Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 370 (2006) 34–42
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Human Endogenous Retrovirus (HERV)-R family in primates: Chromosomal location, gene expression, and evolution

Heui-Soo Kim [a,b,*], Joo-Mi Yi [a], Hirohisa Hirai [c], Jae-Won Huh [a], Mi-Suk Jeong [d], Se-Bok Jang [d], Choong-Gon Kim [e], Naruya Saitou [f], Byung-Hwa Hyun [g], Won-Ho Lee [a]

[a] *Division of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Korea*
[b] *PBBRC, Interdisciplinary Program of Bioinformatics, Pusan National University, Busan 609-735, Korea*
[c] *Department of Cellular and Molecular Biology, Primate Research Institute, Kyoto University, Inuyama, Aichi 484-8506, Japan*
[d] *Korea Nanobiotechnology Center, Pusan National University, Busan 609-735, Korea*
[e] *Marine Living Resources Research Division, Korea Ocean Research and Development Institute, Ansan 426-744, Korea*
[f] *Division of Population Genetics, National Institute of Genetics, Mishima 411-8540, Japan*
[g] *National Primate Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-333, Korea*

Received 19 July 2005; received in revised form 31 October 2005; accepted 2 November 2005
Available online 27 January 2006
Received by T. Sekiya

Abstract

Hitherto, full-length endogenous retrovirus (HERV)-R has been located at human chromosome 7q11.2, and mRNA and envelope proteins have been detected in placenta and a variety of other cell types. In the present study, using a probe derived from the gorilla fosmid library, we detected the paralogous locus (7q31.3) of the HERV-R *env* gene in human chromosome 7q11.2, and also determined the chromosomal location in apes and Old World monkeys. The HERV-R gene was not detected in New World monkeys or prosimians with FISH and PCR analyses. We determined the sequences of the HERV-R *env* genes obtained from the genomic DNA of primates using PCR and sequencing tools. Except for a HERV-R *env* sequence derived from gorilla DNA, the functional domains of putative envelope proteins are conserved, suggesting that those domains could have a functional capacity in the primate genome. In addition, we investigated the *env* gene expression of HERV-R in various human tissues and cancer cells. An RT-PCR approach indicated that the *env* gene was expressed in several human tissues (brain, prostate, testis, kidney, placenta, thymus, and uterus) and cancer cells (RT4, BT-474, MCF7, OVCAR-3, LOX-IMVI, and AZ521). Taken together, our data could be of great use for understanding the evolutionary dynamics of HERV-R through primate radiation as well as the implications of its functional role in human tissues and cancers cells.

Keywords: HERV-R; Primates; FISH; Expression; Human tissues; Cancer cells

1. Introduction

Some of the human endogenous retroviruses (HERVs) have been fixed in the human population, and have also been inherited as stable genomic components, through generations in the evolution of species. Approximately 8% of the entire human genome consists of sequences of retroviral origin including retrotransposons (International Human Genome Sequencing Consortium, 2001). They are thought to be remnants of ancient germ line infections by exogenous retroviruses. Most HERVs seem to have entered the genome between 10 and 50 million years ago, and they comprise over 200 distinct groups and subgroups (Jurka, 2000). They can be classified into over 22 HERV families, containing various types from a single copy to one thousand copies (Tristem, 2000). Previously, phylogenetic significance and evolutionary studies have suggested that HERVs inserted into the genome mostly early in primate evolution (Costas, 2001; Yi and Kim, 2004). Most of them were

Abbreviations: HERV-R, human endogenous retrovirus R; FISH, Fluorescence in Situ Hybridization.

* Corresponding author. Division of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Korea. Tel.: +82 51 510 2259; fax: +82 51 581 2962.
E-mail address: khs307@pusan.ac.kr (H.-S. Kim).

0378-1119/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2005.11.008

highly defective with large deletions, stop codons, and frameshifts in the open reading frames (ORFs). However, structural genes from some HERV families are expressed preferentially in human placenta (Mi et al., 2000; Venables et al., 1995) and several cancer cell lines (Armbruester et al., 2002; Yi et al., 2004). This is growing evidence that HERVs could be involved in important cellular processes such as gene regulation, differentiation, immuno-modulation, cell fusion, and tumorigenesis (Krieg and Steinberg, 1990; Larsson and Andersson, 1998). Expression of HERVs can influence the outcome of infections in different ways that can be either beneficial or detrimental to the host. The multiple-copy families scattered throughout the genome have been reported to have regulatory functions on the gene expression of nearby located genes (Akopov et al., 1998). Full-length retroviral sequences might 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 (Leib-Mosch et al., 1993). Some HERVs have been implicated in certain autoimmune diseases, psychiatric diseases, and cancers. Furthermore, they might have a role in disease etiology and pathology (Nelson et al., 1999; Karlsson et al., 2001).

HERV-R (ERV3) was identified by low-stringency hybridization to be derived from the *pol* region of the chimpanzee endogenous provirus CH2 together with the long terminal repeat (LTR) of the baboon endogenous retrovirus (O'Connell et al., 1984). HERV-R has inserted into the genomes of all great apes and Old World monkeys during primate evolution (Shih et al., 1991). In addition, HERV-R is a provirus containing the *pol* and *env* genes and a 3′ LTR. It is present as a single copy located on chromosome 7, and conserves a full-length sequence of Type C retrovirus containing a long ORF with mRNA expression in the *env* region (O'Connell and Cohen, 1984; O'Connell et al., 1984; Cohen et al., 1985). Based on the HERV-R expression, a HERV-R proviral DNA clone was isolated from a human fetal liver library, and demonstrated that HERV-R mRNAs of 9, 7.3, and 3.5 kb are expressed abundantly in the placental chorionic villi throughout gestation (O'Connell et al., 1984; Kato et al., 1987). Levels of HERV-R mRNA expression for the *env* region differ among tissues and different individuals (Kato et al., 1987; Cohen et al., 1988).

