The role of human endogenous retroviral long terminal repeat sequences in human cancer (Review)

HONG-LIAN YU1, ZHAN-KUI ZHAO2 and FAN ZHU1

1Department of Medical Microbiology, School of Medicine; 2Department of Urology, Zhongnan Hospital, Wuhan University, Wuhan, Hubei 430071, P.R. China

Received March 17, 2013; Accepted May 30, 2013

DOI: 10.3892/ijmm.2013.1460

Abstract. Human endogenous retrovirus (HERV) and solitary long terminal repeats (LTRs) constitute 8% of the human genome. Although most HERV genes are partially deleted and not intact, HERV LTRs comprise features including promoters, enhancers, selective splicer sites and polyadenylation sites in order to regulate the expression of neighboring genes. Owing to the genetic instability of LTRs, their wide distributions along human chromosomes are not only non-random, but are also correlated with gene density. Considerable evidence indicates that HERV LTRs regulate the expression of their adjacent viral and cellular genes in placental development and tumorigenesis. However, the regulatory mechanism of HERV LTRs on the expression of its neighboring cancer-associated genes in human cancers remains to be elucidated. Insertional mutagenesis, recombination and polymorphism are three principal factors of LTR that contribute to its genetic instability. Moreover, genetic instability, hypomethylation, transactivation and the antisense transcript of LTRs enhance the activity of LTRs and regulate the expression of their adjacent genes in human cancers. Therefore, in the present review, we examined the mechanism of HERV LTRs in tumorigenesis in combination with the structure and function of LTRs.

Contents

1. Introduction
2. Functions of HERV LTR in physiology
3. Activation of HERV LTR in tumorigenesis
4. Effect mechanisms of HERV LTR in cancer
5. Conclusion

1. Introduction

Almost half of the human genome is derived from transposable elements (Fig. 1A) (1). The majority of transposable elements consist of retrotransposon. Retrotransposon is divided according to whether or not it contains long terminal repeats (LTRs). Human endogenous retroviruses (HERV) and solitary LTR represent ~8% of the human genome as compared to non-LTR retrotransposon elements which represent 34% of the human genome. HERVs are classified into Class I (HERV-W and HERV-H), Class II (HERV-K) and Class III (HERV-L and HERV-S) following phylogenetic analyses of the POL gene with respect to the established classification of animal retroviruses (γ-, β- and spuma retrovirus) (2). Standard HERVs share several general structures with retroviruses, such as 5’LTR, GAG, POL, ENV and 3’LTR (Fig. 1B).

The number and location of LTRs from several main HERV families are shown in Table I. Different HERV families possess distinct copies. The HERV-K family is considered the most active HERV family and has numerous solitary LTRs (3,4). Although HERV LTRs are widely distributed in the human genome, their distributions along the human chromosome are not only non-random but also associated with gene density (5,6). HERV LTRs are enriched in the region of transcription units. LTRs can be located in the sense and antisense orientation of its adjacent gene (Fig. 2). HERV LTRs within introns of annotated genes exhibit a strong orientation bias, such that HERV LTRs are usually oriented opposite to the transcription direction of the corresponding host genes (7). LTRs are uniformly distributed among sequences of a variety of GC content as they are less abundant only in the most GC-rich regions (6). In addition, LTRs have apparently been inserted randomly into these regions. LTRs may be located at any region of the adjacent genes, including 5’UTR, intron, exon and 3’UTR (Fig. 2). These distributions provide favorable conditions for LTRs regulating the expression of their neighboring genes in different ways.

Although the majority of HERV genes are highly defective with large deletions, stop codons and frameshifts in the open reading frames, HERV LTR still retain their functions, such as promoter, enhancer and transcriptional factor-binding site and potentially regulate their neighboring viral and cellular genes. Non-LTR retrotransposon has been reported to cause human cancer through insertional mutagenesis of
genes, retrotransposition-associated mutagenesis, non-allelic homologous recombination and hypomethylation of retroelement promoters (8). Previous studies have reported that LTR elements affected adjacent genes to contribute to cancer (9-11). However, the molecular mechanism of LTR in tumorigenesis has yet to be reported. Herein, we briefly review the characteristics of HERV LTRs, their distribution, as well as the available evidence for LTR biological significance and function in

Table I. Distribution of LTRs from HERV families.

