ORIGINAL ARTICLE



Identification and expression analysis of human endogenous retrovirus Y (HERV-Y) in various human tissues

Jeong-An Gim^{1,2} · Kyudong Han³ · Heui-Soo Kim^{1,2}

Received: 17 December 2014/Accepted: 7 June 2015/Published online: 20 June 2015 © Springer-Verlag Wien 2015

Abstract Human endogenous retroviruses (HERVs) account for approximately 8 % of the human genome. To date, several HERV families have been identified in the human genome, with some being valid biomarkers for specific disease states. In this study, we have identified three HERV-Y elements in the human genome and characterized their structure and expression in various human tissues. New HERV-Y elements (HERV-Y101, HERV-Y102, and HERV-Y103) were detected on human chromosomes 8 and 13. In a *pol*-based phylogenetic tree, HERV-Y elements were closely grouped with HERV-I, -T, -E, and -R. The HERV-Y pol gene was expressed ubiquitously in all examined tissues, and it was dominantly expressed in the pons among the 12 different brain regions investigated. These results will allow future studies to elucidate the potential functional roles of HERVs in the brain and other tissues.

Electronic supplementary material The online version of this article (doi:10.1007/s00705-015-2486-z) contains supplementary material, which is available to authorized users.

- ¹ Department of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Republic of Korea
- ² Genetic Engineering Institute, Pusan National University, Busan 609-735, Republic of Korea
- ³ Department of Nanobiomedical Science and WCU Research Center, Dankook University, Cheonan 330-714, Republic of Korea

Introduction

As a vestige of ancient retroviruses, human endogenous retroviruses (HERVs) have been inserted into human genome and account for ~8 % of the whole genome. HERVs have been scattered amongst all chromosomes through germ line infection of ancient exogenous retroviruses and have various functional roles. Some of these roles are specific to disease states, and thus, HERV-derived transcripts can sometimes be used as biomarkers [1, 2]. HERVs are classified based on their primer-binding site (PBS), which binds host tRNA and is involved in replication in the host. To date, many families of HERVs have been identified, and their roles identified [3].

A full-length HERV sequence consists of two long terminal repeats (LTRs), the PBS, and four viral genes (gag, prt, pol, and env). The gag gene codes for structural proteins, the prt gene codes for a protease, the pol gene codes for viral enzymes, and the env gene codes for the surface envelope proteins [4, 5]. HERV elements can influence the host cell in two ways: (1) through the expression of HERVderived viral genes [6, 7] and (2) through the modulation of HERV LTRs as transcriptional regulatory signals [8]. During evolution, most HERV families have been truncated or undergone insertions, deletions, and substitutions [3]. However, some HERV families have conserved their structure in the host genome and express their viral genes. Therefore, it is important to classify HERV families based on their PBS, identify full-length HERV LTRs and viral genes, and confirm the expression of HERV genes in host cells.

Full-length HERVs have long open reading frames, and their transcripts and proteins are expressed in various cell lines and tissues [7, 9, 10]. Some HERV transcripts or proteins are more highly expressed in tumors than in

Heui-Soo Kim khs307@pusan.ac.kr

normal tissues [7, 9] and are more highly expressed in placenta than other tissues [11, 12]. In addition, several studies have indicated that HERV LTR sequences can regulate gene expression by acting as alternative promoters [13, 14], enhancers [15, 16], silencers [17], hormone-responsive elements [18], and polyadenylation signals. In this study, we identified HERV-Y in the human genome, confirmed its structure, and assessed its expression in 20 human tissues and 11 human brain regions, including an Alzheimer's disease sample.

Materials and methods

RNA samples

Total RNA from normal human tissues were purchased from Clontech (Mountain View, CA, USA). These samples contained 500 ng of RNA. A reverse transcription reaction was performed using reverse transcriptase (RT) at an annealing temperature of 42 °C for 90 min with an RNase inhibitor (Promega, Madison, WI, USA) as described [9].

In silico analysis and primer design

The RetroTector program was used to identify HERVs located at a specific locus within the human genome (GRCh37/hg19). Each LTR sequence was aligned using ClustalW in the MEGA6 program [19]. To avoid redundant amplified sequences, *in silico* PCR was performed with the UCSC *in silico* PCR program using *pol* regions as primer sequences for HERV-Y (Table S1). Primer sequences that make multiple products in the human genome were excluded from the study. In order to differentiate the three loci by RT-PCR and real-time RT-PCR, we selected specific regions of three HERV-Y elements for each primer-binding site (Figure S1).

