	1 / 1	0 / 1	12 / 6	66 / 53	87 / 42
21	0 / 8	1 / 9	2 / 6	0 / 4	0 / 0	0 / 0	3 / 3	1 / 2	2 / 1	1 / 7	8 / 7	14 / 9
22	3 / 2	10 / 3	1 / 6	0 / 4	1 / 5	0 / 4	0 / 1	3 / 3	2 / 2	9 / 14	0 / 8	0 / 7
23		2 / 2	1 / 2		1 / 2	0 / 2		3 / 1	0 / 1		5 / 14	4 / 8
X	18 / 59	9 / 30	16 / 30	2 / 16	4 / 15	2 / 15	10 / 15	9 / 13	0 / 15	60 / 60	72 / 50	70 / 51
Y	7 / 23	7 / 15	0 / 19	2 / 16	12 / 15	0 / 15	1 / 0	1 / 0	8 / 0	81 / 60	137 / 50	0 / 51
Sum	282 / 668	235 / 436	216 / 403	31 / 132	76 / 119	20 / 131	153 / 209	151 / 202	140 / 177	1001 / 945	862 / 903	796 / 830
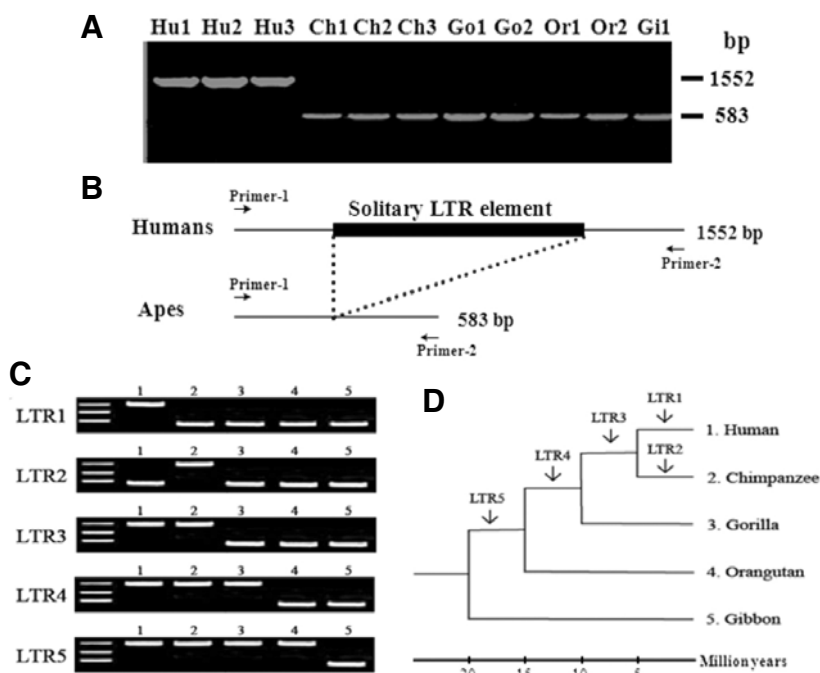
Fig. 2. Schematic representation of the molecular marker for resolving phylogeny and evolution among closely related species. Homologous recombination between the 5′- and 3′-LTRs of a provirus HERV family results in excision of the structural genes and leaves behind a solitary LTR element. The solitary LTR element could be detected by PCR amplification and agarose gel electrophoresis (A, B). Differences in band patterns of the PCR products allow to calculate integration time of solitary LTR element in primate genomes (C, D). Species names are abbreviated as follows: Hu, human; Ch, chimpanzee; Go, gorilla; Or, orangutan. The branching times of the dendrogram should be considered as approximate. Integration event of the solitary LTR element of HERV-K family was modified from Medstrand and Mager (1998) and Buzdin et al. (2003).
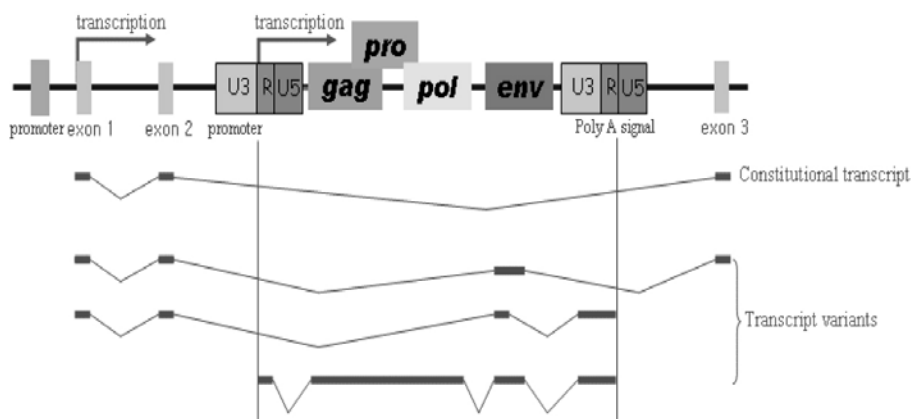
Fig. 3. Schematic representation of genomic structures of HERV elements and functional genes. Alternative promoter of HERV LTR element (U3-R-U5) creates various transcript variants. The structural genes of HERVs are *gag*, *pro*, *pol*, and *env*. The *gag* codes for the structural matrix (MA), capsid (CA), and nucleocapsid (NC) proteins, *pro* codes for protease (PR) protein, *pol* codes for the enzymes of reverse transcriptase (RT) and integrase (IN), and *env* codes for the envelope surface (SU) and transmembrane (TM) proteins.

