Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells


1Division of Biological Sciences, College of Natural Sciences, Pusan National University, Pusan 609-735, Korea

2Laboratory of Biopotency Evaluation, Korea Research Institute of Bioscience and Biotechnology, Taejeon 305-600, Korea

We examined the structural genes (gag, pol and env) of the human endogenous retrovirus (HERV-W) family from 12 normal human tissues and 18 human cancer cell lines using RT-PCR. For the gag and pol genes, their expression patterns showed tissue or cell specificity, depending on the samples, whereas the env gene was expressed in all tissues and cancer cells examined. To identify active HERV-W elements in tissues and cancer cells, the RT-PCR products were cloned and sequenced. Ninety-five clones for the env gene, 55 clones for the pol gene and 17 clones for the gag gene of the HERV-W family were isolated from human tissues and sequenced, while 85 clones for the env gene, 61 clones for the pol gene and no clones for the gag gene of the HERV-W family were isolated and sequenced from cancer cells. Among these clones, 50 clones from tissues and 44 from cancer cells showed putative amino acids of the HERV-W env gene, while only two clones from cancer cells showed putative amino acids of the HERV-W pol gene. Phylogenetic analysis indicated that several clones identified previously from human monochromosomes had sister relationships with the clones from the different tissues and cancer cells. These data suggest that HERV-W elements are actively expressed in human tissues and cancer cells. These active HERV-W elements deserve further investigation as potential causative agents of various human diseases including cancers.

INTRODUCTION

Human endogenous retroviruses (HERVs) and long terminal repeat (LTR)-like elements are dispersed over 8% of the whole human genome (International Human Genome Sequencing Consortium, 2001). There are at least 22 independent HERV families within the human genome (Tristem, 2000), which originated from germ-cell infection by the exogenous retrovirus during primate evolution (Lower et al., 1996). Full-length retroviral sequences may interact with cellular oncogenes (Varmus, 1982), and retroviral LTR elements have the capacity to exert a regulatory influence as promoters and enhancers of cellular genes. Elucidation of the HERV elements should provide information about fundamental cellular activities and the pathogenesis of multifactorial diseases such as cancer and autoimmune disease (Nakagawa et al., 1997; Mayer, 2001).

HERV-W has been one of the most widely investigated HERV elements following the isolation of an HERV-W-related retrovirus (multiple sclerosis-associated retrovirus, MSRV) from retroviral particles produced in cell cultures from patients with multiple sclerosis (Perron et al., 1997). The new HERV-W family was identified by successive overlapping cDNA clones from the pol region of MSRV (Komurian-Pradel et al., 1999). It has been shown that the 7q21-22 region of human chromosome 7, which has already been identified as potentially being involved in genetic susceptibility to multiple sclerosis (Charmley et al., 1991), contains an RGH proviral copy (HERV-H/RGH) and a complete HERV-W proviral copy named HERV-W/7q. The env gene of HERV-W/7q encodes a protein expressed preferentially in foetal tissues and in the placenta (Alliel et al., 1998). Recently, the HERV-W env protein, syncytin, was shown to be involved in placental morphogenesis (Mi et al., 2000). The major sites of syncytin expression were placental syncytiotrophoblasts and multinucleated cells that originated from foetal trophoblasts. The env gene sequences of the HERV-W family from human monochromosomes and several human cancer cell lines were identified and analysed in our previous study (Kim & Lee, 2001; Yi et al., 2002). In this study, we investigated the
expression, sequences and phylogenetic relationships of the structural genes (gag, pol and env) belonging to the HERV-W family in various human tissues and cancer cells.

METHODS

Cell culture. 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 and C-33A) were grown in Dulbecco's modified Eagle's medium supplemented with 10% heat-inactivated fetal calf serum, 2 mM glutamine, 1 mM non-essential amino acids, 1 mM sodium pyruvate, 100 U penicillin ml\(^{-1}\) and 0-1 mg streptomycin ml\(^{-1}\) at 37 °C in a 5% CO\(_2\) incubator.

