Friends-Enemies: Endogenous Retroviruses Are Major Transcriptional Regulators of Human DNA

Anton A. Buzdin [1,2]*, Vladimir Prassolov [1] and Andrew V. Garazha [3,4]

[1] Department of Cell Biology, Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia, [2] Centre for Convergence of Nano-, Bio-, Information and Cognitive Sciences and Technologies, National Research Centre "Kurchatov Institute," Moscow, Russia, [3] Group for Genomic Regulation of Cell Signaling Systems, Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia, [4] Department of Biomedicine, Moscow Institute of Physics and Technology, Moscow, Russia

Endogenous retroviruses are mobile genetic elements hardly distinguishable from infectious, or "exogenous," retroviruses at the time of insertion in the host DNA. Human endogenous retroviruses (HERVs) are not rare. They gave rise to multiple families of closely related mobile elements that occupy ∼8% of the human genome. Together, they shape genomic regulatory landscape by providing at least ∼320,000 human transcription factor binding sites (TFBS) located on ∼110,000 individual HERV elements. The HERVs host as many as 155,000 mapped DNaseI hypersensitivity sites, which denote loci active in the regulation of gene expression or chromatin structure. The contemporary view of the HERVs evolutionary dynamics suggests that at the early stages after insertion, the HERV is treated by the host cells as a foreign genetic element, and is likely to be suppressed by the targeted methylation and mutations. However, at the later stages, when significant number of mutations has been already accumulated and when the retroviral genes are broken, the regulatory potential of a HERV may be released and recruited to modify the genomic balance of transcription factor binding sites. This process goes together with further accumulation and selection of mutations, which reshape the regulatory landscape of the human DNA. However, developmental reprogramming, stress or pathological conditions like cancer, inflammation and infectious diseases, can remove the blocks limiting expression and HERV-mediated host gene regulation. This, in turn, can dramatically alter the gene expression equilibrium and shift it to a newer state, thus further amplifying instability and exacerbating the stressful situation.

Keywords: retrovirus, gene expression regulation, pathology, cancer, inflammation, stress, stability, infection

Human endogenous retroviruses (HERVs) and related genetic elements occupy ∼8% of human genome. They are thought to be remnants of multiple ancient retroviral infections (Sverdlov, 2000; Belshaw et al., 2004; Buzdin, 2007). HERV insertions occurred in the ancestral germ cell lineage, fixed in the genome and became inheritable (Buzdin et al., 2003; Dewannieux and Heidmann, 2013). In the human DNA, HERVs are represented by 504 groups including 717.778 individual fragments (RepeatMasker, hg19). The individual HERV copies are frequently interrupted by other sequences, such as transposable elements, and may each represent two or more genomic fragments.

Older HERVs have accumulated more mutations, including indels, and thus are more fragmented then the evolutionary young elements. For example, the MER-41-int element located at the position chr1:26,952,949-26,962,938 (hg19 assembly) is broken into four fragments in the genome, but biologically this was a single HERV.

Many families of HERVs are highly transcriptionally active in human tissues (Buzdin et al., 2006a; Maliniemi et al., 2013). Genomic copies of HERVs are of particular interest because in addition to viral genes they also have various regulatory sequences concentrated in their *long terminal repeats* (LTRs)— about 1 kb long fragments of DNA flanking the "body" of an element (**Figure 1**). The LTRs serve as promoters (Buzdin et al., 2006a), enhancers (Chuong et al., 2013; Suntsova et al., 2013), polyadenylation signals (Suntsova et al., 2015), chromatin folding reshapers (Schumann et al., 2010), and binding sites for various nuclear proteins (Young et al., 2013). Importantly, most of HERVs reside in the human genome as solitary LTRs arisen due to homologous recombinations between the two 5′- and 3′-flanking LTRs of the same full-length element (Hughes and Coffin, 2004). In turn, further recombinations between the different HERVs may cause genomic instability (Trombetta et al., 2016). For example, this mechanism may be responsible for at least 78 copy number variation cases encompassing known human genes (Campbell et al., 2014).

Most of the newly inserted HERVs harbor functional retroviral genes, such as those encoding for the reverse transcriptase/integrase, the structural polyprotein Gag and the envelope polyprotein Env, and the canonical function of an LTR is the regulation of retroviral expression. However, the LTRs may also drive the transcription of closely located genomic sequences and human genes (Buzdin et al., 2006a). In this minireview, we pay attention to the regulatory function of HERVs which donated multitude of functional sequences to the human genome.