Phylogenetic analysis

Neighbor-joining trees were generated using the MEGA6 program [19] with 100 bootstrap replicates. The percentage of bootstrap replicates supporting the branch is indicated at each node. Sequences for HERV-I on chromosome 7 (AC007276), HERV-T on chromosome 19 (AC078899), HERV-E on chromosome 19 (AC010329), HERV-F on chromosome 11 (AC123788), HERV-W on chromosome 7 (AC007566), HERV-H on chromosome 2 (AC020550), HERV-FRD on chromosome 6 (AL136139), HERV-S on chromosome X (AC233296), HERV-L on chromosome 16 (AC003003), and HERV-M on chromosome 7 (AC004614) were retrieved from GenBank. HERV-R on chromosome 7, known as ERV3-1, was obtained from the UCSC genome

browser. HERV-K on chromosome 6 was represented by the human-specific HERV-K109 sequences described previously [20]. The RetroTector program was used to predict the HERV structure from these sequences. The *pol*-RT region was identified by a BLASTx search [21].

Quantitative real-time RT-PCR

The products of the HERV-Y viral transcript were detected by quantitative real-time RT-PCR using the primers indicated in Table S1, and experiments were performed on a Rotor Gene O (OIAGEN, Hilden, Germany) with a QuantiTect SYBR Green PCR Kit (QIAGEN). In each reaction, the melting curve of amplified samples was a single peak, which indicated one specific PCR product. No template control was amplified, and primer dimers were not detected. For normalization, the ACTB gene was amplified from human β-actin (NP_001092.1). Real-time RT-PCR amplification of HERV-Y elements and the ACTB gene were conducted for 45 cycles of 95 °C for 10 s, 58 °C for 15 s, and 72 °C for 15 s. Then, melting curve analysis was performed for 30 s at 55-99 °C. Each primer set yielded a single, sharp peak, indicating amplification of a single product [9]. All samples were amplified in triplicate to ensure reproducibility.

Results

Structure analysis of three HERV-Y elements

RetroTector was used to screen the human genome for HERVs, and full-length HERVs with the primer-binding site (PBS) for tRNA-Tyr were selected (Figure 1). This PBS is homologous to that of equine endogenous retrovirus (EqERV), which uses tRNA-Tyr for complementary DNA synthesis [22]. The PBS sequence to which the tRNA-Tyr binds had not been described previously in the human genome, and we therefore temporarily named this new HERV family HERV-Y.

We identified the loci and structure of three full-length HERV-Y elements (Figure 2), and each PBS region located between the 5'-LTR and *gag* gene was revealed to have a high-degree of homology to the tRNA-Tyr. The PBS follows the 5' LTR, with each PBS matching at least 15 of the 18 nucleotides at the 3' end of the human tyrosine tRNA. Each HERV-Y element was ordered according to the chromosome number and location, and the full-length HERV-Y family is located on chromosome 8 and chromosome 13 (Table 1). All three HERV-Y elements have a 5' LTR-*gag-prt-pol-env-3'* LTR structure. We characterized the HERV-Y elements according to their size. Full-length HERV-K elements consist of 8.4–9.6 kb, and the

Fig. 1 LTR sequences of three HERV-Y elements and the reverse complement of PBS sequences. Both the 5' and 3'LTRs were identified, and they were aligned against each other. The dashes represent missing residues. The reverse complements of PBS sequences and tRNA-Tyr sequences were aligned. The sites of promoter or polyadenylation sequences are boxed