RNA samples and RT-PCR analysis. Total RNA from human cancer cells was extracted using a High Pure RNA isolation kit (Roche). Using pure mRNA only, at a concentration of 1 µg µl\(^{-1}\), expression patterns of the HERV-W genes (gag, pol and env) were examined by the Titan One Tube RT-PCR System (Roche). As a standard control, the human glyceraldehyde-3-phosphate dehydrogenase (G3PDH) gene (GenBank accession no. AC068657) was amplified by the primers GPH-S (5'-CAAGTGTGACATGGATGACCC-3', nt 31721–31740) and GPH-AS (5'-CCATGGAGAAGGCTGGGG-3', nt 31898–31915). Using the RT-PCR approach, the HERV-W family was identified from mRNA of human tissues and cancer cell lines. The HERV-W env fragments (GenBank accession no. AF072506) were amplified using the primer pair HS46 (5'-TCCCTGTAACCTGAAATGG-3', nt 1360–1379) and HY76 (5'-CTTTCAGCGGTTAGCAAGTC-3', nt 1920–1939). The HERV-W pol fragments (GenBank accession no. AF009668) were amplified by the primer pair HS48 (5'-ACTACCTGTGGCTACAAGGT-3', nt 1030–1050) and HY78 (5'-AAGTGGCGAAGCTCAAGCA-3', nt 1703–1720). The HERV-W gag fragments (GenBank accession no. AF072500) were amplified using the primer pair JM76 (5'-AAACGCCCTGGCTACAGGCA-3', nt 611–630) and JM28 (5'-GGTGTATTACAGGTGGAG-3', nt 1151–1170). The RT-PCR conditions followed the standard protocol of the Titan One Tube RT-PCR System (Roche) with an annealing temperature of 56 °C.

Molecular cloning of RT-PCR products. RT-PCR products were separated on a 1-8% agarose gel, purified with the QIAEX II gel extraction kit (Qiagen) and cloned into the pGEM T-easy vector (Promega). The cloned cDNA was isolated by the alkaline lysis method using the High Pure plasmid isolation kit (Roche).

Sequencing and data analysis. Individual plasmid DNA was screened for inserts by PCR. Positive samples were subjected to sequence analysis 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 analysis was performed using GAP, PILEUP and PRETTY programs from the GCG package (University of Wisconsin). The neighbour-joining phylogenetic analysis was performed with the MEGA2 program (Kumar et al., 2001). Statistical significance evaluation of the branching pattern was performed with 100 replications.

RESULTS

Expression analyses of HERV-W genes (gag, pol and env) in 12 human tissues and 18 human cancer cell lines were performed by RT-PCR and sequencing. Expression of the env gene was detected in all human tissues examined (brain, prostate, testis, heart, kidney, liver, lung, placenta, skeletal muscle, spleen, thymus and uterus). It was also detected in all cancer cell lines examined (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 and C-33A). The pol gene was expressed in all the human tissues and cancer cell lines tested, except for heart and uterus tissues and PC3, LOX-IMVI and 2F7 cancer cells. The gag gene was expressed specifically in brain, testis, placenta and spleen tissues, whereas no expression was detected in the cancer cell lines (Fig. 1A and B). To eliminate the possibility of DNA contamination, PCR amplification was performed without the reverse transcriptase reaction using the pure mRNA samples, which indicated that the prepared mRNA samples from tissues and cells did not contain genomic DNA (data not shown). The RT-PCR products were cloned and sequenced. The clones of the expressed HERV-W sequences are presented in Tables 1 and 2: 17 clones for gag, 53 clones for pol and 95 clones for env from different human tissues, and 60 clones for pol and 85 clones for env from human cancer cells. The env fragments from human tissues and cancer cells showed a high degree of nucleotide sequence similarity (92.0–99.6%) to that of human HERV-W (AF072506). Deduced amino acid sequences of the 95 clones from multiple human tissues and the 85 clones from cancer cell lines indicated that no translation interruptions as a result of point mutations or deletions/insertions could be found in 50 of the clones from human tissues and 44 of the clones from cancer cell lines (highlighted clones in Tables 1 and 2). In addition, the gag fragments from human tissues also had a high sequence similarity (96.4–99.2%) to that of human HERV-W (AF072500), while the pol fragments from human tissues and cancer cells showed 83.5–91.8% sequence similarity to that of human HERV-W (AF009668). The putative amino acid sequences of the pol gene sequences exhibited multiple frame shifts and termination codons caused by deletions/insertions or point mutations in all identified clones, except for CWP10-6 (from MCF7 cancer cells) and CWP15-8 (from AZ521 cancer cells). The amino acid sequences of the pol fragments of the HERV-W family from cancer cells were aligned and analysed against our previous data (Kim, 2001) (Fig. 2). Notably, four amino acids (ALKK) were deleted in the CWP15-8 clone. Synonymous and non-synonymous substitutions within the HERV-W family were examined to discover the evolutionary forces at work. In the env fragments from human tissues and cancer cells, the mean synonymous substitution ($K_s$) ranged from 0 to 5.9% and 0 to 10%, respectively, whereas the mean non-synonymous substitution ($K_a$) ranged from 0 to 8.6% and 0 to 14%, respectively.