<table>
<thead>
<tr>
<th>Family name</th>
<th>Complete provirus</th>
<th>Solitary LTR (copy)</th>
<th>Integration time (million years ago)</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>HERV-I</td>
<td>85</td>
<td>101-4</td>
<td>30-48</td>
<td>(71)</td>
</tr>
<tr>
<td>HERV-H</td>
<td>100</td>
<td>1000</td>
<td>30-35</td>
<td>(72,73)</td>
</tr>
<tr>
<td>HERV-W</td>
<td>115</td>
<td>1100</td>
<td>30</td>
<td>(74)</td>
</tr>
<tr>
<td>HERV-L</td>
<td>575</td>
<td>6000</td>
<td>70</td>
<td>(75)</td>
</tr>
<tr>
<td>ERV-9</td>
<td>70</td>
<td>3000-4000</td>
<td>33-40</td>
<td>(76)</td>
</tr>
<tr>
<td>HERV-K</td>
<td>20-50</td>
<td>10000-25000</td>
<td>8-15</td>
<td>(2,3,77)</td>
</tr>
</tbody>
</table>

HERV, human endogenous retrovirus; LTR, long terminal repeats.

Figure 1. The classification and structure of transposable elements in the human genome. (A) Approximately 45% of the human genome is considered to be composed of transposable elements. Transposable elements can be classified as: DNA transposon, non-long terminal repeats (LTR) retrotransposons and LTR retrotransposon. LTR retrotransposon consists of three types of HERV (Class I, II and III) according to the classification of retroviruses. (B) Structure of the various transposons. The DNA transposon comprises the transposase gene and two short direct repeats (DRs). The retrotransposon includes the canonical L1 element, Alu element, SVA element, HERV and solitary LTR. The L1 element consists of two open reading frames (ORF) flanked by 5' and 3'UTRs and ends with a polyadenylation [poly(A)] tail. The Alu element consists of two related monomers (A and B grey boxes), A-rich linker region (ATACmA) and ends with a poly(A). The SVA element has a composite structure consisting of a (CCCTCT)n hexamer repeat region, an Alu-like region, a variable number of tandem repeats (VNTR) region, HERV-K10-like sequence and a poly(A) tail. HERV contains 5LTR-GAG-POL-ENV-3LTR. The majority of these genes are not intact. Solitary LTR is formed through excision provirus by homologous recombination. The retrovirus contains 5' and 3'LTR flanking sequences that produce proteins (GAG, POL and ENV) necessary for mobilization.
physiological and pathological processes. The focus of this review is on the role of HERV LTR in human tumorigenesis.

2. Functions of HERV LTR in physiology

HERV LTR elements are retained in the vicinity of genes in order to regulate their expression during primate evolution. LTR elements and nearby genes appear to have forged a mutually beneficial relationship. LTR elements can benefit from genes by surviving from one generation to the next, while genes can benefit from the characteristics of LTR. For example, genes could also obtain their enhancer (12), promoter (sometimes bidirectional) (12-16), polyadenylation [poly(A)] sites (17,18) from LTR. Alteration of the expression of adjacent genes usually results from LTR. For example, an explanation of LTR promoter activity involves the expression of amylase in the human parotid glands, where integration of HERV-E in reverse orientation upstream of the pancreatic amylase gene promotes its expression and release into saliva (19). In addition, bidirectional promoter activity is observed in HERV-K LTR for the neighboring reporter gene (20). As an alternative promoter LTR is capable of enhancing the transcription from a native promoter of apolipoprotein CI and the endothelin B receptor genes (13). HERV-K LTR revealed its tissue-specific
enhancer activity (21). LTR may lead to polyadenylation of spliced chromosomal transcripts in human genes (HHLA2 and HHLA3) as it contains a poly(A) sequence (17). LTR can provide alternative and aberrant sites for transcript splicing and encode an additional carboxy-terminal sequence in the human leptin receptor (22). Similarly, LTR acts as one of at least two alternative promoters for the human \( \beta_{1,3} \)-galactosyltransferase 5 gene and is the dominant promoter in the colon (15). These functions of LTR are essential for LTR to regulate adjacent gene expression.