(a)	HERV-Y3 5LTR HERV-Y3 3LTR	TGTCTAGCTAAAGGTTTGTAAATGCACCCAGTCAGCACTCTGTTAAAATGGACCAATCAGCACTCTGTAAAATGGATCAAT TGTCTAGCTAAAAGGTTTGTAAATGCACCAATCAGCACTCTGTAAAAATGGACCAATCAGCACTCGTAAAAATGGACCAAT
	HERV-Y3 5LTR HERV-Y3 3LTR	CAATGCTCTGTAAAATGGAC-AATCAGCAGGAGGTGGGCAGGGCCAAAGAAGGGAGTAAAAGCTGGCCACCCGAGCCAGC CAGTGTTCCATAAAATGGACCAATCAGCAGGATATGGGTGGG
	HERV-Y3 5LTR HERV-Y3 3LTR	AGCGACAACCTGCTCAGGTCCCCTTCCACGCTGTGGAAGCTTTGTTCTTTTGCTCATCAC2ATAAATCTTGCTGCTGCTGC AGTGGCAACCCTCCAGGTCCCCTTCCATCCTGTGGAAGCTTTGTTCTTTCACTCTTCACCATAAATCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
	HERV-Y3 5LTR HERV-Y3 3LTR	A A
	tRNATyr HERV-Y101	TTGGCgACCACgAAGGGA PBS TTGGCaACCACaAAGGGA
(b)	HERV-Y4 5LTR HERV-Y4 3LTR	TGTGGGTGGGGCCAAATAAGGGAATAAAAACTGGCCACCCGAGCCAGCAGTGGCAACCTGCTTGGGTCCCCCTTTCATGCT TGTGGGCGGGGCCAAATAAGGG <u>AATAAA</u> AGCTGGCCACCAGCAGGAGCAACCTGCTCAGGTCCCCTTCCATACT
	HERV-Y4 5LTR HERV-Y4 3LTR	GTGGAAGCTTTGTTCTTTTGCTCTTCAC2ATAAATCTTGCTGCTGCTCACTCTTTGGGTCTGCATTACCTTTAAGAGCTG ATGGAAGCTTTGTTCTTTCGCTCTTCAC2ATAAATCTTGCTGCTGCTCACTCTTTGGGTCAGCACTACCTTTAAGAGCTC
	HERV-Y4 5LTR HERV-Y4 3LTR	TAACACTCACTGCAAAGGTCTGTGGCTTCACTCCTGAAGTCA TAACACTCACTGTGAAGGTCTGTGGCTTCACTCCTGAAGTCA
	tRNATyr	TTGGCGAcCACGAAGGGA
		77.0 77.0 77.1 77.1 77.1 77.1 77.1 77.1
	HERV-Y102	PBS -TGGCGATCACGAAGGGA
(c)	HERV-Y102 HERV-Y7 5LTR HERV-Y7 3LTR	PBS -TGGCGATCACGAAGGGA TGGGTCGGGTGGGGACCTGCAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCCATCAGCACTATGTGTCTAGCT TGGGTCGGGTGGGGACTTGGAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCAATAAGCACTCTGTGTCTAGCT
(c)	HERV-Y7 5LTR HERV-Y7 3LTR HERV-Y7 5LTR HERV-Y7 5LTR HERV-Y7 3LTR	PBS -TGGCGATCACGAAGGGA TGGGTCGGGTGGGGACCTGCAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCCATCAGCACTATGTGTCTAGCT TGGGTCGGGTGGGGACTTGGAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACAC
(c)	HERV-Y7 5LTR HERV-Y7 5LTR HERV-Y7 5LTR HERV-Y7 5LTR HERV-Y7 5LTR HERV-Y7 5LTR	PBS -TGGCGATCACGAAGGGA TGGGTCGGGTGGGGACCTGCAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCCATCAGCACTATGTGTCTAGCT TGGGTCGGGTGGGGACTTTGGAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACAC
(c)	HERV-Y7 5LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 5LTR HERV-Y7 5LTR HERV-Y7 5LTR HERV-Y7 3LTR	PBS -TGGCGATCACGAAGGGA TGGGTCGGGTGGGGACCTGCAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCCATCAGCACTATGTGTCTAGCT TGGGTCGGGTGGGGACTTGGAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACAC
(c)	HERV-Y7 5LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 5LTR HERV-Y7 3LTR	PBS -TGGCGATCACGAAGGGA TGGGTCGGGTGGGGACCTGCAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCCATCAGCACTATGTGTCTAGCT TGGGTCGGGTGGGGACTTGGAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCAATAAGGACTCTGTGTCAGCT AAAGGTTTGTAAACGCACCTATCAGCACTCTGTAAAAACGAACCAATCAGGCACTCCGTGTCCAGCCAAAGGTTTGTAAAC AAAGTCTGTAAACACACCAATCAGCACTCTGTAAAAATGCACCAATCAGGCACTCGTGTCCAGCCAAAGGTTTGTAAAT ACACCAATCAGCACTCTGTAAAAATGCACCAATCAGAGCTCTGTGTCTAGCTAAAGGTTTGTAAATGCACCAATCAGCAC GCACCAATCAGCACTCTGTAAAAATGCACCAATCAGAGCTCTGTGTCTAGCTAAAGGTTTGTAAATGCACCAATCAGCAC GCACCAATCAGCACTCTGTAAAAATGCACCAATCGGCACTCTGT TCTGTAAAAATGGACCTCTGTAAAAATGCACCAATCGGCCACGGGCCAAATAAGGTATTAAA ACACCAATCAGCACTCTGTAAAAATGCACCAATCGGCCAACTCGGCACTCTGT TCTGTAAAAATGGACCAATCAGCAGGACGTGGGCCGAGGCCAAATAAGGTAATAAA AGCCCACTCAGGACCAATCAGCAGGATGTGGGCCGAGGCCAAATAAGGTAATAAA GCTGGTCGCCCACTCCACCACTGTGGAAGCTTTGTTCTTTTGCTCTTCACAATAAATCTTGCTGCTGCTGCTCACTCT CAACCCACTTGGGTCCCCTTCCACGCTGTGGAAGCTTTGTTCTTTCGCTCTTCACAATAAATCTTGCTGCTGCTGCTCACTCT
(c)	HERV-Y7 5LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 3LTR HERV-Y7 5LTR HERV-Y7 5LTR HERV-Y7 5LTR HERV-Y7 5LTR HERV-Y7 3LTR	PBS -TGGCGATCACGAAGGGA TGGGTCGGGTGGGGACCTGCAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCCATCAGCACTATGTGTCTAGCT TGGGTCGGGTGGGGACTTGGAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCCAATAAGCACTCTGTGTCTAGCT AAAGGTTTGTAAACGCACCTATCAGCACTCTGTAAAAAACGAACCAATCAGCACTCCGTGTCCAGCACAAGGGTTTGTAAAC AAAGTCTGTAAACACACCAATCAGCACTCTGTAAAAAATGCATCAGTGCTCTGTGTCAGCCAAAGGTTTGTAAAC ACACCAATCAGCACTCTGTAAAAATGCACCAATCAGAGCTCTGTGTCTAGCTAAAGGTTTGTAAATGCACCAATCAGGCC GACCAATCAGGACCTATCAGCAGGACGTGGGCAAGGGCCAAATAAAGGGAATAAAAGGCTGGCCACCCAAGCCAGCAGCAGCG CCCCACTCAGGACCAATCAGCAGGACGTGGGCCGAGGCCAAATAAAGGGAATAAAAGCTGGCCACCCAAGCCAGCAGCAGG CAACCCACTCAGGACCAATCAGCAGGACGTGGGCCGAGGCCAAATAAAGGGAATAAAAGCTGGCCACCCGAGCCAGCAGCGG CAACCCACTCAGGTCCCCTTCCACACTGTGGAAGCTTTGTTCTTTTGCTCTTCACAATAAATCTTGCTGCTGCTGCTGCTCACTCT TTGGGTCCACACTACCTTTATGAGCTGT-AACACTCACCGCAAGGGTCTGCA TTGGGTCCACACTACCTTTATAAGCGTTAACACTCACCACAGAGGGTCTGCA
(c)	HERV-Y7 5LTR HERV-Y7 3LTR HERV-Y7 3LTR	PBS -TGGCGATCACGAAGGGA TGGGTCGGGTGGGGACCTGCAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCCATCAGCACTATGTGTCTAGCT TGGGTCGGGTGGGGACTTGGAGAACTTTTCTGTCTAGCTAAAGGATTGTAAACACACCCAATAAGCACTCTGTGTCTAGCT AAAGGTTTGTAAACGCACCCATCAGGACCTCTGTAAAAAAGGAACCAATCAGGCACTCGGTGTCCAGCACAAGGGTTTGTAAAC AAAGTCTGTAAACACACCAATCAGGACCTCTGTAAAAAAGGACCAATCAGGCCTCTGTGTCCAGCCAAAGGTTTGTAAAC ACACCAATCAGGACCTCTGTAAAAATGCACCAATCAGAGCCTCTGTGTCTAGCTAAAGGTTTGTAAATGGACCAATCAGGAC GACCAATCAGGACCTATCAGGAGGACGAGGGCAAGGGCCAAATAAGGGAATTAAAGGCTGGCCACCCAAGGCAGGAGCAG CCCCACTCAGGACCAATCAGGAGGACGGGGCGGGGCGAAATAAGGGAATTAAAGGCGCCCCCCAGGCCAGCAGGAGCGG CAACCCACTCAGGACCAATCAGGAGGATGTGGCCGAGGCCAAATAAGGGAATTAAAAGCTGGCCACCCAAGCCAGGCAGCGG CAACCCACTCAGGTCCCCTTCCACACTGTGGAAGCTTTGTTCTTTTGCTCTTCACAATAAATCTGCTGCTGCTGCTCACTCT CAACCCACTCAGGTCCCCTTCCACGCTGTGGAAGCTTTGTTCTTTGCTCTTCACAATAAATCTTGCTGCTGCTGCTCACTCT TTGGGTCCACACTCACCTTTATGAGCTGT-AACACCCCCCCAGGGGTCTGCA TTGGCCGACCCaCGAaGGGAG