To understand the evolutionary relationships within the HERV-W family in human tissues and cancer cells, we carried out phylogenetic analyses with the clones that had no interruptions from framemshifts or stop codons for the env gene sequences of the HERV-W family. Using all HERV-W env members including our previous data (from a human monochromosomal DNA panel), a neighbour-joining tree was constructed. The tree indicated that HERV-W members have been continuously amplified and
expanded by duplication events during primate evolution and cluster randomly between the clones from various human tissues and cancer cells (Fig. 3). On the tree, we found interesting features within several groups of W-1, W-2, W-3 and W-4. Among these clones of the HERV-W family from human tissues and cancer cell lines, they showed 100% nucleotide sequence similarity to each other, even though they were derived from different tissues or cancer cells. Particularly, clones WE1-5 from brain, WE3-8 from testis, WE5-4 from kidney, WE7-2 from lung, WE8-1 from placenta, WE9-3 from skeletal muscle and WE10-14 from uterus of normal tissues were completely identical to W-7-1 from human chromosome 7 in the W-1 group. Likewise, HERV-W families from different tissues (uterus, spleen, skeletal muscle and prostate) and cancer cells (Jurkat, RT4, HCT-116, A549, MIA-PaCa-2, LOX-IMVI) belonging to the W-2, W-3 and W-4 groups showed completely identical sequences to W-12-2 (from human chromosome 12) and W-14-1 (from human chromosome 14), W-17-5 (from human chromosome 17) and W-X-3 (from human chromosome X), respectively.

**DISCUSSION**

The HERV-W env gene, encoding the protein syncytin, has been shown to be expressed in human placenta and testis with two transcripts of 4 and 8 kb in Northern blot analysis (Mi et al., 2000), and has been found to mediate placental cytotrophoblast fusion connected to placental development. Particles of HERV-W have been recovered from monocyte cultures from patients with multiple sclerosis (Perron et al., 1997). Multiple sclerosis-associated retrovirus (MSRV, a
member of the HERV-W family) has been reported in the serum of patients with multiple sclerosis (Garson et al., 1998). Recently, the HERV-W family was also identified in the cell-free cerebrospinal fluids of 1 of 20 individuals with chronic schizophrenia (Karlsson et al., 2001). Transcriptional activation of the HERV-W family within the central nervous system could be associated with the development of schizophrenia in some individuals. Other HERV families are expressed in normal tissues and cancer cells. The HERV-E family on human chromosome 17q11 is expressed specifically in human pancreas and the thyroid gland, with two major transcripts of 3.3 and 4.1 kb detected by Northern blot analysis (Shiroma et al., 2001). HERV-H env transcripts have been detected in T cell leukaemia cell lines and lymphocytes from healthy blood donors using RT-PCR (Lindeskog & Blomberg, 1997). Recently, the tissue-specific expression of two HERV transcripts (HERV-K cORFrec and HERV-R env) was found in normal human tissues (kidney, tongue, heart, liver and central nervous system) during embryogenesis, suggesting a possible role in the development and differentiation of human tissues (Andersson et al., 2002). Our present data have indicated that retroviral genes are specifically expressed in tissues or cancer cells (Fig. 1). In comparative analysis between normal tissues and cancer cells, the HERV-W gag gene was expressed in brain tissue, whereas no expression appeared in PFSK-1 cancer cells derived from brain tissue. The HERV-W pol gene was also expressed in prostate tissue, but PC3 cancer cells derived from prostate tissue did not show any expression products (Fig. 1C). In the case of the HERV-E env gene, it was actively expressed in prostate carcinoma tissues and cell lines, but not in normal prostate tissues and cells using RT-PCR, RNA in situ hybridization and Northern blot hybridization (Wang-Johanning et al., 2003). No expression of the HERV-E env gene was detected in breast tissue. In contrast, HERV-K env transcripts have been detected in most breast cancer cell lines and many breast tumour tissues (Wang-Johanning et al., 2001). HERV-K gag transcripts were found in gonorocytes of all gonadoblastomas and in neoplastic germ cell tumours (Herbst et al., 1999). The transcriptional activity of the HERV-K gag gene was found to be five- to tenfold higher in the blood cells of leukaemia patients than in normal cells (Depil et al., 2002).