The most abundant transcriptional activation of HERV LTR may be observed in the placenta and embryonic tissues. Early studies of LTR transcript in placenta emphasized the expression of primary placenta of LTR of different HERV families [ERV3 (23); HERV-F (24); HERV-W (25,26)]. For example, 5'LTR regulates the expression of its downstream viral env gene in placenta trophoblast (26), while the ERV-9 regulatory region acts as a promoter to form a novel transcript of p63 in order to replenish cutaneous epithelial stem cells and maintain the fidelity of the female germ line (27). Further examples include the placenta-specific and tissue-specific expression of adjacent genes driven by LTR (28). These findings suggest that HERV LTR expression in human placenta is frequently active in human physical biology.

### 3. Activation of HERV LTR in tumorigenesis

**Activation of HERV LTR in non-tumor disease.** LTR activation is involved in diseases, such as rheumatoid arthritis (29,30), type I diabetes (30), and schizophrenia (31). For example, the presence of LTR of the HERV-K family is associated with certain DQB alleles. Linkage disequilibrium with DQB1 alleles can contribute to susceptibility to rheumatoid arthritis (29). Expression of the GABBR1 gene is downregulated by HERV-W LTR in schizophrenia (31). Two members of the HERV-I family induce AZFa gene microdeletions in azoospermia patients (32). These abovementioned reports demonstrate that HERV LTR may be active in many diseases.

**Activation of HERV LTR in tumor.** Previous studies (33-37) have described the detection of RNA transcripts from various HERV LTRs in various types of human tumors and cell lines. Elevated HERV-K 5'LTR mRNA is significantly associated with tested prostate cancer tissues (33). It is also reported that the transcripts of the HERV-H LTR-derived promoter are widely distributed in various human tissues and cancer cells (11). Based on the function of LTRs on their adjacent genes in physiology described above, the function of the adjacent gene was investigated in order to elaborate the role of the LTR in tumorigenesis. The cancer-related LTR adjacent genes include viral and cellular genes. Considerable cancer-associated genes regulated by HERV LTR in tumorigenesis are shown in Table II. For example, the transcriptional start sites of HERV-K Rec and Np9 oncogenes are regulated by 5'LTR (Table II). HERV-K Np9 is preferentially expressed in various tumor tissues (34) and the expression of Np9 significantly promoted the growth of leukemia cells in vitro and in vivo (35). HERV-K Rec expression was involved in the process of melanoma and germ cell tumor (36,37). The GSDML gene with the alternative promoters from HERV-H LTR was able to promote
cell proliferation and was correlated with carcinogenesis and the progression of uterine cervix cancer (Table II). DNAJC15 gene also possesses an alternative transcript provided by LTR33 and LTR7 elements. The LTR-related transcripts are only revealed in some cancer cells (HCT106, MCF-3, TE-1, HeLa and CCHM) compared to human tissues (38). Neuronal apoptosis inhibitory protein gene (NAIP), which contains LTR of endogenous retroviral elements as tissue-specific promoter, inhibits apoptosis in the neuron (39). Thus, HERV LTR is capable of regulating the expression of tumor-related genes in different tumor tissues or cells.

4. Effect of possible mechanisms of HERV LTR in cancer

The role of HERV LTR in tumorigenesis as well as the possible mechanisms of how HERV LTR contributes to the development of human cancer was investigated. An outline of potential mechanisms of HERV LTR in human tumorigenesis is provided in Fig. 2. The four aspects of genetic instability, methylation, transactivation and RNA interference by antisense transcripts are discussed below.