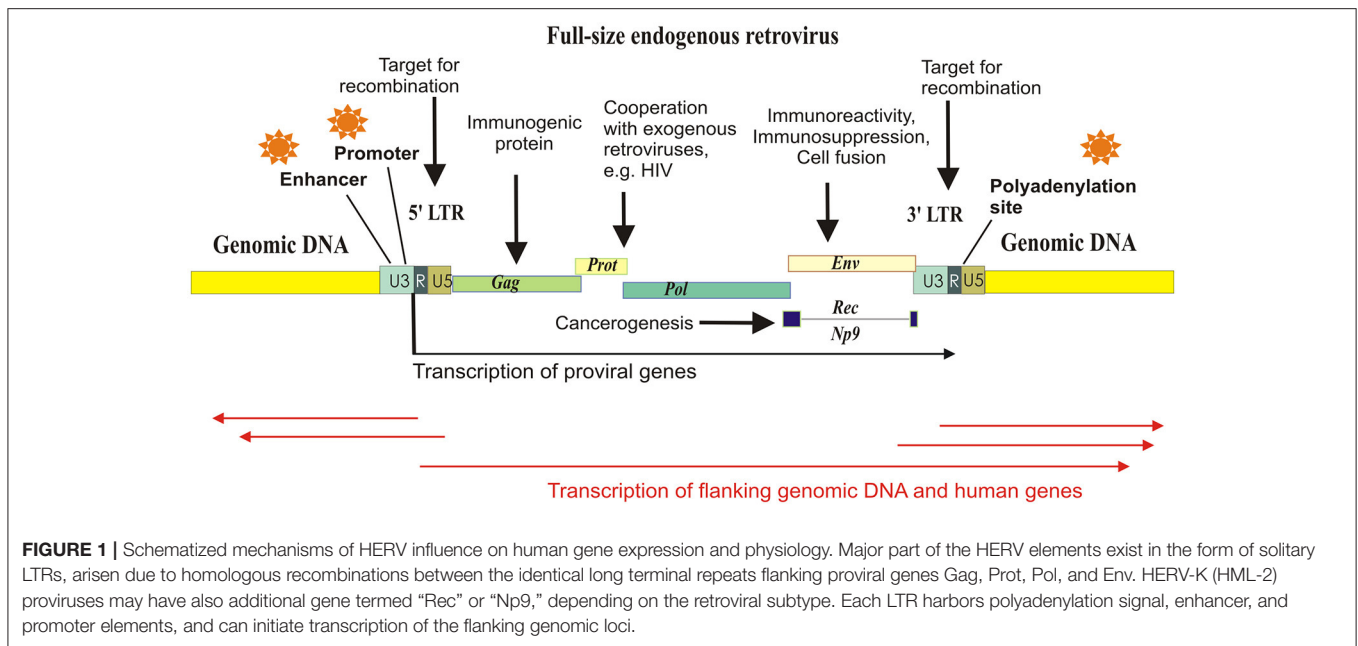## STRUCTURE OF LTR AND BINDING OF NUCLEAR PROTEINS

Most of the HERVs exist in the form of solitary LTRs. The LTRs include promoter elements, enhancers, transcriptional factor binding sites, splice sites, and polyadenylation signals, and are thought to serve as the major transcriptional regulators of HERVs. LTRs specifically bind host cell nuclear proteins (Trubetskoy et al., 2002) and serve in the following five pathways of human transcriptional regulation: (i) LTRs may have enhancer/repressor activities (Domansky et al., 2000; Hughes and Coffin, 2004; Ruda et al., 2004; Suntsova et al., 2013); (ii) LTRs may be promoter active; (iii) LTR may provide polyadenylation sites to terminate read-through transcripts; (iv) LTRs may provide splice sites; (v) LTRs may regulate host genes by RNA interference (Gogvadze et al., 2009).

Mapping DNaseI hypersensitivity sites (DHS) is the method of choice for the high-throughput identification of the regulatory genomic regions. Similarly, transcription factor binding sites (TFBS) denote fragments of DNA with nuclear protein binding

capacities (Ho et al., 2012). We combined investigation of both DHS and TFBS content of HERVs on a genomic scale (Garazha et al., 2015). To this end, we annotated all the genomic copies of HERVs and devised a bioinformatic algorithm mapping relevant TFBS and DHS features. For the entire set of HERVs, ~140,000 individual inserts (~19%) had at least one DHS and ~110,000 inserts (~15%)—at least one TFBS. Totally, there were identified ~155,000 and ~320,000 HERV-related DHS and TFBS, respectively (Garazha et al., 2015). This directly evidences potential implication of HERVs in the regulation of thousands of human genes. This is also in line with the previous finding that ~30% of all p53 binding sites localized by chromatin immunoprecipitation approach in the human genome fall within the HERV elements (Wang et al., 2007). Finally, as much as ~31.4% of all human transcription start sites were mapped within various transposable elements, including the HERVs (Faulkner et al., 2009).