A large number of endogenous retroviral sequences are present in the human genome and some of them have been shown to influence the regulation of normal cellular genes; the ERV9 and ERV3 have altered the expression of Krüppel-related zinc finger genes (Di Cristofano et al., 1995; Kato et al., 1990). H-plk (Human-proviral-linked Krüppel) fused with HERV-R was strongly expressed in the placenta. The activation of H-plk could be due to insertion of the HERV-R upstream of the gene (Abrink et al., 1998). Recently, tissue-specific expression of two HERV transcripts (HERV-K *cORF*/*rec* and HERV-R *env*) was found in human normal tissues (kidney, tongue, heart, liver, central nervous system) during embryogenesis, suggesting a possible role in the development and differentiation of human tissues (Andersson et al., 2002).

In the present study, based on the chromosomal localization of HERV-R on 7q11.2 by radiation hybrid mapping (Kim et al., 2000), we detected and identified the paralogous locus of HERV-R on human 7q31.3 and that of various primate species. To understand the molecular and evolutionary features of HERV-R, its sequences and expression were analyzed in relation to functional capacity in various primate genomes, human tissues, and cancer cells.

2. Materials and methods

2.1. Isolation of genomic DNAs from primates and PCR amplification

Genomic DNA segments were isolated from heparinized blood samples of the following species using a standard protocol: hominoid primates; human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), bonobo (*Pan paniscus*), gorilla (*Gorilla gorilla*), orangutan (*Pongo pygmaeus*), and gibbon (*Hylobates agilis*); Old World monkey, Japanese monkey (*Macaca fuscata*), New World monkeys, night monkey (*Aotus trivirgatus*), and common marmoset (*Callithrix jacchus*); prosimian, ring-tailed lemur (*Lemur catta*). The genomic DNA samples were subjected to PCR amplification. The *env* fragments of the HERV-R were amplified by the primer pairs HS41 (5′-GCCTTTACAGACCCAGTAGG-3′, bases 1741-1760) and DS10 (5′-AGGGCCATTAGTCGTGTGGT-3′ bases 2671-2690) from HERV-R (GenBank, accession no. M12140). The PCR conditions were performed as previously described (Kim et al., 1996) with a modified annealing temperature of 56 °C.

2.2. Cell culture and total RNA isolation

Human cancer cells (RT4, PFSK-1, BT-474, HCT-116, TE-1, UO-31, Jurkat, HepG2, A549, MCF7, OVCAR-3, MIA-PaCa-2, PC3, LOX-IMVI, AZ521, 2F7, U-937, C-33A) were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum, 2 mM glutamine, 1 mM nonessential amino acids, 1 mM sodium pyruvate, 100 U/ml penicillin, 0.1 mg/ml streptomycin at 37 °C, and 5% CO_2 incubator. Total RNA from human cancer cells was extracted by High Pure RNA isolation kit (Roche). Total RNA from human tissues (brain, prostate, testis, heart, kidney, liver, lung, placenta, skeletal muscle, spleen, thymus, uterus) was purchased from Roche company.

2.3. Molecular cloning of PCR products

The PCR products were separated on a 1.5% agarose gel, purified with the QIAEX II gel extraction kit (Qiagen) and cloned into the pGEM T-easy vector (Promega). The cloned products were isolated by the alkali lysis method using the high pure plasmid isolation kit (Roche).

2.4. RT-PCR analysis for HERV-R Env gene

Using the pure 1 ug/ul mRNA only, expression patterns of the HERV-R *env* gene were examined by the Titan One Tube

RT-PCR System (Roche). We performed PCR amplification without reverse transcription reaction with pure mRNA samples, indicating that the prepared mRNA samples from the human tissues and cancer cells did not contain the genomic DNA. Using the RT-PCR approach, the HERV-R *env* gene was amplified by the same primer pairs HS41 and DS10. The RT-PCR conditions followed were standard protocol of Titan One Tube RT-PCR System with an annealing temperature of 56 °C. As a standard control, G3PDH was amplified by the primers GPH-S (5′-CAAAGTTGTCATGGATGACC-3′, bases 31721-31740) and GPH-AS (5′-CCATGGAGAAGGCTGGGG-3′, bases 31898-31915) from the human G3PDH (GenBank accession no. AC068657).

2.5. In vitro transcription and translation assay

The HERV-R gene encoding residues 270-834 was amplified by PCR technique with forward (5′-CCGGAATTCATGACTAAAACCCTGTTGTATCAC-3′) and reverse (5′-CCGCTCGAGCTATCCTTTCCAAGTCTGAACTG-3′) primers from human clone (R-HSenv37). The PCR product was purified, digested with *EcoR I* and *Xho I* and then ligated into the pGEX-4T1. *E. coli* strain DH5α was used for cloning and maintaining the constructs. The resulting vector was introduced into *E. coli* BL21(DE3) and BL21(DE3)RIL strain. The expressions of GST-HERV-R fusion proteins were introduced by 1 mM isopropyl-D-thiogalacto pyranoside (IPTG) at optical densities of 0.5–0.6 at 37 °C for 5 h and at 16 °C for 16 h. Bacterial lysates were prepared by sonication in buffer A (50 mM Tris-HCl pH 7.5 and 200 mM NaCl). The supernatant and pellet were used in SDS-PAGE analyses to determine the solubilities of the expressed proteins.