Genetic instability. Genetic instability is one of the key features associated with cancer causation and progression (40,41). Instability exists at two distinct levels. In a small subset of tumors, instability is observed at the nucleotide level and results in base substitutions, deletions or insertions of a few nucleotides, such as polymorphisms. In the majority of other cancers, instability is observed at the chromosome level, resulting in losses and gains of whole chromosomes or large portions, for example insertion mutation and recombination. As a mobile element HERV LTR possesses the feature of genetic instability. Therefore, HERV LTR affects human genome instability involvement in tumorigenesis via three principal mechanisms: insertion, recombination and polymorphism (Fig. 2).

Retroviruses mediate malignant transformation via insertion mutations or expression of viral genes. Similarly, HERV is also involved in human tumorigenesis via the mechanism of insertion (42). HERV LTR amplifies in the human genome by generating new integration events. The amplification of HERV integration in the human genome consists of three processes: replication, excision and transduction (Fig. 2). Insertions are therefore considered beneficial, negative or neutral. HERV LTR elements insert into genomic regions and contain crucial biological functions, resulting in insertional mutagenesis during evolution. Abnormal expression of novel gene and adjacent gene are also caused by the insertion of LTR. Among the examples of human cancers caused by LTR-mediated insertion mutagenesis, the expression of PTN gene contributes to the highly aggressive growth of human choriocarcinoma (42). LTR-associated insertional mutagenesis can contribute to cancer via germ-line and somatic mutations, either of which can directly lead to the onset of malignant transformation.

Solitary LTR is formed by homologous recombination between two full-length HERV families. A variety of different recombination products are shown in Fig. 2 (43), including recombination between the two LTRs of a single provirus, homologous recombination between two HERV on the same chromosome, recombination between the 3’ and 5’LTRs of one HERV and gene conversion between the non-homologous genes. Homologous recombination between two proviruses results in substantial deletion and rearrangements of cellular DNA. Gene conversion leads to non-homologous gene exchange with no proviral loss. This occurs predominantly through contributions to recombination, which is frequently detected in cancer (44). LTR33 and LTR7 are formed by recombination and can regulate DNAJC15 gene expression in cancer cells (38). Therefore, LTR recombination may be crucial for LTR to be involved in cancer progression.

The existence of polymorphism provides one explanation of how a ubiquitous gene causes disease in only a proportion of individuals. The polymorphism is subdivided into two broad categories: sequence and insertional polymorphisms. Only a few sequence polymorphisms have been described, partly due to the difficulty in identifying them against the background of closely similar proviruses. Using specific PCR primers spanning 5’LTRs of K115 and K113, Jha et al (45) reported the presence of three single nucleotide polymorphism sites in the K113 5’LTR and four in the K115 5’LTR that together constituted four haplotypes for K113. At present, there is little convincing evidence that any sequence polymorphism of a HERV provides susceptibility to human disease. For insertion polymorphisms, although most HERV families became integrated into the human genome millions of years ago, a new class ofinsertionally polymorphic HERV-K family members has recently been described (46,47). HERV-K113 and HER-K115 insertion has been reported to be involved in various types of cancer (48-51). New insertion polymorphism of HERVs could serve as novel genetic risk factors and thus provide new insight for research into HERV LTR and cancer.

LTR affects the expression of adjacent genes by hypomethylation. Genetic instability of LTR leads to the its wide distribution in the human genome, while the activation of LTR can be controlled by DNA methylation. For example, Gimenez et al (52) have recently reported that hypomethylation of the promoter domain of the HERV U3 element appears to be a prerequisite for the increased expression in tumor tissues compared to normal tissues. Colon cancer cells are treated with DNA methylation and histone deacetylase inhibitors, and RT-PCR results show that the expression patterns of HERV-H are significantly altered in several colon cancer cells. The finding suggests that the hypomethylation context affects the expression of HERV-H elements in colon cancer cells (53). Increased HERV-K expression in melanomas may be due to increased promoter activity and demethylation of the 5’LTR (54). Thus, overexpression of the HERV sequence in cell lines is correlated with the demethylation of LTR.