into the human genome approximately 1.2 million years ago. Therefore, solitary LTR elements could provide valuable information regarding the evolutionary history of retroelements within the primate genome. Further investigation into HERV families will cast light on their biological function in relation to the gene regulatory network system and transcriptional activity.

Transcript variants and alternative promoters of functional genes of the HERV family

Transcript variants of functional genes are induced by HERV elements, which are capable of altering expression (Fig. 3). Further, HERV LTR elements have been suggested to be alternative gene promoters, as alternatively spliced variants of genes show different expression patterns depending on tissue or cell type. Segmental duplication of LTR fusion genes (LTR element and functional gene) can result in the generation of an alternative promoter and transcriptional variants. Recently, placenta-specific expression of IL-2 receptor by the HERV LTR was detected (Cohen et al., 2011). Other examples include the placenta-specific, LTR-driven transcription of endothelin B receptor (*EDNRB*) gene by tissue-specific transcription factors (Landry and Mager, 2003) as well as placenta-specific transcription driven by the alternative LTR promoter of *CYP19* gene located 100 kb upstream of the coding region (Kamat et al., 2002). The HERV-E LTR element has also been used as an alternative promoter for the endothelin B receptor and apolipoprotein C-I genes in humans (Medstrand et al., 2001), whereas the alternative promoter of the HERV-H LTR for the gasdermin-like (*GSDML*) gene was shown to be present in human tissues and cancer cells (Sin et al., 2006). The HERV-K LTR element acts as a switch of alternative splicing and encodes the 67 carboxy-terminal amino acid residues in human leptin receptor protein (Kapitonov and Jurka, 1999). The human *NDUFV1* gene containing a solitary HERV-K LTR element was demonstrated to have strong tissue-specific promoter activity in germ-derived cells (Tera-1 and EP2102), whereas it has no apparent activity in HEK293 cells (Kuzmin et al., 2010). On the other hand, insertion of HERV-H LTR element provides the polyadenylation signals HHLA2 and HHLA3 (Mager et al., 1999). In the case of NAD synthetase 1(*NADSYN1*) gene, the HERV-H LTR integrates into the 3′ flanking region. Thereafter, the polyadenylation signal within the LTR element encourages the formation of alternative transcripts (Sin et al., 2007).