two LTRs range in size between 0.2 and 0.5 kb (Table 1). With respect to HERV-Y101, one potential promoter (ATAAAT) site was identified. The PBS of this element matches 16 of the 18 bases at the 3' end of the human tyrosine tRNA. A putative promoter (ATAAAT) and polyadenylation site (AATAAA) were detected in the HERV-Y102 and HERV-Y103 LTR sequences. These LTR elements could regulate transcription of adjacent genes [8, 11, 23]. Based on this information, the roles of HERV-Y LTR sequences and viral genes could be elucidated in further studies.

Phylogenetic analysis of three HERV-Y elements

Using phylogenetic analysis, we compared the relationship of HERV-Y to other HERV families (Figure 3). The phylogenetic tree was constructed by the neighbor-joining method with HERV family sequences obtained from the GenBank database. Our three HERV-Y elements grouped with HERV-I, -T, -E, and -R, consistent with prior reports [24, 25], and these elements are classified as gammaretroviruses [26]. HERV-L and HERV-S are classified as spumaviruses [4, 26] and are closely related as an outgroup of the group including HERV-Y [3, 27]. HERV-K, also known as HML, is classified as a betaretrovirus and comprised an outgroup of the HERV phylogeny [3, 26, 28]. Two families of HERV (HERV-K, HERV-M) were also independently classified. Our dendrogram illustrates the phylogenetic relationships between the HERV-Y elements and other HERV families.

Expression analysis of HERV-Y

To determine the role of HERV-Y, we confirmed the presence of the pol genes of HERV-Y. In the HERV genome, the *pol* gene encodes the RT and integrase (IN), which are responsible for the synthesis of viral DNA from viral RNA and the integration that DNA into the host genome, respectively [29, 30]. The *pol* gene is therefore essential for viral activity. HERV-Y101 pol was expressed



Fig. 2 Structural analysis of full-length HERV-Y located in specific loci on human chromosomes. The depicted structures were analyzed using HERV-Y sequences from the human genome hg19 assembly with the RetroTector10 program. Those HERVs harbor a 5' LTR-gag-

prt-pol-env-3′ LTR structure. Subdomains detected by RetroTector10 are indicated in three frames in the upper diagrams. Two LTRs, four viral genes, and the PBS are indicated below each diagram

Table 1 List of novel HERV-Y families identified and	Classification	Location	Strand	Total length (kb)	5'/3' LTR (bp)
characterized in this paper	HERV-Y101	chr8:39,733,726-39,742,985	(-)	9.3	240/241
	HERV-Y102	chr8:42,415,006-42,423,430	(-)	8.4	202/196
	HERV-Y103	chr13:111,841,835-111,851,453	(+)	9.6	451/510

ubiquitously in all examined tissues, and HERV-Y102 pol and HERV-Y103 pol were dominantly expressed in the brain (Figure 4). In the case of the HERV-W element, a placenta-specific expression pattern has been observed [12], but HERV-Y elements were dominantly expressed in the brain. Based on our data, HERV-Y101 pol was ubiquitously expressed in all tissues but dominantly expressed in cerebellum and whole brain (Figure 4a, left). HERV-Y102 pol was dominantly expressed in the cerebellum, fetal brain, and prostate (Figure 4b, left). HERV-Y103 pol was also ubiquitously expressed, and the highest level of expression was in the adrenal gland and cerebellum (Figure 4c, left). HERV-Y pol genes were expressed in various tissues, suggesting that they can have biological roles in the host. Given that HERV-Y pol genes were most highly expressed within the brain, the expression pattern within the brain was assessed. Eleven different areas of the brain as well as an Alzheimer's disease brain sample were analyzed, and HERV-Y pol genes were expressed at a higher

level in the pons than in other brain regions (Figure 4, right).

Discussion

Here, we report the identification and expression analysis of HERV-Y, a new multicopy family of HERVs in the human genome. In mammalian genomes, two ERVs predicted to use tRNA-Tyr were detected by *in silico* analysis in cow and horse, as bovine endogenous retrovirus (BoERV) and EqERV, respectively [22, 31]. When identifying retroviral sequences in the host genome, finding the LTRs is important, as it plays an important role in the initiation of transcription. A specific tRNA derived from the host cell binds to the complementary PBS region, located between the 5' LTR and the *gag* region. We used the program RetroTector to identify ERVs that use tRNA-Tyr, and we then predicted the genomic sequence and



Fig. 3 Phylogenetic trees based on the pol RT sequences of HERV-Y and other HERV families. The tree was constructed by the neighborjoining method. Bootstrap evaluation of the branching patterns was performed with 100 replicates. Branch lengths correspond to genetic distance

structure of these elements in the human genome. Viral LTR sequences have promoter activity that induces viral RNA synthesis, and ERVs have promoter activity that can initiate viral gene expression within the host cell, as well as expression of adjacent genes [15, 23]. We identified LTR sequences, performed alignments between 5' and 3' LTR sequences, and then identified promoter and polyadenylation signals. HERV-K [32], HERV-H [33], and HERV-W [17] LTR sequences have been confirmed by their promoter activity in an *in vitro* reporter system, and we expect to investigate the promoter activity of HERV-Y LTRs in future studies.

In this study, we identified three HERV-Y elements in the human genome. HERV-Y101 and HERV-Y102 were detected in chromosome 8, and HERV-Y103 was detected in chromosome 13. These elements have the typical retroviral structure at the genomic level, and mRNA expression patterns were found by real-time RT-PCR. The three HERV-Y elements have well-conserved genomic elements that are similar in length, and they grouped with HERV-I, -T, -E, and -R in a phylogeny tree. The phylogeny indicated that the HERV-Y elements should be classified as gammaretrovirus-like class I HERVs [4, 26]. Class I HERVs have similar expression patterns and are specifically expressed in the brain and reproductive tissues [4]. From this analysis, we reasoned that these three HERV-Y elements could play roles similar to those of class I HERVs in the host.