All the 504 known HERV groups were characterized with regard to their TFBS content and showed very different results (available at http://herv.pparser.net/TotalStatistic.php). The families differed in their copy number, ranging from several copies as for the HERV-F, to more than 22,000 members as for the THE1B family. The total number of TFBS was also strikingly different—from zero (LTR5, LTR7A) to ~13,000 (MLT1K). The densities of TFBS also varied among the families. This is also important to quantitate absolute numbers of TFBS in each family. For example, the LTR12 family had the biggest proportion of TFBS-positive members and donated ~1,300 TFBS to the human DNA, whereas the family MLT1K contributed the greatest number of TFBS (~13,000), but had a small occurrence of TFBS-positive members. Interestingly, the TFBS and DHS tended to commonly appear in the same HERV elements. The probability that a particular element had DHS, was proportionate to the number of TFBS mapped herein (Garazha et al., 2015). Although, these findings provide clues for identification and functional annotation of multiple previously unknown human regulatory sequences, they are most likely still an underestimation of the HERV-generated TFBS pool. The repetitive nature of HERVs in many cases did not allow to directly attribute TFBS or DHS to any particular HERV element (Garazha et al., 2015).

Importantly, all the interrogated transcriptional factors had TFBS in the HERVs. This can explain extremely diverse and sometimes strongly tissue-specific influence of the different HERVs on the gene expression. For example, the LTR of the most recent HERV family HERV-K (HML-2) containing many human-specific and even polymorphic members, shows very high promoter and enhancer activities in the human germ cells and the corresponding tumors (seminoma), being transcriptionally silent in the other tissues (Domansky et al., 2000; Ruda et al., 2004). The promoter activity of the HERV-K (HML-2) inserts also provided the first evidence for the human specific antisense regulation of gene expression (Gogvadze et al., 2009). The human-specific LTRs located in the introns of genes *SLC4A8* and *IFT172* (for sodium bicarbonate cotransporter and intraflagellar transport protein

FIGURE 1 | Schematized mechanisms of HERV influence on human gene expression and physiology. Major part of the HERV elements exist in the form of solitary LTRs, arisen due to homologous recombinations between the identical long terminal repeats flanking proviral genes Gag, Prot, Pol, and Env. HERV-K (HML-2) proviruses may have also additional gene termed "Rec" or "Np9," depending on the retroviral subtype. Each LTR harbors polyadenylation signal, enhancer, and promoter elements, and can initiate transcription of the flanking genomic loci.

172, respectively) can *in vivo* generate transcripts that are reverse-complementary to the exons of those genes. Overexpression of the antisense transcripts resulted in approximately three- to four-fold decrease in mRNA levels for these genes (Gogvadze et al., 2009).

The HERVs can also provide polyadenylation signals for the regulation of gene expression. For example, mRNA for 8-kDa human protein similar to transcription factor GON4L is polyadenylated using the HERV-K (HML-2) LTR sequence (Baust et al., 2000). Another human transcription factor gene, *ZNF195*, utilizes the HERV-F LTR as the alternative polyadenylation site (Kjellman et al., 1999).

FUNCTIONAL INTERPLAY OF HERVs AND HUMAN GENOME

Expression of HERVs is tightly controlled by the host cell because it may be deleterious. Even the physical presence of the repetitive sequences in the genome can generate genomic instability due to homologous recombination between the HERV elements. HERVs can bias normal gene regulatory networks (Suntsova et al., 2015; reviewed by Rebollo et al., 2012). Expression of HERV proteins may result in dangerous inflammatory or immunosuppressive effects (Cho et al., 2008). In mammals, endogenous retroviruses are transcriptionally repressed using the KRAB domain zinc finger proteins and their cofactor TRIM28, which recruit methylation machinery to HERV copies (Turelli et al., 2014). In embryonic cells, a zinc finger protein Yin Yang 1 may serve as another repressor of HERV transcription by suppressing promoter activities of the LTRs (Schlesinger et al., 2013). Besides DNA methylation, histone modification is considered an alternative mechanism of endogenous retroviral repression in embryonal stem cells with the proteins SETDB1

(methyltransferase responsible for H3K9 trimethylation) and H3K4 demethylase LSD1/KDM1A involved (reviewed by Rebollo et al., 2012).