Functional LTR transcription start sites are located between the R and U5 region (3′ termini of the R region) (Kovalskaya et al., 2006). LTR elements have strong bidirectional promoter

activity and contain a negative regulatory element in their U5 region (Akopov et al., 1998; Domansky et al., 2000; Schon et al., 2001). The ERV1 solitary LTR element acts as a bidirectional promoter for the human Down syndrome critical region 4 (*DSCR4*) and *DSCR8* genes. Their shared LTR promoter is more active in the sense than antisense orientation (Dunn et al., 2006). In the case of the Opitz syndrome gene *Mid1*, it has been shown to be transcribed from the HERV-E LTR promoter, indicating that the HERV-E LTR element plays functional roles as both an alternative tissue-specific promoter and enhancer in placenta and embryonic kidney cell lines for the *Mid1* gene (Landry et al., 2002). A human-specific solitary LTR element (L47334) was previously shown to have enhancer activity in Tera-1 human testicular carcinoma cells (Ruda et al., 2004). Promoter activities of human-specific LTR elements may depend on the epigenetic status of their CpG dinucleotides (Ego et al., 2005; Park et al., 2010; Swindle et al., 2004). The varying genetic structure of alternative promoters results in different functional effects, which include stimulating different transcription patterns depending on the tissue type or developmental stage (Gerlo et al., 2006). Use of alternative promoters also could regulate alternative transcript variants, resulting in the production of different protein isoforms (Xin et al., 2008). These changes may result in loss or gain of alternative functional domains. Further, the use of multiple promoters could be an important evolutionary mechanism that provides flexibility in the regulation of gene expression. Specifically, multiple promoters compete with the host promoter for the use of its transcriptional machinery. Thereafter, the integrated proviral elements evolve new biological functions during evolution, thereby regulating transcriptional potential. Therefore, expression of these elements varies significantly among cell lines, in some cases showing strict cell type specificity. Accumulated changes in the LTR elements involved in gene regulation are likely to have consequences with regard to diversification, speciation, and evolution.

CONCLUSIONS

HERV elements have undergone insertion into the primate genome and constitute a substantial part of the human genome. These LTRs change the transcriptional regulation of neighboring genes by supplying new promoters, and definitely contribute to genomic plasticity during primate evolution, which indicates that they could be used as genetic markers for understanding evolutionary history. Multiple copy numbers of HERV elements can create new functional exons, alternative splicing products,

and miRNAs via integration and adaption events, which are implicated in several human diseases such as male infertility, schizophrenia, and cancer. It was also determined that human-specific HERV or solitary LTR elements are involved in the specific activation and antisense regulation of neighboring functional genes. Alternative promoters and enhancers provided by LTR elements contribute to species diversity and transcriptional gene regulation, and their expression tends to be specific depending on cell type, tissue, and developmental stage. Further, tissue-specific promoters may be utilized to develop cell culture models or gene therapies. In some cases, LTR elements do not act as a promoter or enhancer and instead may be implicated in epigenetic regulation. LTR consensus sequences contain multiple binding sites for transcription factors such as p53, CTCF, Pou5F1-Sox2, and ESR1 (Bourque et al., 2008; Cohen et al., 2009). As such, epigenetic regulation of these binding sites is directly correlated with gene silencing and genomic instability. Tissue-specific epigenetic modifications in the LTR promoter region are associated with the formation of transcript variants and gene expression. Therefore, epigenome analyses of the LTR promoter region will revolutionize our understanding of the biological function of HERVs and neighboring genes. Taken together, elucidation of the various roles of HERVs and solitary LTR elements will provide insights into hominid genetic traits in relation to segmental duplication, copy number variation, alternative promoters, epigenetic events, and transcriptional regulation.