2.6. Fluorescence in situ hybridization (FISH) of HERV-R Env gene in primates

Fosmid clone GGP42M6 derived from gorilla genomic DNA (17.5 μg/μl) was denatured with boiling for 5 min, and labeled with BioPrime DNA labeling system containing biotin-14-dCTP as a hapten (Invitrogen). For DNA precipitation after labeling, a total 740 μl mixture of 50 μl the labeled probe and 690 μl containing 465 μl 99.5% cold (−20 °C) ethanol, 0.5 M NaCl and 19 μg salmon testes DNA was centrifuged at 21,500 *g* after keeping at 4 °C for 60 min. The labeled probe DNA pellet was saturated with 50 μl formamide. Hybridization solution was made with 1 μl probe DNA (350 ng), 8 μl formamide, 1 μl cot I DNA (10 μg), an 15 μl hybridization buffer (3 : 1 30% dextran sulfate : 20×SSC), then was denatured at 72 °C for 10 min before starting hybridization. The chromosome DNA was denatured with 0.05 M NaOH (pH 12.5) in 2×SSC for 4.5 min, followed by dehydration with 70% and 99.5% ethanol for 5 min each. After drying, the denatured probe DNA was applied onto the denatured chromosomes slide preparation. The slide covered with parafilm (cut to size) was incubated in a moist chamber at 37 °C for 12–16 h. Signals were detected with FITC-avidin DCS (Vector) after washing adequately and observed recorded under a fluorescence microscope (Axioplan 2, Zeiss) mounted CCD camera connected to a personal computer (Apple G4) running the IPLab imaging software (Scanalytics, Inc.). These procedures were modified a previously described technique for BAC DNA clones (Hirai and Hirai, 2004). To identify human chromosome 7, a painting probe of biotin-labeled human chromosome 7 (Cambio) was used according to a manufacture's protocol, and the signals were detected with avidin-Rhodamin (Leinco Technologies, Inc.) or avidin-FITC (Vector).

2.7. DNA sequencing and data analyses

Individual plasmid DNA was digested by restriction enzyme *EcoR*I. 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 Biosystems). Nucleotide sequence analyses were performed using the GAP, PILEUP, and PRETTY from the GCG software (University of Wisconsin) and MEGA2 program (Kumar et al., 2001).

2.8. Nucleotide sequence accession numbers

The nucleotide sequence data reported in this paper appear in the DDBJ/EMBL/GenBank nucleotide sequence databases with the accession numbers: CH (AB198934), GO (AB198935), OR (AB198936), GI (AB198937), JM (AB198938).

3. Results

3.1. Structure and chromosomal location of HERV-R family

HERV-R (ERV3) is a provirus containing the *pol*, *env* genes and a 3′ LTR (accession no. M12140). We determined the detailed structure of the HERV-R using a protein domain search by functional domain information from the Pfam HMM database (Fig. 1A). In our previous study, HERV-R (M12140) was been mapped on the human chromosome 7q11.2 using radiation hybrid mapping (Fig. 1B; Kim et al., 2000). In order to confirm the presence of the HERV-R *env* fragment, we performed PCR amplification using genomic DNAs from various primates. The PCR products appeared in the samples of apes and Old World monkeys, but not in those of New World monkeys or a prosimian, indicating that the HERV-R family was integrated into the primate genome before the divergence of apes and Old World monkeys (Fig. 2A). Each of the PCR products from the genomic DNA of the various primates was cloned and sequenced. They showed a high sequence similarity (95.7–98.6%) to that of HERV-R (M12140), except for the sequences derived from gorilla genomic DNA. These sequences showed a low similarity (76.9%) to that of human DNA (Table 1). The gorilla samples contained a 35 bp insertion, unlike those of the other primates. Likewise, five different individuals among the gorilla samples exhibited such an insertion event (Fig. 2B). This information motivated us to investigate the evolutional implications of HERV-R for primate evolution.

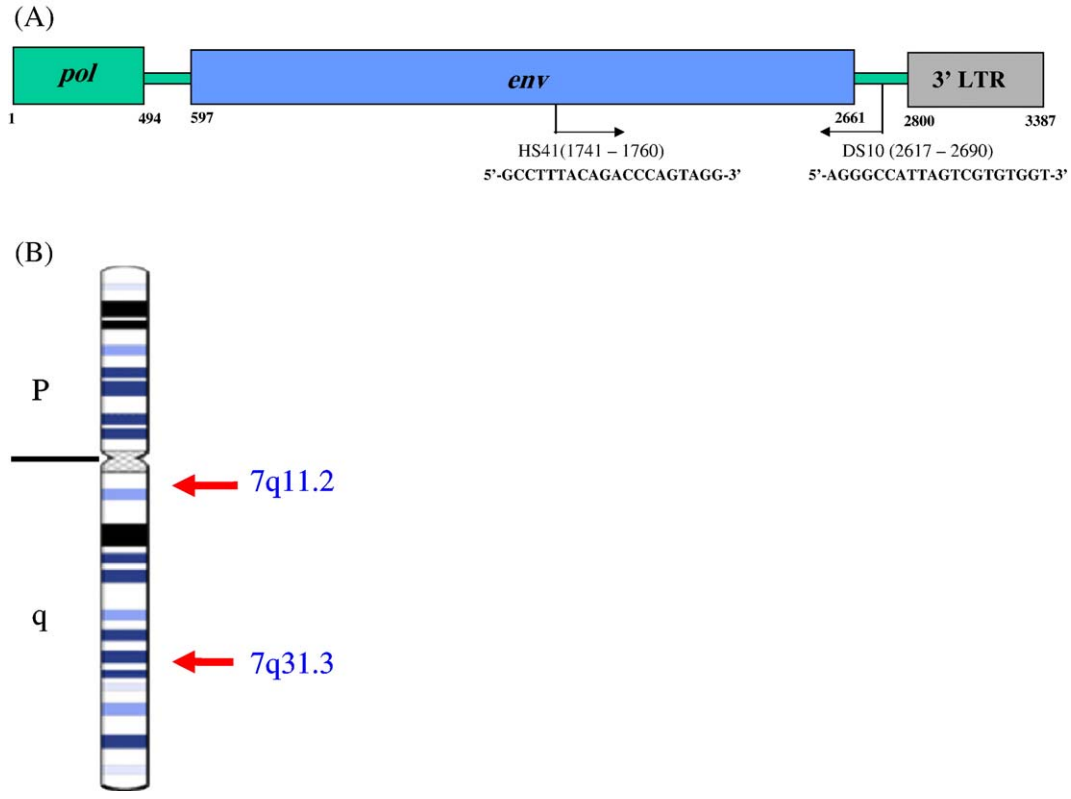

Fig. 1. (A) Putative structure of the HERV-R family (M12140) and (B) chromosomal location of the HERV-R family.