APOBEC3 protein family has another function in suppression of HERVs and retroviruses. APOBEC3G (hA3G) inhibits the retroviruses by entering viral particles and inducing hypermutation of viral genome during reverse transcription, leading to G to A substitutions (Bae and Jung, 2014). In concert, the protein hA3F induces viral hypermutation by deaminating minus-strand of viral cDNA during reverse transcription (Bae and Jung, 2014). Taken together, these factors induce epigenetic silencing and hypermutation of HERVs. Indeed, the LTRs have a bigger mutation rate than the rest of non-coding fraction of the human genome (Romano et al., 2006).

Conversely, the content of TFBS among the HERVs decreases with their evolutionary age (Garazha et al., 2015). For the heavily mutated, highly diverged (>20%) HERV elements, this content is approximately six-fold lower compared to the top evolutionary young elements. This observation may suggest that genomic "domestication" of HERVs involved reformatting of the active TFBS profiles and their further "standardization" upon accumulation of mutations, until they get equilibrated with the rest of non-coding DNA (Garazha et al., 2015). However, this type of analysis can be biased by the higher fragmentation in the evolutionary older HERVs, because each fragment is considered as an independent element. Further studies are, therefore, needed to explore the TFBS accumulation trends in linkage to the evolutionary dynamics of the human genome.

Sometimes co-evolution with the human genome resulted in a recruitment of certain HERV regulatory modules by the host organism (**Table 1**). The best-known example is the acquisition of salivary expression of the carbohydrate digestive enzyme amylase

TABLE 1 | Implication of HERV transcriptional regulation in human physiology and pathology (selected examples).

HERV element	Function	Mechanism	References
NORMAL PHYSIOLOGY			
HERV-E	Expression of amylase genes *AMY1A, AMY1B, AMY1C* in salivary glands	Creates tissue-specific enhancer	Ting et al., 1992
HERV-K (HML-2)	Expression of proline dehydrogenase gene *PRODH* in hippocampus	Creates tissue-specific enhancer	Suntsova et al., 2013
HERV-H	Maintaining pluripotency in stem cells	Recruits transcriptional activators by initiating transcription of intergenic RNAs	Ohnuki et al., 2014
HUERS-P1	Maintaining pluripotency in stem cells	Promotes transcription of a non-coding RNA serving as a molecular sponge for miR let-7 microRNAs	Durruthy-Durruthy et al., 2015
MER39	Expression of Prolactin during pregnancy	Creates tissue-specific promoter	Emera et al., 2012
HERV9	Control of fetal and adult expression of globin locus	Recruits transcriptional factors to the downstream Beta-globin promoter	Tuan and Pi, 2014
HERV-K (HML-2)	Control of *SLC4A8* and *IFT172* gene expression	Promotes negative regulator antisense RNAs	Gogvadze et al., 2009
HERV-W	Fusion of trophoblast cells in placenta	Encodes protein Syncytin	Frendo et al., 2003
CANCER			
MaLR LTR	Survival of Hodgkin's lymphoma cells by upregulation of *CSF1R* gene	Creates alternative promoter	Lamprecht et al., 2010
HERV-K (HML-2)	Survival of Chronic Lymphocytic Leukemia cells	Encodes protein NP9 with possible oncogenic functions	Fischer et al., 2014
HERV-W	Tumor growth and metastasis via immunosuppression	Encodes protein Syncytin	Kassiotis, 2014
INFECTIOUS DISEASES			
HERV-L/HERV16	Suppression of Varicela zoster virus and HIV infection	Creates HLA Complex P5 gene	Crosslin et al., 2015
HERV-K	dUTPase activity for HIV life cycle	Encodes endogenous retroviral dUTPase	Mayer and Meese, 2003
AUTOIMMUNE DISEASES			
HERV-E	Possible role in promotion of systemic lupus erythematosus (SLE)	Encodes potentially immunogenic retroviral proteins	Wu et al., 2015
HERV-K10	Possible role in promotion of rheumatoid arthritis	Encodes potentially immunogenic retroviral protein HERV-K10 Gag	Nelson et al., 2014
HERV-W	Possible role in promotion of osteoarthritis	Encodes potentially immunogenic retroviral protein Syncytin	Bendiksen et al., 2014
HERV-K18	Possible role in promotion of osteoarthritis	Encodes potentially immunogenic retroviral proteins	Garcia-Montojo et al., 2013
NEUROLOGICAL DISORDERS			
Multiple HERVs	Possible role in promotion of multiple sclerosis	Encode potentially immunogenic proteins and induce autoimmunoreactiviry	Libbey et al., 2014; Manghera et al., 2015
HERV-W	Possible role in promotion of schizophrenia and bipolar disorder	Encodes potentially immunogenic retroviral protein Syncytin	Diem et al., 2012
HERV-K (HML-2)	Possible role in regulation of proline dehydrogenase in schizophrenia	Creates tissue-specific enhancer for gene *PRODH*	Suntsova et al., 2013
HERV-H	Induction of hypotonia and motor, language, and cognitive delays	Due to recombinations, mediate 3q13.2-q13.31 deletions	Shuvarikov et al., 2013

from a HERV element inserted in the common ancestor of great apes (Ting et al., 1992).