ACKNOWLEDGMENTS

REFERENCES

Akopov, S.B., Nikolaev, L.G., Khil, P.P., Lebedev, Y.B., and Sverd-lov, E.D. (1998). Long terminal repeats of human endogenous retrovirus K family (HERV-K) specifically bind host cell nuclear proteins. FEBS Lett. *421*, 229-223.

Barbulescu, M., Turner, G., Seaman, M.I., Deinard, A.S., Kidd, K.K., and Lenz, J. (1999). Many human endogenous retrovirus K (HERV-K) proviruses are unique to humans. Curr. Biol. *9*, 861-868.

Bourque, G., Leong, B., Vega, V.B., Chen, X., Lee, Y.L., Srinivasan, K.G., Chew, J.L., Ruan, Y., Wei, C.L., Ng, H.H., et al. (2008). Evolution of the mammalian transcription factor binding repertoire via transposable elements. Genome Res. *18*, 1752-1762.

Buzdin, A., Ustyugova, S., Khodosevich, K., Mamedov, I., Lebedev, Y., Hunsmann, G., and Sverdlov, E. (2003). Human-specific subfamilies of HERV-K (HML-2) long terminal repeats: three master genes were active simultaneously during branching of hominoid lineages. Genomics *81*, 149-156.

Buzdin, A., Vinogradova, T., Lebedev, Y., and Sverdlov, E. (2005). Genome-wide experimental identification and functional analysis of human specific retroelements. Cytogenet. Genome Res. *110*, 468-474.

Cohen, C.J., Lock, W.M., and Mager, D.L. (2009). Endogenous retroviral LTRs as promoters for human genes: a critical assessment. Gene *448*, 105-114.

Cohen, C.J., Rebollo, R., Babovic, S., Dai, E.L., Robinson, W.P., and Mager, D.L. (2011). Placenta-specific expression of the interleukin-2 (IL-2) receptor β subunit from an endogenous retroviral promoter. J. Biol. Chem. *286*, 35543-35552.

Costas, J. (2000). Characterization of the intragenomic spread of the human endogenous retrovirus family HERV-W. Mol. Biol. Evol. *19*, 526-533.

Dangel, A.W., Baker, B.J., Mendoza, A.R., and Yu, C.Y. (1995). Complement component C4 gene intron 9 as a phylogenetic marker for primates: long terminal repeats of the endogenous retrovirus ERV-K(C4) are a molecular clock of evolution. Immunogenetics *42*, 41-52.

de Parseval, N., Casella, J., Gressin, L., and Heidmann, T. (2001). Characterization of the three HERV-H proviruses with an open envelope reading frame encompassing the immunosuppressive domain and evolutionary history in primates. Virology *279*, 558-569.

Domansky, A.N., Kopantzev, E.P., Snezhkov, E.V., Lebedev, Y.B., Leib-Mosch, C., and Sverdlov, E.D. (2000). Solitary HERV-K LTRs possess bi-directional promoter activity and contain a negative regulatory element in the U5 region. FEBS Lett. *472*, 191-195.

Dunn, C.A., Romanish, M.T., Gutierrez, L.E., van de Lagemaat, L.N., and Mager, D.L. (2006). Transcription of two human genes from a bidirectional endogenous retrovirus promoter. Gene *366*, 335-342.

Ego, T., Tanaka, T., and Shimotohno, K. (2005). Interaction of HTLV-1 Tax and methyl-CpG-binding domain 2 positively regulates the gene expression from the hypermethylated LTR. Oncogene *24*, 1914-1923.

Gerlo, S., Davis, J.R., Mager, D.L., and Kooijman, R. (2006). Prolactin in man: a tale of two promoters. Bioessays *28*, 1051-1055.

Ha, H.S., Chung, W.K., Ahn, K., Bae, J.H., Park, S.J., Moon, J.W., Nam, K.H., Han, K., Cho, H., and Kim, H.S. (2011). Development of GEBRET: a web-based analysis tool for retroelements in primate genomes. Genes Genomics *33*, 679-684.