Kim et al. (2003) have constructed and characterized a fosmid clone library from gorilla DNA with a PCR screening method, suggesting that it is useful for comparative genomics of human and apes. Based on this system, one fosmid clone (GGP42M6) from gorilla genomic DNA was identified. We performed a FISH analysis to determine the chromosomal location of the HERV-R element in various primate species (chimpanzee, gorilla, orangutan, gibbon, Japanese monkey, New World monkey, and prosimian), including humans. The clone GGP42M6 (50 kb in length) containing the HERV-R element was located in the bottom region of human chromosome 7q31.3 (Fig. 3a, b). A more exact location might be 7q31.32–33. This location of HERV-R, newly detected at chromosome 7q31.3, is the paralogous locus of human chromosome 7q11.2 mapped by radiation hybrid mapping (Kim et al., 2000). The localization was limited to the orthologous locus in all species used here, that is, the long arm of chromosome 6 of chimpanzees (Fig. 3c, d, e) and gorillas (Fig. 3f, g), chromosome 10 of orangutans (Fig. 3h, i), chromosome 1 of gibbons (Fig. 3j, k, l, m), and chromosome 2 of macaques (Fig. 3n, o, p) (Yunis and Prakash, 1982; Weber et al., 1986; Jauch et al., 1992; Wienberg et al., 1992). Because representatives of the New World monkeys (night monkey, marmoset, and lemur) and prosimians were not hybridized with the probe, the HERV-R element seemingly was inserted into the genomic site of the apes and Old World monkeys lineage approximately 30 million years ago.

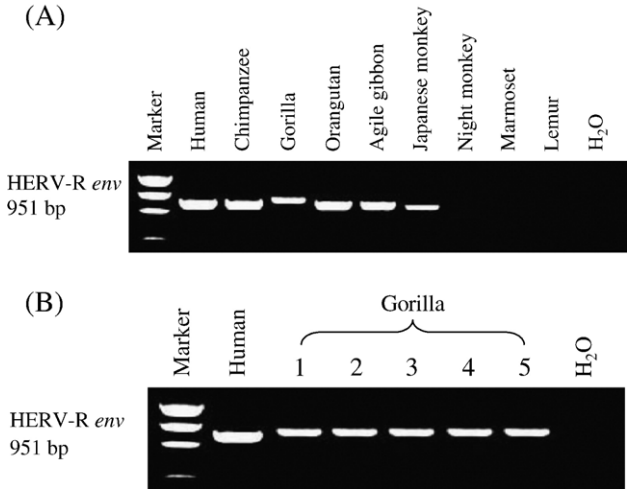

Fig. 2. (A) PCR analysis of genomic DNAs for the presence of the *env* gene of the HERV-R family in primates and (B) five different individuals of gorilla.

Table 1
Percentage similarity of nucleotide sequences of *env* fragments

HERV-R	1	2	3	4	5	6
1. HU	–					
2. CH	98.6	–				
3. GO	76.9	77.0	–			
4. OR	97.3	97.8	76.1	–		
5. GI	96.6	97.4	76.5	97.0	–	
6. JM	95.7	96.3	76.3	96.2	96.3	–

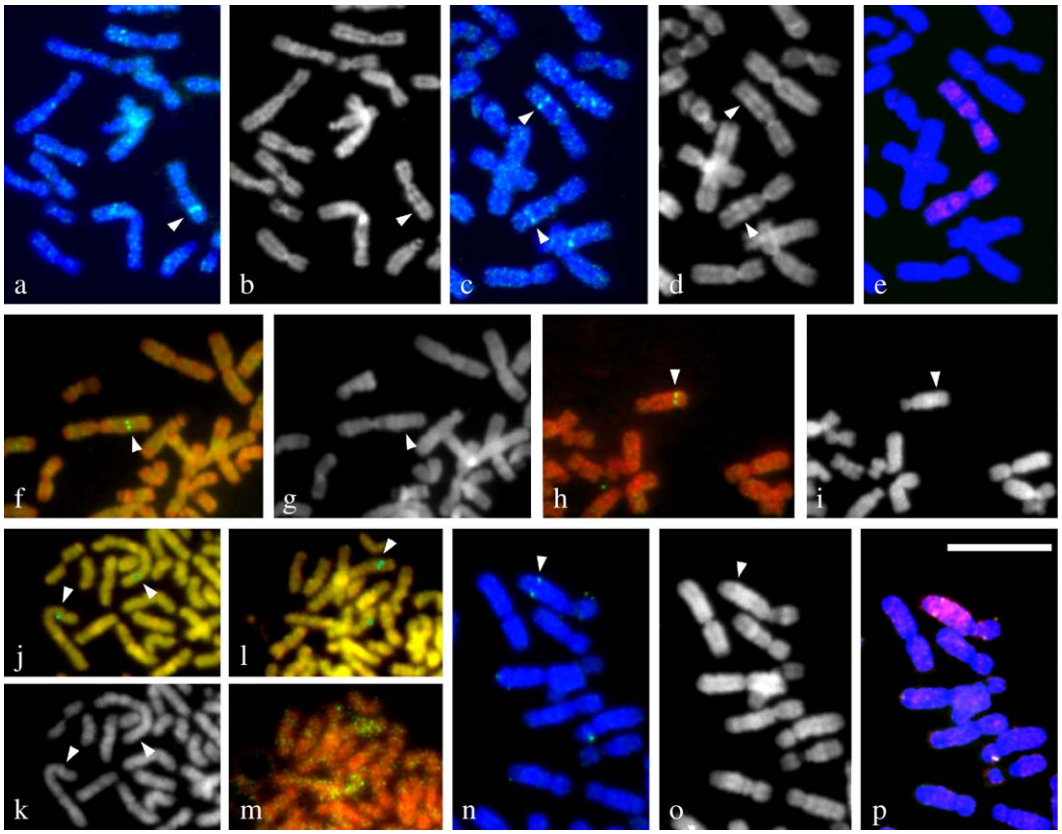

Fig. 3. FISH mapping with a fosmid clone GGP42M6 derived from gorilla genomic DNA on metaphase chromosomes of human (a, b), chimpanzee (c–e), gorilla (f, g), orangutan (h, i), gibbon (j–m), and Japanese monkey (n–p). Some species were checked homologous chromosome with human using a painting probe of human chromosome 7 (pink or green) (e, m, p). Arrowheads indicate locations of signal. The location corresponds to human 7q31.3 (32–33). Scale is 10 μm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Identification and sequence analysis of HERV-R env gene in primates