On the other hand, HERV-H is a family expressed preferentially in human embryonal stem cells. Surprisingly, these are the HERV-H LTRs that appeared to be the primary mediators of cell fate reprogramming using famous "Yamanaka cocktail" (by overexpressing OCT3/4, SOX2, and KLF4 proteins), due to regulatory HERV-H-driven intergenic non-coding RNAs that help to recruit the transcriptional activator genes by serving as the scaffold (Ohnuki et al., 2014). Another human long non-coding RNA (human pluripotency-associated transcript 5, *HPAT5*) derived from both a HERV element HUERS-P1 and an Alu retrotransposon, was shown to promote pluripotency by functioning as a molecular sponge for the let-7 family of

microRNAs (Durruthy-Durruthy et al., 2015; Chuong et al., 2017).

The element MER39 forms an endometrium-specific promoter that regulates expression of Prolactin during pregnancy (Emera et al., 2012). The developmental switch from fetal to adult beta-globin gene expression in human is controlled by a copy of HERV9 element (Tuan and Pi, 2014). In hippocampus, transcription of gene *PRODH* is regulated human-specifically by a HERV-K (HML-2) LTR (Suntsova et al., 2013). *PRODH* metabolizes neuromediator molecules and has a strong implication in higher nervous activity and neurological disorders, and its deregulation might have an important impact on human evolution (Suntsova et al., 2013).

HERV-MEDIATED REGULATION OF GENE EXPRESSION IN PATHOLOGY

Proliferative Disorders

Recent findings indicate that HERV-mediated control of gene expression may be involved in various human diseases including cancer (Kassiotis, 2014). The role of HERVs in cancer is most likely limited to regulation of gene expression (Hohn et al., 2013). The data from cancer genome sequencing identified over 180 somatic insertions caused by LINE-1 retrotransposon activity, vs. only a single integration of a short HERV fragment, most likely replicated due to microhomology-mediated DNA repair mechanism (Lee et al., 2012). Many HERVs are abnormally expressed in cancer. For instance, HERV-K (HML-2) elements are up to ~3,000 times overexpressed in germ cell tumors and in melanoma (Buzdin et al., 2006b; Schmitt et al., 2013). Upregulation of HERVs can be mediated by either biased content of the specific transcription factors or by disruption of the anti-retroviral suppression mechanisms, such as aberrant demethylation (Conti et al., 2016) and decreased expression of APOBEC3 proteins (Shepelin et al., 2016). HERVs, in turn, may promote cellular transformation by regulating downstream human genes. For example, a demethylated copy of MaLR LTR can act as an alternative promoter to transcriptionally derepress the gene *CSF1R*, encoding colony stimulating factor-1 receptor, which is linked with survival of the Hodgkin's lymphoma cells (Lamprecht et al., 2010). More examples can be found in the other specific reviews (Babaian and Mager, 2016; Gonzalez-Cao et al., 2016; Anwar et al., 2017).

Infectious Diseases

The evolution of human pathogens might generate mechanisms involving transcriptional interactions of endogenous and exogenous retroviruses. For example, in HIV-infected patients, the HERV-K (HML-2) proviruses are expressed in peripheral blood mononuclear cells at higher levels compared to the non-infected individuals (Bhardwaj et al., 2014). The antibodies against HERV-K (HML-2) Env protein in blood were proposed as the new biomarker of HIV-1 infection, because HIV-1 can upregulate expression of a fully N-glycosylated HERV-K (HML-2) envelope protein on the cell surface (Michaud et al., 2014). Moreover, the HERV-K (HML-2)-specific T-cells from the HIV-1 infected patients *in vitro* completely eliminated the human cells infected with a panel of globally diverse HIV isolates. The mechanism of HIV-1 induced activation of human transposable elements possibly involves the activity of an HIV-1 Tat protein (Jones et al., 2013). Recent studies showed that out of 91 annotated HERV-K (HML-2) proviruses, Tat could activate expression of 26 proviruses, silenced 12, and did not change the expression of the others (Gonzalez-Hernandez et al., 2014). In addition, HIV infection may cause transactivation of HERV-W elements with their Env genes and Syncytin (Uleri et al., 2014). However, a controversial data were reported on the presence of HERV-K (HML-2) viral particles in the plasma of HIV-infected patients—higher levels of HERV-K (HML-2) RNA were detected in the HIV patients from Uganda, but not from the USA (Li et al., 2013). Of note, the recent association