The activity and location of HERVs can be detected by various *in silico* techniques. Furthermore, HERV activity

can be measured and characterized by RT or *env* gene activity [34–36]. An RT-PCR-based method was first introduced by Silver *et al.* [37], and subsequent methods have used real-time RT-PCR [9, 34] to detect RT or *env* gene activity. In this study, we used *in silico* analysis to identify HERV-Y elements in the human genome. Then, we assessed the expression of the *pol* genes of these elements and found tissue-specific expression patterns.

Many HERV elements have been identified, and their roles in the host have been elucidated. Some HERV elements may play pathological roles in cancer or other diseases and could be useful as biomarkers. In the case of class I HERVs, members of the HERV-E family are not expressed in normal lung and liver tissue but are expressed in the HepG2 and A549 cell lines [38]. HERV-R is expressed in placental tissue as well as in some cancer tissues and cancer cell lines [39]. It is also significantly expressed in normal human placenta as well as liver and lung cancer tissues [9]. HERV-F transcripts have been detected in human leukemia cell lines [40], and they are dominantly expressed in lung cancer [34]. HERV-H is also detected in leukemia cells [40], as well as in the plasma of people with multiple sclerosis [41]. Tissue-specific HERV expression indicates that HERVs might have a biological function. As a syncytin gene, HERV-W is preferentially expressed in the human placenta, where it mediates placental cytotrophoblast fusion during placental development [12].

As one of the best-studied HERV elements, HERV-K is very active within the host genome and is classified as a betaretrovirus-like class II HERV. The activity of HERV-K has been detected in cancer tissues, and it is currently used as a cancer biomarker [42]. HERV-S is closely related to ERV-L and is classified as a spuma-like class III HERV [4, 26, 43]. It is present on the X chromosome [3]. The pol gene of the HERV-S element is expressed within the brain and thymus, as well as in some cancer cell lines. The HERV-S pol gene is well conserved in non-prosimian primates [27]. HERV-V, one of the recently identified HERV elements in chromosome 19, possesses simianspecific elements and has undergone purifying selection during evolution [44]. Over the past two decades, many HERV families and elements have been identified and classified, and their functions have been characterized. The HERV-Y elements are full-length and expressed in various human tissues. The expression patterns of the HERV-Y elements in cancer will need to be studied in the near future, as well as their role in the brain.

The expression of HERVs is related to brain-related diseases such as schizophrenia [45, 46], bipolar disorder [47], and multiple sclerosis [48]. The results of this study suggest that HERV-Y elements have a tendency to be highly expressed in the brain. This information motivated



Fig. 4 Quantitative real-time RT-PCR analysis of HERV-Y *pol* genes in 20 normal human tissues and 12 human brain regions. Relative expression levels of each HERV-Y gene was normalized to the expression levels of the human *ACTB* gene

us to assess the expression of HERV-Y in 12 different brain regions. HERV-Y elements were highly expressed in the pons and thalamus (Figure 4, right), which are known to be crucial nodes of neuronal signals [49]. This expression pattern is similar to the expression pattern for the HERV-W element [50]. The pons is a part of the brainstem that controls the basic vital functions such as heartbeat, breathing, and blood pressure. Importantly, the pons carries signals from the cerebrum to the cerebellum and medulla, and passes sensory signals on to the thalamus [49, 51]. These data suggest that HERV-Y elements have a possible role in controlling the transmission of sensory signals.

In the human genome, this study is the first to report the identification of the HERV-Y family. Following identification, the locations and genomic structure of these elements were determined. We confirmed HERV-Y expression in 20 tissues, 11 brain tissues, and a brain tissue sample from a patient with Alzheimer's disease, using real-time RT-PCR. In summary, we identified new HERV-Y

(101, 102, and 103) families in chromosomes 8 and 13. They were ubiquitously expressed and were especially dominant in the pons. The results presented here provide insight into the biological functions of HERV-Y in those tissues.

Acknowledgments This work was supported by a research program funded by the Korea Centers for Disease Control and Prevention (2013-E7200201; 4800-4845-301).

Conflict of interest The authors declare no conflicts of interest.