A full-length endogenous provirus was localized on human chromosome 7q11.2 (O'Connell et al., 1984; Kim et al., 2000). We determined the nucleotide sequences of the HERV-R *env* fragments in various primates (human, chimpanzee, gorilla, orangutan, gibbon, and Japanese monkey). The *env* sequences showed a high degree of sequence similarity (95.7–98.6%) to that of human HERV-R (M12140) except for the *env* sequences (76.9%) from gorilla DNA (Table 1). Amino acid sequences of the *env* fragments of the HERV-R from the various primates were aligned and analyzed (Fig. 4). The putative amino acid sequences of the *env* fragments indicated no translation interruptions by point mutations or deletion/insertions in any of the clones identified in this study with the exception of the *env* sequences from gorilla. The percentage identity of the 246 amino acid sequences of the HERV-R *env* fragments ranged from 95.9% to 97.6% among primates (Table 2A). In order to discover the evolutionary forces at work, we also analyzed synonymous and non-synonymous substitutions within the *env* fragments of HERV-R. As shown in Table 2B, the mean synonymous substitution (K_s) ranged from 0.8% to 7.9%, whereas the mean non-synonymous substitution (K_a) ranged from 1.6% to 3.1%. In K_a/K_s ratio, most of the values in pairwise comparisons were <1. The data indicates that negative selection has been acting on these sequences of the HERV-R family during primate radiation.

3.3. HERV-R Env gene expression in human tissues and cancer cells

The expression pattern of the HERV-R *env* gene in twelve human normal tissues and eighteen human cancer cell lines was examined using an RT-PCR approach. The HERV-R *env* gene was expressed in brain, prostate, testis, kidney, placenta, thymus and uterus of the human normal tissues (Fig. 5A). In the case of the cancer cells, the expression pattern was only detected in the following cell lines: RT4, BT-474, MCF7, OVCAR-3, LOX-IMVI, and AZ521 (Fig. 5B). The RT-PCR products were cloned and sequenced (data not shown), indicating that all of the sequences identified from human tissues and cancer cells originated from HERV-R (M12140) on human chromosome 7q11.2 by primer specificity.

3.4. In vitro transcription and translation analysis of HERV-R Env gene

Active HERVs can serve multiple functions including early placental differentiation, fetal development, and immunosuppression. Along with these functions, a possible role for

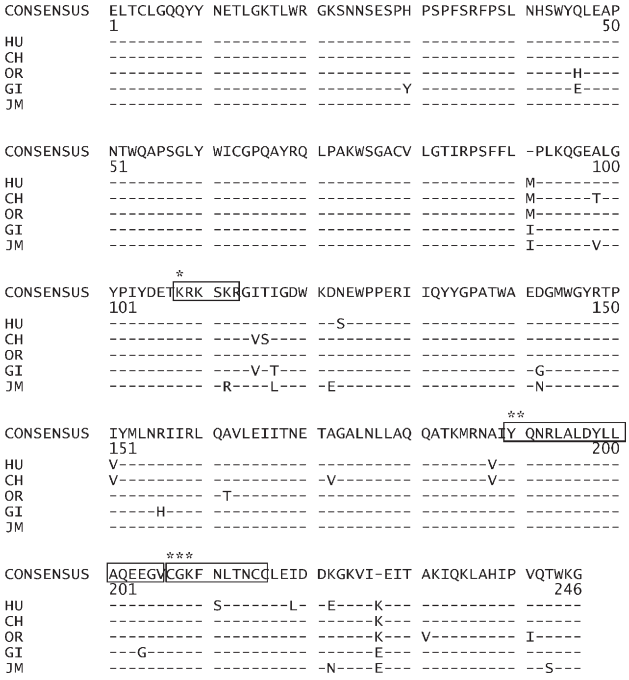

Fig. 4. Amino acid sequence alignments of the HERV-R *env* fragments could encode functional proteins without nonsense or frameshift mutations. Consensus sequences are shown on the top row. Dashes indicate residues identical to the consensus sequences and dot indicates gap introduced to maximize the alignments. *, **, and *** indicate the SU/TM cleavage signal, immunosuppressive domain CKS-17 peptide, and cysteine loop, respectively.

HERV-R in the differentiation of syncytiotrophoblast and U-937 cells has been proposed (Larsson et al., 1996). Therefore, we examined in vitro transcription/translation assays using HERV-R *env* derived from human clone (R-HSenv37), which has only one transition mutation compared to the M12140 sequences. In order to compare the expression efficiencies, two kinds of *E. coli* strains (BL21(DE3) and BL21 (DE3)RIL) and a fusion system (glutathione *S*-transferase [GST]) were used. An SDS-PAGE analysis of the recombinant clones containing all of the two plasmids indicated high levels of production of the recombinant proteins after induction with 1 mM IPTG at 37 °C for 5 h and at 16 °C for 16 h, respectively. The recombinant proteins were found to be insoluble, since they were detected in the insoluble fractions of the *E. coli* sonicate (Fig. 6). Growth temperature and *E. coli* strains did not affect the expression level or the solubility of fusion proteins. As a result of a comparison of the expression level and the solubility, the expression rate for GST fusion systems were high, but the fusion proteins were in insoluble forms.