study showed that susceptibility to infection with varicella zoster virus is linked with the non-coding gene HLA Complex P5 in the major histocompatibility complex. This gene is a copy of an endogenous retrovirus that may have a potential to suppress viral activity through indirect regulatory mechanisms. In previous studies, particular genetic variants of this region were associated with delay in development of AIDS in HIV-infected individuals (Crosslin et al., 2015).

Autoimmunity

The biased expression of HERVs is considered as one of the triggers of autoimmune disorders (Suntsova et al., 2015), which is evidenced by increased proviral RNA levels (Ehlhardt et al., 2006) and anti-HERV protein antibodies in sera from several types of patients (Bannert and Kurth, 2004). Immune reactivity against ERV proteins can be experimentally induced in mice and non-human primates, evidencing that immunological tolerance to endogenous retroviral products is not complete (Kassiotis, 2014). The HERV overexpression may be linked with massive DNA hypomethylation as seen for T-cells in systemic lupus erythematosus (SLE) patients (Wu et al., 2015).

Compared to the normal controls, in the patients with rheumatoid arthritis, increased antibody response was detected against the HERV-K10 Gag protein (Nelson et al., 2014). HERV-W transcripts and protein isoforms of Syncytin were overexpressed in cartilage of osteoarthritis patients (Bendiksen et al., 2014). In osteoarthritis, the patient's individual disease severity index was correlated with the expression of HERV-K18 provirus (Garcia-Montojo et al., 2013). However, inflammatory diseases may be also associated with the decreased expression of HERVs (**Table 1**).

Neurological Diseases

Expression of HERVs may serve as the biomarker for various neurological diseases (**Table 1**). For example, the HERV expression may be inducible in human astrocytes and neurons under inflammatory conditions in an IFNγ-dependent manner (Manghera et al., 2015). For multiple sclerosis (MS), a hypothesis was proposed that HERV-encoded proteins can act as the powerful immune stimulators inducing disease progression following neurodegeneration (Libbey et al., 2014). Indeed, genetic variants in some genes restricting retroviral infections were statistically linked with the risk of getting MS, as shown for the TRIM5, TRIM22, and BST2 genes (Nexo et al., 2013).

The abnormally high levels of the HERV-W Env gene product were detected in the plasma of the patients with schizophrenia and bipolar disorder (Diem et al., 2012), and in the active lesions in multiple sclerosis (van Horssen et al., 2016) and in the biopsies from the chronic inflammatory demyelinating polyradiculoneuropathies (Faucard et al., 2016). The increased expression of endogenous HERV-K (HML-2) proviral Env gene, in turn, may contribute to the development of amyotrophic lateral sclerosis by inducing neurodegeneration (Li et al., 2015). Finally, HERVs may also cause neurological disorders due to HERV-linked genomic rearrangements (**Table 1**). The human-specific enhancer activity of a HERV-K (HML-2) provirus on schizophrenia-associated gene *PRODH* may be another

active mechanism of HERV involvement in schizophrenia (Suntsova et al., 2013). Recently, a link was discovered between schizophrenia risk and the complement C4 system (Sekar et al., 2016). The individuals having a polymorphic HERV intronic insertion have elevated C4 expression, which in turn may cause neuronal synapse over-pruning, a phenotype that is associated with schizophrenia. Although this evidence is still indirect, this case is intriguing in light of previous observations of an association between schizophrenia and elevated ERV transcriptional activity (Chuong et al., 2017).

CONCLUSIONS

Taken together, these findings suggest that at the early stages after insertion, the HERV is treated by the host cells as a foreign genetic element, and is likely to be suppressed by the targeted methylation and mutations. However, at the later stages, when significant number of mutations has been already accumulated and when the retroviral genes are broken, the regulatory potential of a HERV may be released and recruited

to modify the genomic balance of transcription factor binding sites. This process goes together with further accumulation and selection of mutations, which reshape the regulatory landscape of the human DNA. However, developmental reprogramming, stress or pathological conditions like cancer, inflammation and infectious diseases, can remove the blocks limiting expression and HERV-mediated host gene regulation. This, in turn, can dramatically alter the gene expression equilibrium and shift it to a newer state, thus further exacerbating the stressful or unstable situation.