References

- Katoh I, Kurata SI (2013) Association of endogenous retroviruses and long terminal repeats with human disorders. Front Oncol 3:234
- Laska MJ, Nissen KK, Nexo BA (2013) (Some) cellular mechanisms influencing the transcription of human endogenous retrovirus, HERV-Fc1. PLoS One 8:e53895

- Tristem M (2000) Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database. J Virol 74:3715–3730
- Blikstad V, Benachenhou F, Sperber GO, Blomberg J (2008) Evolution of human endogenous retroviral sequences: a conceptual account. Cell Mol Life Sci 65:3348–3365
- Urnovitz HB, Murphy WH (1996) Human endogenous retroviruses: nature, occurrence, and clinical implications in human disease. Clin Microbiol Rev 9:72–99
- Ruprecht K, Mayer J, Sauter M, Roemer K, Mueller-Lantzsch N (2008) Endogenous retroviruses and cancer. Cell Mol Life Sci 65:3366–3382
- Zhao J, Rycaj K, Geng S, Li M, Plummer JB, Yin B, Liu H, Xu X, Zhang Y, Yan Y, Glynn SA, Dorsey TH, Ambs S, Johanning GL, Gu L, Wang-Johanning F (2011) Expression of human endogenous retrovirus type K envelope protein is a novel candidate prognostic marker for human breast cancer. Genes Cancer 2:914–922
- Sverdlov ED (1998) Perpetually mobile footprints of ancient infections in human genome. FEBS Lett 428:1–6
- Ahn K, Kim HS (2009) Structural and quantitative expression analyses of HERV gene family in human tissues. Mol Cells 28:99–103
- Reis BS, Jungbluth AA, Frosina D, Holz M, Ritter E, Nakayama E, Ishida T, Obata Y, Carver BS, Scher HI, Scardino PT, Slovin SF, Subudhi SK, Reuter VE, Savage C, Allison JP, Melamed J, Jager E, Ritter G, Old L, Gnjatic S (2013) Prostate cancer progression correlates with increased humoral immune response to a human endogenous retrovirus GAG protein. Clin Cancer Res 19:6112–6125
- 11. Huh JW, Ha HS, Kim DS, Kim HS (2008) Placenta-restricted expression of LTR-derived NOS3. Placenta 29:602–608
- Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, LaVallie E, Tang XY, Edouard P, Howes S, Keith JC Jr, McCoy JM (2000) Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature 403:785–789
- Dunn CA, Medstrand P, Mager DL (2003) An endogenous retroviral long terminal repeat is the dominant promoter for human beta1,3-galactosyltransferase 5 in the colon. Proc Natl Acad Sci USA 100:12841–12846
- Jung YD, Lee JR, Kim YJ, Ha HS, Oh KB, Im GS, Choi BH, Kim HS (2013) Promoter activity analysis and methylation characterization of LTR elements of PERVs in NIH miniature pig. Genes Genet Syst 88:135–142
- Chang NT, Yang WK, Huang HC, Yeh KW, Wu CW (2007) The transcriptional activity of HERV-I LTR is negatively regulated by its cis-elements and wild type p53 tumor suppressor protein. J Biomed Sci 14:211–222
- McGee-Estrada K, Fan H (2007) Comparison of LTR enhancer elements in sheep beta retroviruses: insights into the basis for tissue-specific expression. Virus Genes 35:303–312
- Lee WJ, Kwun HJ, Jang KL (2003) Analysis of transcriptional regulatory sequences in the human endogenous retrovirus W long terminal repeat. J Gen Virol 84:2229–2235
- Goering W, Ribarska T, Schulz WA (2011) Selective changes of retroelement expression in human prostate cancer. Carcinogenesis 32:1484–1492
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725–2729
- 20. Shin W, Lee J, Son SY, Ahn K, Kim HS, Han K (2013) Humanspecific HERV-K insertion causes genomic variations in the human genome. PLoS One 8:e60605
- Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL (2008) NCBI BLAST: a better web interface. Nucl Acids Res 36:W5–W9