Hughes, J.F., and Coffin, J.M. (2004). Human endogenous retrovirus K solo-LTR formation and insertional polymorphisms: implications for human and viral evolution. Proc. Natl. Acad. Sci. USA *101*, 1668-1672.

Huh, J.W., Ha, H.S., Kim, D.S., and Kim, H.S. (2008). Placenta-restricted expression of LTR-derived NOS3. Placenta *29*, 602-608.

Huh, J.W., Kim, Y.H., Lee, S.R., Kim, H., Kim, D.S., Kim, H.S., Kang, H.S., and Chang, K.T. (2009). Gain of new exons and promoters by lineage-specific transposable elements-integration and conservation event on CHRM3 gene. Mol. Cells *28*, 111-117.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Kamat, A., Hinshelwood, M.M., Murry, B.A., and Mendelson, C.R. (2002). Mechanisms in tissue-specific regulation of estrogen biosynthesis in humans. Trends Endocrinol. Metab. *13*, 122-128.

Kamp, C., Hirschmann, P., Voss, H., Huellen, K., and Vogt, P.H. (2000). Two long homologous retroviral sequence blocks in proximal Yq11 cause AZFa microdeletions as a result of intrachromosomal recombination events. Hum. Mol. Genet. *9*, 2563-2572.

Kapitonov, V.V., and Jurka, J. (1999). The long terminal repeat of an endogenous retrovirus induces alternative splicing and encode an additional carboxy-terminal sequence in the human leptin receptor. J. Mol. Evol. *48*, 248-251.

Kim, H.S. (2001). Sequence and phylogeny of HERV-W pol fragments. AIDS Res. Hum. Retroviruses *17*, 1665-1671.

Kim, H.S., Wadekar, R.V., Takenaka, O., Winstanley, C., Mitsunaga, F., Kageyama, T., Hyun, B.H., and Crow, T.J. (1999). SINE-R.C2 (a Homo sapiens specific retroposon) is homologous to CDNA from postmortem brain in schizophrenia and to two loci in the Xq21.3/Yp block linked to handedness and psychosis. Am. J. Med. Genet. *88*, 560-566.

Kovalskaya, E., Buzdin, A., Gogvadze, E., Vinogradova, T., and Sverdlov, E. (2006). Functional human endogenous retroviral LTR transcription start sites are located between the R and U5 regions. Virology *346*, 373-378.

Kuzmin, D., Gogvadze, E., Kholodenko, R., Grzela, D.P., Mityaev, M., Vinogradova, T., Kopantzev, E., Malakhova, G., Suntsova, M., Sokov, D., et al. (2010). Novel strong tissue specific promoter for gene expression in human germ cells. BMC Biotechnol. *17*, 58-68.

Landry, J.R., and Mager, D.L. (2003). Functional analysis of the endogenous retroviral promoter of the human endothelin B receptor gene. J. Virol. *77*, 7459-7466.

Landry, J.R., Rouhi, A., Medstrand, P., and Mager, D.L. (2002). The Opitz syndrome gene Mid1 is transcribed from a human endogenous retroviral promoter. Mol. Biol. Evol. *19*, 1934-1942.

Larsson, E., Kato, N., and Cohen, M. (1989). Human endogenous proviruses. Curr. Top. Microbiol. Immunol. *148*, 115-132.

Locke, D.P., Hillier, L.W., Warren, W.C., Worley, K.C., Nazareth, L.V., Muzny, D.M., Yang, S.P., Wang, Z., Chinwalla, A.T., Minx, P., et al. (2011). Comparative and demographic analysis of oran-

gutan genomes. Nature *469*, 529-533.

Lower, R., Lower, 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.

Macfarlane, C., and Simmonds, P. (2004). Allelic variation of HERV-K(HML-2) endogenous retroviral elements in human populations. J. Mol. Evol. *59*, 642-656.