AUTHOR CONTRIBUTIONS

All authors listed have made substantial, direct and intellectual contributions to the work, and approved it for publication.

FUNDING

REFERENCES

Anwar, S. L., Wulaningsih, W., and Lehmann, U. (2017). Transposable elements in human cancer: causes and consequences of deregulation. *Int. J. Mol. Sci.* 18:E974. doi: 10.3390/ijms18050974

Babaian, A., and Mager, D. L. (2016). Endogenous retroviral promoter exaptation in human cancer. *Mob. DNA* 7:24. doi: 10.1186/s13100-016-0080-x

Bae, E. H., and Jung, Y.-T. (2014). Comparison of the effects of retroviral restriction factors involved in resistance to porcine endogenous retrovirus. *J. Microbiol. Biotechnol.* 24, 577–583. doi: 10.4014/jmb.1312.12079

Bannert, N., and Kurth, R. (2004). Retroelements and the human genome: new perspectives on an old relation. *Proc. Natl. Acad. Sci. U.S.A.* 101, 14572–14579. doi: 10.1073/pnas.0404838101

Baust, C., Seifarth, W., Germaier, H., Hehlmann, R., and Leib-Mösch, C. (2000). HERV-K-T47D-Related long terminal repeats mediate polyadenylation of cellular transcripts. *Genomics* 66, 98–103. doi: 10.1006/geno.2000.6175

Belshaw, R., Pereira, V., Katzourakis, A., Talbot, G., Paces, J., Burt, A., et al. (2004). Long-term reinfection of the human genome by endogenous retroviruses. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4894–4899. doi: 10.1073/pnas.0307800101

Bendiksen, S., Martinez-Zubiavrra, I., Tummler, C., Knutsen, G., Elvenes, J., Olsen, E., et al. (2014). Human endogenous retrovirus W activity in cartilage of osteoarthritis patients. *Biomed. Res. Int.* 2014:698609. doi: 10.1155/2014/698609

Bhardwaj, N., Maldarelli, F., Mellors, J., and Coffin, J. M. (2014). HIV-1 infection leads to increased transcription of human endogenous retrovirus HERV-K (HML-2) proviruses *in vivo* but not to increased virion production. *J. Virol.* 88, 11108–11120. doi: 10.1128/JVI.01623-14

Buzdin, A. (2007). Human-specific endogenous retroviruses. *ScientificWorldJournal* 7, 1848–1868. doi: 10.1100/tsw.2007.270

Buzdin, A., Kovalskaya-Alexandrova, E., Gogvadze, E., and Sverdlov, E. (2006a). At least 50% of human-specific HERV-K (HML-2) long terminal repeats serve *in vivo* as active promoters for host nonrepetitive DNA transcription. *J. Virol.* 80, 10752–10762. doi: 10.1128/JVI.00871-06

Buzdin, A., Kovalskaya-Alexandrova, E., Gogvadze, E., and Sverdlov, E. (2006b). GREM, a technique for genome-wide isolation and quantitative analysis of promoter active repeats. *Nucleic Acids Res.* 34:e67. doi: 10.1093/nar/gkl335

Buzdin, A., Ustyugova, S., Khodosevich, K., Mamedov, I., Lebedev, Y., Hunsmann, G., et al. (2003). Human-specific subfamilies of HERV-K (HML-2) long terminal repeats: three master genes were active simultaneously during branching of hominoid lineages. *Genomics* 81, 149–156. doi: 10.1016/S0888-7543(02)00027-7

Campbell, I. M., Gambin, T., Dittwald, P., Beck, C. R., Shuvarikov, A., Hixson, P., et al. (2014). Human endogenous retroviral elements promote genome instability via non-allelic homologous recombination. *BMC Biol.* 12:74. doi: 10.1186/s12915-014-0074-4

Cho, K., Lee, Y.-K., and Greenhalgh, D. G. (2008). Endogenous retroviruses in systemic response to stress signals. *Shock* 30, 105–116. doi: 10.1097/SHK.0b013e31816a363f

Chuong, E. B., Elde, N. C., and Feschotte, C. (2017). Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.* 18, 71–86. doi: 10.1038/nrg.2016.139

Chuong, E. B., Rumi, M. A. K., Soares, M. J., and Baker, J. C. (2013). Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat. Genet.* 45, 325–329. doi: 10.1038/ng.2553