- Garcia-Etxebarria K, Jugo BM (2012) Detection and characterization of endogenous retroviruses in the horse genome by in silico analysis. Virology 434:59–67
- Cohen CJ, Lock WM, Mager DL (2009) Endogenous retroviral LTRs as promoters for human genes: a critical assessment. Gene 448:105–114
- Bénit L, Dessen P, Heidmann T (2001) Identification, phylogeny, and evolution of retroviral elements based on their envelope genes. J Virol 75:11709–11719
- 25. Yi JM, Kim HS (2007) Expression and phylogenetic analyses of human endogenous retrovirus HC2 belonging to the HERV-T family in human tissues and cancer cells. J Hum Genet 52:285–296
- 26. van der Kuyl AC (2012) HIV infection and HERV expression: a review. Retrovirology 9:6
- 27. Yi JM, Kim TH, Huh JW, Park KS, Jang SB, Kim HM, Kim HS (2004) Human endogenous retroviral elements belonging to the HERV-S family from human tissues, cancer cells, and primates: expression, structure, phylogeny and evolution. Gene 342:283–292
- 28. Weiss RA (2006) The discovery of endogenous retroviruses. Retrovirology 3:67
- 29. Lower R, Lower J, 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
- Qari SH, Magre S, Garcia-Lerma JG, Hussain AI, Takeuchi Y, Patience C, Weiss RA, Heneine W (2001) Susceptibility of the porcine endogenous retrovirus to reverse transcriptase and protease inhibitors. J Virol 75:1048–1053
- Garcia-Etxebarria K, Jugo BM (2010) Genome-wide detection and characterization of endogenous retroviruses in Bos taurus. J Virol 84:10852–10862
- 32. Domansky AN, Kopantzev EP, Snezhkov EV, Lebedev YB, Leib-Mosch C, Sverdlov ED (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
- Feuchter A, Mager D (1990) Functional heterogeneity of a large family of human LTR-like promoters and enhancers. Nucl Acids Res 18:1261–1270
- Ahn K, Han K, Kim H-S (2011) Quantitative analysis of the HERV pol gene in human tissues. Genes Genom 33:439–443
- Berkhout B, Jebbink M, Zsíros J (1999) Identification of an active reverse transcriptase enzyme encoded by a human endogenous HERV-K retrovirus. J Virol 73:2365–2375
- 36. Huang W-J, Liu Z-C, Wei W, Wang G-H, Wu J-G, Zhu F (2006) Human endogenous retroviral pol RNA and protein detected and identified in the blood of individuals with schizophrenia. Schizophr Res 83:193–199
- Silver J, Maudru T, Fujita K, Repaske R (1993) An RT-PCR assay for the enzyme activity of reverse transcriptase capable of detecting single virions. Nucl Acids Res 21:3593
- Yi JM, Kim HS (2007) Molecular phylogenetic analysis of the human endogenous retrovirus E (HERV-E) family in human tissues and human cancers. Genes Genet Syst 82:89–98
- 39. Kim HS, Yi JM, Hirai H, Huh JW, Jeong MS, Jang SB, Kim CG, Saitou N, Hyun BH, Lee WH (2006) Human endogenous retrovirus (HERV)-R family in primates: chromosomal location, gene expression, and evolution. Gene 370:34–42
- 40. Patzke S, Lindeskog M, Munthe E, Aasheim H-C (2002) Characterization of a novel human endogenous retrovirus, HERV-H/F, expressed in human leukemia cell lines. Virology 303:164–173
- Christensen T, Dissing SP, Riemann H, Hansen H, Munch M, Haahr S, Møller-Larsen A (2000) Molecular characterization of HERV-H variants associated with multiple sclerosis. Acta Neurol Scand 101:229–238

- 42. Wang-Johanning F, Frost AR, Jian B, Epp L, Lu DW, Johanning GL (2003) Quantitation of HERV-K env gene expression and splicing in human breast cancer. Oncogene 22:1528–1535
- 43. Bénit L, De Parseval N, Casella J-F, Callebaut I, Cordonnier A, Heidmann T (1997) Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a gag coding sequence closely related to the Fv1 restriction gene. J Virol 71:5652–5657
- 44. Kjeldbjerg AL, Villesen P, Aagaard L, Pedersen FS (2008) Gene conversion and purifying selection of a placenta-specific ERV-V envelope gene during simian evolution. BMC Evol Biol 8:266
- 45. Karlsson H, Schroder J, Bachmann S, Bottmer C, Yolken RH (2004) HERV-W-related RNA detected in plasma from individuals with recent-onset schizophrenia or schizoaffective disorder. Mol Psychiatry 9:12–13
- Yolken RH, Karlsson H, Yee F, Johnston-Wilson NL, Torrey EF (2000) Endogenous retroviruses and schizophrenia. Brain Res Brain Res Rev 31:193–199

- 47. Frank O, Giehl M, Zheng C, Hehlmann R, Leib-Mosch C, Seifarth W (2005) Human endogenous retrovirus expression profiles in samples from brains of patients with schizophrenia and bipolar disorders. J Virol 79:10890–10901
- 48. Antony JM, Ellestad KK, Hammond R, Imaizumi K, Mallet F, Warren KG, Power C (2007) The human endogenous retrovirus envelope glycoprotein, syncytin-1, regulates neuroinflammation and its receptor expression in multiple sclerosis: a role for endoplasmic reticulum chaperones in astrocytes. J Immunol 179:1210–1224
- Kolmac CI, Mitrofanis J (1998) Patterns of brainstem projection to the thalamic reticular nucleus. J Comp Neurol 396:531–543
- Kim HS, Ahn K, Kim DS (2008) Quantitative expression of the HERV-W env gene in human tissues. Arch Virol 153:1587–1591
- Schmahmann JD, Pandya DN (1997) Anatomic organization of the basilar pontine projections from prefrontal cortices in rhesus monkey. J Neurosci 17:438–458