Table 2A
Percentage identity of 246-amino-acid sequences of *env* fragments

HERV-R	1	2	4	5	6
1. HU	–				
2. CH	97.6	–			
3. OR	97.2	97.2	–		
4. GI	95.9	95.9	96.3	–	
5. JM	96.3	96.3	96.3	97.2	–

Table 2B
Synonymous substitutions per site (K_s), nonsynonymous substitutions per site (K_a) and their ratio (K_a/K_s) in *env* fragments

HERV-R	K_s	K_a	K_a/K_s
HU and CH	0.8	1.6	2.00
HU and OR	5.2	2.0	0.39
HU and GI	4.6	3.1	0.67
HU and JM	6.2	3.0	0.48
CH and OR	4.4	2.0	0.46
CH and GI	3.8	2.6	0.69
CH and JM	5.4	3.0	0.55
OR and GI	6.2	2.6	0.42
OR and JM	7.9	2.5	0.32
GI and JM	5.8	2.8	0.48

4. Discussion

Most HERVs entered into the genome between 10 and 50 million years ago during primate evolution (Jurka, 2000). HERV-R, one of the HERV-ERI supergroup, was integrated into the primate genomes before the divergence of apes and Old World monkeys about 30 million years ago (O'Connell et al., 1984; Shih et al., 1991). Ten members of HERV-R including related sequences have been reported (Tristem, 2000). The exact location, expression, and ORF sequences of the structural genes of the HERV-R provirus in primate species are not known. In the present study, using the probe GGP42M6 derived from the gorilla fosmid library (Kim et al., 2003), we determined the chromosomal location of HERV-R on 7q31.3 in human DNA. It seems to be proliferated by duplication event, suggesting that the HERV-R on 7q31.3 is paralogous to human chromosome 7q11.2, mapped previously by radiation hybrid

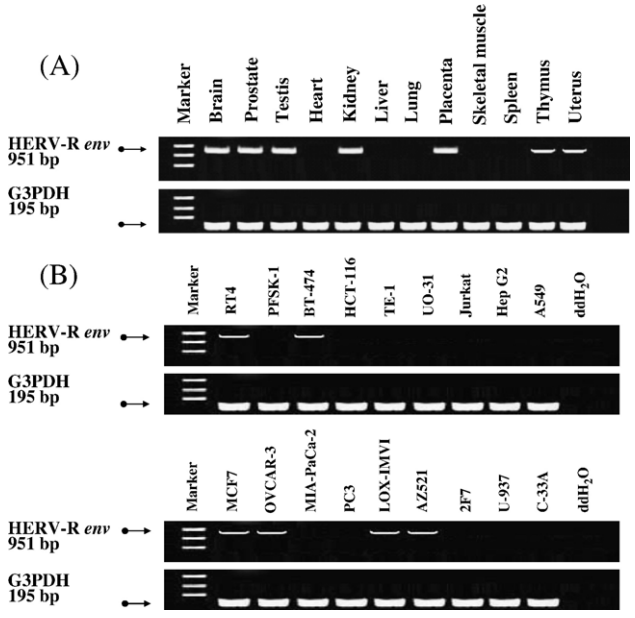

Fig. 5. Expression analysis of HERV-R *env* gene from human tissues (A) and cancer cells (B) by RT-PCR approach.

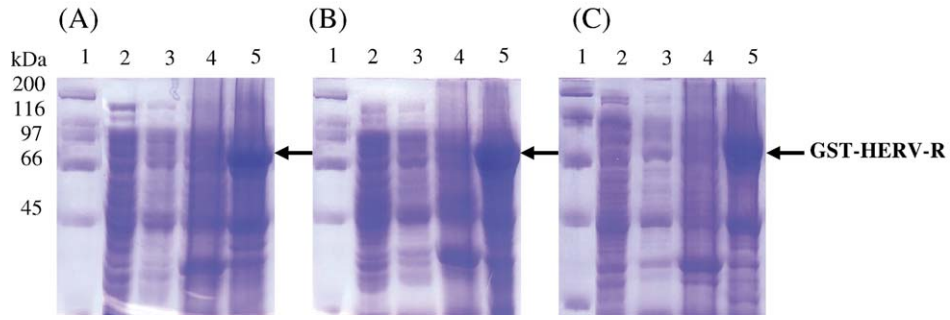

Fig. 6. SDS-PAGE analysis of the expressed proteins in *E. coli*. (A) Expressed GST-HERV-R fusion protein in *E. coli* BL21(DE3) at 37 °C. (B) Expressed GST-HERV-R fusion protein in *E. coli* BL21(DE3)RIL. (C) Expressed GST-HERV-R fusion protein in *E. coli* BL21(DE3) at 16 °C. Lane 1, molecular weight markers; Lane 2, supernatant of lysed bacteria before induction by IPTG; Lane 3, supernatant after induction by IPTG; Lane 4, pellet of lysed bacteria before induction by IPTG; Lane 5, pellet of after induction by IPTG.

mapping (Kim et al., 2000). HERV-R was clearly detected in chimpanzee, gorilla, orangutan, gibbon, and Japanese monkey, but not in New World monkeys or prosimian, by FISH approach. Taken together, HERV-R on the 7q31.3 and 7q11.2 loci of the human chromosome is conserved in the genomes of apes and Old World monkeys, as shown in previous data (O'Connell et al., 1984; Shih et al., 1991).