Mager, D.L., and Freeman, J.D. (1995). HERV-H endogenous retroviruses: presence in the New World branch but amplification in the Old World primate lineage. Virology *213*, 395-404.

Mager, D.L., Hunter, D.G., Schertzer, M., and Freeman, J.D. (1999). Endogenous retroviruses provide the primary polyadenylation signal for two new human genes (HHLA2 and HHLA3). Genomics *59*, 255-263.

Medstrand, P., and Mager, D.L. (1998). Human-specific integrations of the HERV-K endogenous retrovirus family. J. Virol. *72*, 9782-9787.

Medstrand, P., Landry, J.R., and Mager, D.L. (2001). Long terminal repeats are used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans. J. Biol. Chem. *276*, 1896-1903.

Ono, M. (1986). Molecular cloning and long terminal repeat sequences of human endogenous retrovirus genes related to types A and B retrovirus genes. J. Virol. *58*, 937-944.

Park, S.J., Huh, J.W., Kim, D.S., Ha, H.S., Jung, Y.D., Ahn, K., Oh, K.B., Park, E.W., Chang, K.T., and Kim, H.S. (2010). Analysis of the molecular and regulatory properties of active porcine endogenous retrovirus gamma-1 long terminal repeats in kidney tissues of the NIH-miniature pig. Mol. Cells *30*, 319-325.

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

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.

Schulte, A.M., Lai, S., Kurtz, A., Czubayko, F., Riegel, A.T., and Wellstein, A. (1996). Human trophoblast and choriocarcinoma expression of the growth factor pleiotrophin attributable to germ-line insertion of an endogenous retrovirus. Proc. Natl. Acad. Sci. USA *93*, 14759-14764.

Sin, H.S., Huh, J.W., Kim, D.S., Kang, D.W., Min, D.S., Kim, T.H., Ha, H.S., Kim, H.H., Lee, S.Y., and Kim, H.S. (2006). Transcriptional control of the HERV-H LTR element of the GSDML gene in human tissues and cancer cells. Arch. Virol. *151*, 1985-1994.

Sin, H.S., Huh, J.W., Kim, W.Y., Kim, D.S., Ahn, K., Ha, H.S., and Kim, H.S. (2007). Long terminal repeats provide alternative polyadenylation signals to NADSYN1 gene. Korean J. Genet. *29*, 395-401.

Sverdlov, E.D. (2000). Retroviruses and primate evolution. Bioessays *22*, 161-171.

Swindle, C.S., Kim, H.G., and Klug, C.A. (2004). Mutation of CpGs in the murine stem cell virus retroviral vector long terminal repeat represses silencing in embryonic stem cells. J. Biol. Chem. *279*, 34-41.

Varmus, H.E. (1982). Form and function of retroviral proviruses. Science *216*, 812-820.

Venables, P.J., Brookes, S.M., Griffiths, D., Weiss, R.A., and Boyd, M.T. (1995). Abundance of an endogenous retroviral envelope protein in placental trophoblasts suggests a biological function. Virology *211*, 589-592.

Voisset, C., Bouton, O., Bedin, F., Duret, L., Mandrand, B., Mallet, F., and Paranhos-Baccala, G. (2000). Chromosomal distribution and coding capacity of the human endogenous retrovirus HERV-W family. AIDS Res. Hum. Retroviruses *16*, 731-740.

Xin, D., Hu, L., and Kong, X. (2008). Alternative promoters influence alternative splicing at the genomic level. PLoS One *3*, e2377

Yi, J.M., Kim, H.M., and Kim, H.S. (2004). Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells. J. Gen. Virol. *85*, 1203-1210.

Yi, J.M., Kim, H.M., and Kim, H.S. (2006). Human endogenous retrovirus HERV-H family in human tissues and cancer cells: expression, identification, and phylogeny. Cancer Lett. *231*, 228-239.