LTR affects the expression of adjacent genes by RNA interference. HERV and LTR are usually oriented opposite to the transcription direction of corresponding host genes (55-57). Antisense transcripts affect the sense partner gene function by modifying the transcriptional and post-transcriptional regulation processes. Findings of a previous study (55) demonstrated that from 10 HERV-K LTRs which were localized in introns of unique human genes, nine exhibited opposite orientation to the transcription direction of the corresponding human genes. A hypothesis was propounded...
that LTR affects the gene expression by initiation of the antisense RNA synthesis (55). Subsequently, several reports supported the hypothesis (57-61). Antisense transcript of HERV-Ec1 affected the expression of cytosolic phospholipase A2 in urothelial carcinoma (57). A novel exon cassette is derived from the antisense transcript of the HERV-K element (58). Intronic RNAs arising from U3 of ERV-9 LTR are expressed as both sense and antisense transcripts, with the antisense transcript being expressed at higher levels compared to the sense expression in malignant cells (59). LTR from exogenous retrovirus HTLV can also generate antisense manuscripts such as the HTLV-1 basic leucine zipper factor (HBZ) (60). It has been reported that HBZ is consistently expressed and remains intact in all ATL cases and HTLV-1-infected individuals, where it promotes cell proliferation (61). Antisense transcripts are able to form double-stranded RNA and may recruit RNAi machinery. The abovementioned studies suggest that the abnormal expression of antisense transcripts of LTR retrotransposon may be a causative factor for tumorigenesis.

**Transactivation of LTR.** HERV LTR could also be reactivated by environmental factors including cytokines (62), radiation (63,64) and proteins of exogenous retroviruses (65,66). Radiation induced the epigenetic regulation of HERV-R 5’LTR and upregulated HERV-R env expression (63). HSV-1 activated the LTR activity of HERV-W to enhance their potential oligodendrotoxic and immunopathogenic effects (67). HSV-1 infection also induced the LTR-directed transcription of the HERV-K. HSV-1 immediate-early ICP0 protein was able to upregulate the activity of HERV-K LTR (66). Relative levels of transcripts encoding HERV-W elements and cellular genes are transactivated by viral infection in different cell lines by regulating the transcriptional activity (68). Several HERV LTRs could be activated by HTLV Tax and were involved in diseases (65). High-level expression of HERV-K has been demonstrated to be activated by the MITF-M gene in melanomas, breast cancers and teratocarcinomas (69). Of note, these factors can also target ERVs, for example, a HERV-W LTR-directed transcription of the HERV-W env expression (63). HSV-1 infection also induced the LTR-directed transcription of the HERV-K. HSV-1 immediate-early ICP0 protein was able to upregulate the activity of HERV-K LTR (66).

5. Conclusion

In summary, cumulative evidences indicate that LTR may be involved in the process of tumorigenesis at various levels. The potential role of LTR in human cancer appears much more complex (Fig. 2). LTR can be involved in tumorigenesis in four distinct ways, i.e., genetic instability, hypomethylation, transactivation and RNA interference. Insertional mutagenesis, recombination and polymorphism of LTR have been found to contribute to its genetic instability. Genetic instability can lead to diverse distribution and high copy numbers of LTR in the human genome. LTR can exhibit a sense or antisense orientation of their neighboring genes. By contrast, LTR can also be located in the 5’UTR, intron, exon and 3’UTR region of their neighboring genes. Genome-wide hypomethylation, transactivation of LTR and RNA interference by antisense orientation of LTR can regulate the expression of human cellular, viral and novel genes. Abnormal expression of cancer-associated genes may also contribute to tumorigenesis. Thus, identification of the distribution, structure and functional characteristics of LTR is extremely important to elucidate the mechanism involved in LTR regulation of cancer-associated gene abnormal expression in the process of tumorigenesis and development. Given the strong evidence for the abnormal expression of LTR adjacent genes in human tumors, it can be hypothesized that similar unknown genes or other tumor-associated genes may also affect various types of human cancer. Although a causative role of HERV LTR in human tumors has not been reported, HERV LTR is a potential contributory factor in various types of human cancer. Further investigations should include, for example, the novel insertion of HERV LTR in the human genome, RNA interference by antisense transcript of LTR, distribution of HERV LTR in different human chromosomes and the identification of LTR-related tumor genes.

**References**