The HERV-R *env* sequences on the 7q11.2 locus were analyzed by PCR amplification and sequencing from genomic DNA of apes and Old World monkeys, but not from New World monkeys and prosimians, considering an integration event more than approximately 30 million years ago with a subsequent changing in gorillas (Steiper et al., 2004; Fukami-Kobayashi et al., 2005; see also Figs. 2 and 3). We identified and analyzed the HERV-R *env* sequences derived from chimpanzee, orangutan, gibbon, and Japanese monkey DNA. In the case of gorilla DNA, the HERV-R *env* sequences showed a 76.9% similarity to 35 bp insertion compared to humans. This fact implies that the change of HERV-R in gorilla DNA can be explained by the rearrangement or translocation after duplication of the HERV-R family during primate evolution. Most of the sequences derived from the various primates showed higher values for synonymous substitutions than for non-synonymous substitutions in each of the comparisons (Table 2B), suggesting that the HERV-R *env* sequences might have an important biological function in primates. As shown in Fig. 4, the putative amino acid sequences of the HERV-R *env* fragments are well conserved during primate evolution. Interestingly, several functional regions of the protein showed high conservation in primates with only a few changes. These include the immunosuppressive domain, the CKS-17 peptide of mammalian type C retroviruses (amino acids 190 to 206), the cystein loop (amino acids 207 to 216), and the sequence related to the SU/TM cleavage signal (amino acids 108 to 113). Taken together, these HERV-R *env* gene sequences could be associated with an active provirus in primates.

Several HERVs are also highly expressed in steroid-dependent organs such as the placenta, the adrenals, the fetal primitive adrenal cortex, the corpus luteum, the testis and the sebaceous glands. HERV-R is expressed in most human tissues, and its expression is usually found to be low by comparison to that in human placenta, using Northern blot analysis. Elevated expression of HERV-R has recently been reported in sebaceous glands (Andersson et al., 1996). In contrast, the level of HERV-R expression in choriocarcinoma cell lines is extremely low (Kato et al., 1987, 1988). In a different case, the HERV-K *env* gene was expressed in many breast cancer tissues and cell lines, but no expression was detected in normal breast tissues, suggesting that HERV-K *env* expression could be a tumor marker (Wang-Johanning et al., 2003). In a comparative analysis between normal tissues and cancer cells, the HERV-R *env* gene was expressed in brain, prostate, and kidney tissues (Fig. 5A), whereas no expression appeared in PFSK-1, PC3, or UO-31 cancer cells derived from brain, prostate, or kidney tissues, respectively (Fig. 5B). Tissue-specific expression of two HERVs transcripts (HERV-K *cORF*/*rec* and HERV-R *env*) was found in human normal tissues (kidney, tongue, heart, liver, central nervous system) during embryogenesis, suggesting a possible role in the development and differentiation of human tissues (Andersson et al., 2002).

Until now, there has been no clear evidence of the functional roles of HERV-R. There is growing body of evidence that HERV-R might be involved in normal cellular processes. For instance, identification of specific demethylation of the ERV3-H-plk gene in human cell lines and expression levels of the various H-plk transcripts have been reported (Abrink et al., 1998). It has been found, by investigation of the cellular mechanism, that HERV-R *env* expression affects the expression of β-hCG and cell proliferation (Lin et al., 2000). It has been strongly suggested that enhancer and promoter elements in retroviral LTRs influence the transcription of neighboring genes (Kowalski et al., 1999). This can result in transcriptional activation or gene silencing as well as in changes of tissue specificity expression (Ting et al., 1992; Schulte et al., 1996). In addition, the biological role of HERV-R in vivo has been studied using established transgenic rats carrying a full sequence of HERV-R under control of its own LTR promoter (Tanaka et al., 2003).

In this paper, we describe dynamic evolutionary features of the HERV-R family. Their putative amino acids without any mutations such as insertion or deletion in primates and in vitro transcription and translation data, suggest the important biological role of the HERV-R family. Accordingly, the relation of the HERV-R *env* gene to neighboring genes in the genome and to

gene function related to human diseases could be the focus of further research.

Acknowledgments

We thank Yuriko Hirai for her technical assistance for FISH and chromosome painting analyses. This research was supported by the KRIBB Research Initiative Program and by the Grant for the Biodiversity Research of the 21st Century COE (A14).

References

Abrink, M., Larsson, E., Hellman, L., 1998. Demethylation of ERV3, an endogenous retrovirus regulating the Kruppel-related zinc finger gene H-plk, in several human cell lines arrested during early monocyte development. DNA Cell Biol. 17, 27–37.

Akopov, S.B., Nikolaev, L.G., Khil, P.P., Lebedev, Y.B., Sverdlov, 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.

Andersson, M.L., Medstrand, P., Yin, J., Blomberg, J., 1996. Differential expression of human endogenous retroviral sequences similar to mouse mammary tumor virus in normal peripheral blood mononuclear cells. AIDS Res. Hum. Retrovir. 12, 833–840.

Andersson, A.C., Venables, P.J., Tönjes, R.R., Scherer, J., Eriksson, L., Larsson, E., 2002. Developmental expression of HERV-R (ERV3) and HERV-K in human tissue. Virology 297, 220–225.

Armbruester, V., et al., 2002. A novel gene from the human endogenous retrovirus K expressed in transformed cells. Clin. Cancer Res. 8, 1800–1807.

Cohen, M., Powers, M., O'Connell, C., Kato, N., 1985. The nucleotide sequence of the *env* gene from the human provirus ERV3 and isolation and characterization of an ERV3-specific cDNA. Virology 147, 449–458.

Cohen, M., Kato, N., Larsson, E., 1988. ERV3 human endogenous provirus mRNAs are expressed in normal and malignant tissues and cells, but not in choriocarcinoma tumor cells. J. Cell. Biochem. 36, 121–128.

Costas, J., 2001. Evolutionary dynamics of the human endogenous retrovirus family HERV-K inferred from full-length proviral genomes. J. Mol. Evol. 53, 237–243.

Di Cristofano, A., Strazullo, M., Longo, L., La Mantia, G., 1995. Characterization and genomic mapping of the ZNF80 locus: expression of this zinc-finger gene is driven by a solitary LTR of ERV9 endogenous retroviral family. Nucleic Acids Res. 23, 2823–2830.

Fukami-Kobayashi, K., et al., 2005. Genomic evolution of MHC class I region in primates. Proc. Natl. Acad. Sci. U. S. A. 102, 9230–9234.

Hirai, H., Hirai, Y., 2004. In: Melville, S.E. (Ed.), FISH mapping for helminth genome. Methods in Molecular Biology, Parasite Genomics Protocols, vol. 270. Humana Press, Totowa, pp. 379–394.