To date, BLAST searches for the HERV-W family have shown that 140 sequences, representing 39 HERV-W proviruses, 40 full-length HERV-W retroposons and 61 truncated HERV-W retrosequences, have been identified in DDBJ/EMBL/GenBank databases (Costas, 2002).

### Table 1. Clones of HERV-W structural genes from various human tissues

<table>
<thead>
<tr>
<th>Human tissues</th>
<th>W-env clones</th>
<th>W-pol clones</th>
<th>W-gag clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Brain</td>
<td>WE1-1, WE1-2, WE1-3, WE1-5, WE1-7, WE1-8, WE1-10, WE1-12, WE1-13, WE1-14, WE1-15, WE1-16, WE1-17, WE1-18, WE1-19, WE1-20</td>
<td>WP1-1, WP1-2, WP1-4, WP1-5, WP1-6, WP1-7</td>
<td>HWG1-1, HWG1-2, HWG1-3, HWG1-4, HWG1-5</td>
</tr>
<tr>
<td>2. Prostate</td>
<td>WE2-1, WE2-2, WE2-3, WE2-4, WE2-5, WE2-7, WE2-10, WE2-11, WE2-12, WE2-14, WE2-15, WE2-16</td>
<td>WP2-4, WP2-5, WP2-8, WP2-10</td>
<td></td>
</tr>
<tr>
<td>3. Testis</td>
<td>WE3-5, WE3-6, WE3-8, WE3-13, WE3-14, WE3-16, WE3-17</td>
<td>WP3-1, WP3-2, WP3-3, WP3-4, WP3-8, WP3-9, WP3-10</td>
<td>HWG3-1, HWG3-3, HWG3-4, HWG3-8</td>
</tr>
<tr>
<td>4. Heart</td>
<td>WE4-3, WE4-6</td>
<td>WP4-6</td>
<td></td>
</tr>
<tr>
<td>5. Kidney</td>
<td>WE5-1, WE5-2, WE5-4, WE5-5, WE5-6, WE5-7, WE5-8, WE5-9</td>
<td>WP5-2, WP5-3, WP5-5, WP5-8, WP5-9, WP5-10</td>
<td></td>
</tr>
<tr>
<td>6. Liver</td>
<td>WE6-1, WE6-3, WE6-12</td>
<td>WP6-3</td>
<td></td>
</tr>
<tr>
<td>7. Lung</td>
<td>WE7-1, WE7-2, WE7-3, WE7-4, WE7-5, WE7-7, WE7-10, WE7-11, WE7-13, WE7-15</td>
<td>WP7-1, WP7-3, WP7-5, WP7-6, WP7-7, WP7-9, WP7-10</td>
<td></td>
</tr>
<tr>
<td>8. Placenta</td>
<td>WE8-1</td>
<td>WP8-1, WP8-6, WP8-8, WP8-2</td>
<td>HWG8-1, HWG8-3, HWG8-4, HWG8-6</td>
</tr>
<tr>
<td>10. Spleen</td>
<td>WE10-3, WE10-4, WE10-7, WE10-9, WE10-11, WE10-12, WE10-13, WE10-14</td>
<td>WP10-1, WP10-2, WP10-4, WP10-5, WP10-8, WP10-9</td>
<td>HWG10-2, HWG10-3, HWG10-5, HWG10-7</td>
</tr>
<tr>
<td>11. Thymus</td>
<td>WE11-1, WE11-2, WE11-3, WE11-5, WE11-7, WE11-9, WE11-12, WE11-13, WE11-15</td>
<td>WP11-2, WP11-5, WP11-7, WP11-8, WP11-9</td>
<td></td>
</tr>
<tr>
<td>12. Uterus</td>
<td>WE12-1, WE12-2, WE12-3, WE12-4, WE12-6, WE12-7, WE12-8, WE12-11, WE12-12, WE12-15, WE12-16</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The number of identified HERV-W-related fragments—at least 70 copies for \textit{gag} and 30 copies for \textit{env} per haploid genome—is correlated with the increase in complexity from \textit{env} to \textit{gag} and \textit{pro} regions previously described (Voisset et al., 2000). In our previous study, we identified 15 HERV-W family \textit{env} genes from human chromosomes 1, 3, 4, 5, 6, 7, 12, 14, 17, 20 and X (Kim & Lee, 2001). One of them, clone W-7-1 from chromosome 7, showed identical sequences to the clones from brain, testis, kidney, lung, placenta, skeletal muscle and uterus of normal tissues (WE1-5, WE3-8, WE5-4, WE7-2, WE8-1, WE9-3 and WE10-14), suggesting that they are actively expressed in various normal tissues (Fig. 3, group W-1). In groups W-2, W-3 and W-4, identical sequences appeared among clones from monochromosomes 12, 14, 17 and X, normal tissues (uterus, spleen, skeletal muscle and prostate) and cancer cells (Jurkat, RT4, HCT-116, AZ521, MA-PaCa-2 and LOX-IMVI), as shown in Fig. 3. The data imply that at least the HERV-W family on human chromosomes 12, 14, 17 and X may be associated with human cancers.