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

Jauch, A., et al., 1992. Reconstruction of genomic rearrangements in great apes and gibbons by chromosome painting. Proc. Natl. Acad. Sci. U. S. A. 89, 8611–8613.

Jurka, J., 2000. Repbase update: a database and an electronic journal of repetitive elements. Trends Genet. 16, 418–420.

Karlsson, H., Bachmann, S., Schroder, J., McArthur, J., Torrey, E.F., Yolken, R.H., 2001. Retroviral RNA identified in the cerebrospinal fluids and brains of individuals with schizophrenia. Proc. Natl. Acad. Sci. U. S. A. 98, 4634–4639.

Kato, N., et al., 1987. Tissue-specific expression of human provirus ERV3 mRNA in human placenta: two of the three ERV3 mRNAs contain human cellular sequences. J. Virol. 61, 2182–2191.

Kato, N., Larsson, E., Cohen, M., 1988. Absence of expression of a human endogenous retrovirus is correlated with choriocarcinoma. Int. J. Cancer 41, 380–385.

Kato, N., Shimotohno, K., VanLeeuwen, D., Cohen, M., 1990. Human proviral mRNAs down regulated in choriocarcinoma encode a zinc finger protein related to Kruppel. Mol. Cell Biol. 10, 4401–4405.

Kim, H.S., Hirai, H., Takenaka, O., 1996. Molecular features of the TSPY gene of gibbons and old world monkeys. Chromosom. Res. 4, 500–506.

Kim, H.S., Crow, T.J., Hyun, B.H., 2000. Assignment of the endogenous retrovirus HERV-R (ERV3) to human chromosome 7q11.2 by radiation hybrid mapping. Cytogenet. Cell Genet. 89, 10.

Kim, C.G., Fujiyama, A., Saitou, A., 2003. Construction of a gorilla fosmid library and its PCR screening system. Genomics 82, 571–574.

Kowalski, P.E., Freeman, J.D., Mager, D.L., 1999. Intergenic splicing between a HERV-H endogenous retrovirus and two adjacent human genes. Genomics 57, 371–379.

Krieg, A.M., Steinberg, A.D., 1990. Retroviruses and autoimmunity. J. Autoimmun. 3, 137–166.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17, 1244–1245.

Larsson, E., Andersson, G., 1998. Beneficial role of human endogenous retroviruses: facts and hypotheses. Scand. J. Immunol. 48, 329–338.

Larsson, E., et al., 1996. Expression of the endogenous retrovirus ERV3 (HERV-R) during induced monocytic differentiation in the U-937 cell line. Int. J. Cancer 67, 451–456.

Leib-Mosch, C., et al., 1993. Genomic distribution and transcription of solitary HERV-K LTRs. Genomics 18, 261–269.

Lin, L., Xu, B., Rote, N.S., 2000. The cellular mechanism by which the human endogenous retrovirus ERV-3 *env* gene affects proliferation and differentiation in a human placental trophoblast model, BeWo. Placenta 21, 73–78.

Mi, S., et al., 2000. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature 403, 785–789.

Nelson, P.N., et al., 1999. Molecular investigations implicate human endogenous retroviruses as mediators of anti-retroviral antibodies in autoimmune rheumatic disease. Immunol. Invest. 28, 277–289.

O'Connell, C.D., Cohen, M., 1984. The long terminal repeat sequences of a novel human endogenous retrovirus. Science 226, 1204–1206.

O'Connell, C., O'Brien, S., Nash, W.G., Cohen, M., 1984. ERV3, a full-length human endogenous provirus: chromosomal localization and evolutionary relationships. Virology 138, 225–235.

Schulte, A.M., Lai, S., Kurtz, A., Czubayko, F., Riegel, A.T., 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. U. S. A. 93, 14759–14764.

Shih, A., Coutavas, E.E., Rush, M.G., 1991. Evolutionary implications of primate endogenous retroviruses. Virology 182, 495–502.

Steiper, M.E., Young, N.M., Sukarna, T.Y., 2004. Genomic data support the hominoid slowdown and an Early Oligocene estimate for the hominoid–cercopithecoid divergence. Proc. Natl. Acad. Sci. U. S. A. 101, 17021–17026.

Tanaka, S., Ikeda, H., Otsuka, N., Yamamoto, Y., Sugay, T., Yoshiki, T., 2003. Tissue specific high level expression of a full length human endogenous retrovirus genome transgene, HERV-R, under control of its own promoter in rats. Transgenic Res. 12, 319–328.

Ting, C.N., Rosenberg, M.P., Snow, C.M., Samuelson, L.C., Meisler, M.H., 1992. Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. Genes Dev. 6, 1457–1465.

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.

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., Boyd, M.T., 1995. Abundance of an endogenous retroviral envelope protein in placental trophoblasts suggests a biological function. Virology 211, 589–592.

Wang-Johanning, F., Frost, A.R., Jian, B., Epp, L., Lu, D.W., Johanning, G.L., 2003. Quantitation of HERV-K *env* gene expression and splicing in human breast cancer. Oncogene 22, 1528–1535.

Weber, B., Schempp, W., Wiesner, H., 1986. An evolutionarily conserved early replicating segment on the sex chromosomes of man and the great apes. Cytogenet. Cell Genet. 43, 72–78.

Wienberg, J., Stanyon, R., Jauch, A., Cremer, T., 1992. Homologies in human and *Macaca fuscata* chromosomes revealed by in situ suppression hybridization with human chromosome specific DNA libraries. Chromosoma 101, 265–270.

Yi, J.M., Kim, H.S., 2004. Evolutionary implication of human endogenous retrovirus HERV-H family. J. Hum. Genet. 49, 215–219.

Yi, J.M., Kim, H.M., 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.

Yunis, J.J., Prakash, O., 1982. The origin of man: a chromosomal pictorial legacy. Science 215, 1525–1530.