With advances in functional studies, the product of the HERV-W \textit{env} gene has been indicated as having a direct role in human trophoblast cell fusion and differentiation (Frendo et al., 2003). The HERV-W \textit{env} gene product has also been shown to block infection by an exogenous retrovirus, suggesting that the expressed HERV-W \textit{env} gene could have a beneficial function to the host (Ponferrada et al., 2003). The relationship between HERV elements and human diseases has been much discussed in recent years, with reports of the detection of various HERV mRNAs, HERV proteins and even HERV particles in human diseases, especially cancer. We previously reported the characterization of HERV-W \textit{env} gene sequences in some human cancer cell lines to examine the role of HERV in human cancer (Yi et al., 2002) and we also investigated by comparative analysis the HERV-W families in humans and monkeys to understand better their tissue distribution, evolution and phylogenetic relationships (Yi et al., 2003). In this study, we have identified structural genes (\textit{gag}, \textit{pol} and \textit{env}) of the HERV-W family that were expressed in normal tissues and cancer cell lines using the RT-PCR approach and sequence analysis. In bioinformatic analysis, a BLASTN query on the EST (expressed sequence tag) database with HERV-W showed multiple transcripts in various human tissues such as bladder, brain, breast, colon, hypothalamus, kidney, liver, lung, ovary, parathyroid, placenta, prostate, skin, stomach, testis and uterus (data not shown), supporting the suggestion that the HERV-W families could have biological roles in many human tissues. Taken together,
these data suggest that recently proliferated and expressed HERV-W families are active in the genomes of human tissues and cancer cells and could, therefore, play a functional role in human diseases including cancers.

**ACKNOWLEDGEMENTS**

This study was supported by a grant from the Korea Health 21 R&D Project, Ministry of Health & Welfare, Republic of Korea (02-PJ1-PG3-21001-0004).
REFERENCES