Conti, A., Rota, F., Ragni, E., Favero, C., Motta, V., Lazzari, L., et al. (2016). Hydroquinone induces DNA hypomethylation-independent overexpression of retroelements in human leukemia and hematopoietic stem cells. *Biochem. Biophys. Res. Commun.* 474, 691–695. doi: 10.1016/j.bbrc.2016.05.010

Crosslin, D. R., Carrell, D. S., Burt, A., Kim, D. S., Underwood, J. G., Hanna, D. S., et al. (2015). Genetic variation in the HLA region is associated with susceptibility to herpes zoster. *Genes Immun.* 16, 1–7. doi: 10.1038/gene.2014.51

Dewannieux, M., and Heidmann, T. (2013). Endogenous retroviruses: acquisition, amplification and taming of genome invaders. *Curr. Opin. Virol.* 3, 646–656. doi: 10.1016/j.coviro.2013.08.005

Diem, O., Schaffner, M., Seifarth, W., and Leib-Mosch, C. (2012). Influence of antipsychotic drugs on human endogenous retrovirus (HERV) transcription in brain cells. *PLoS ONE* 7:e30054. doi: 10.1371/journal.pone.0030054

Domansky, A. N., Kopantzev, E. P., Snezhkov, E. V., Lebedev, Y. B., Leib-Mosch, C., and Sverdlov, E. D. (2000). Solitary HERV-K LTRs possess bi-directional promoter activity and contain a negative regulatory element in the U5 region. *FEBS Lett.* 472, 191–195. doi: 10.1016/S0014-5793(00)01460-5

Durruthy-Durruthy, J., Sebastiano, V., Wossidlo, M., Cepeda, D., Cui, J., Grow, E. J., et al. (2015). The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming. *Nat. Genet.* 48, 44–52. doi: 10.1038/ng.3449

Ehlhardt, S., Seifert, M., Schneider, J., Ojak, A., Zang, K. D., and Mehraein, Y. (2006). Human endogenous retrovirus HERV-K (HML-2) Rec expression and transcriptional activities in normal and rheumatoid arthritis synovia. *J. Rheumatol.* 3, 316–323.

Emera, D., Casola, C., Lynch, V. J., Wildman, D. E., Agnew, D., and Wagner, G. P. (2012). Convergent evolution of endometrial prolactin expression in primates,

mice, and elephants through the independent recruitment of transposable elements. *Mol. Biol. Evol.* 29, 239–247. doi: 10.1093/molbev/msr189

Faucard, R., Madeira, A., Gehin, N., Authier, F. J., Panaite, P. A., and Lesage, C. (2016). Human endogenous retrovirus and neuroinflammation in chronic inflammatory demyelinating polyradiculoneuropathy. *EBioMedicine* 6, 190–198. doi: 10.1016/j.ebiom.2016.03.001

Faulkner, G. J., Kimura, Y., Daub, C. O., Wani, S., Plessy, C., Irvine, K. M., et al. (2009). The regulated retrotransposon transcriptome of mammalian cells. *Nat. Genet.* 41, 563–571. doi: 10.1038/ng.368

Fischer, S., Echeverria, N., Moratorio, G., Landoni, A. I., Dighiero, G., Cristina, J., et al. (2014). Human endogenous retrovirus np9 gene is over expressed in chronic lymphocytic leukemia patients. *Leuk. Res. Rep.* 3, 70–72.

Frendo, J. L., Olivier, D., Cheynet, V., Blond, J. L., Bouton, O., Vidaud, M., et al. (2003). Direct involvement of HERV-W Env glycoprotein in human trophoblast cell fusion and differentiation. *Mol. Cell. Biol.* 23, 3566–3574. doi: 10.1016/j.lrr.2014.06.005

Garazha, A., Ivanova, A., Suntsova, M., Malakhova, G., Roumiantsev, S., Zhavoronkov, A., et al. (2015). New bioinformatic tool for quick identification of functionally relevant endogenous retroviral inserts in human genome. *Cell Cycle* 14, 1476–1484. doi: 10.1080/15384101.2015.1022696

Garcia-Montojo, M., Varade, J., Villafuertes, E., de La Hera, B., Hoyas-Fernandez, J., Dominguez-Mozo, M. I., et al. (2013). Expression of human endogenous retrovirus HERV-K18 is associated with clinical severity in osteoarthritis patients. *Scand. J. Rheumatol.* 42, 498–504. doi: 10.3109/03009742.2013.779021

Gogvadze, E., Stukacheva, E., Buzdin, A., and Sverdlov, E. (2009). Human-specific modulation of transcriptional activity provided by endogenous retroviral insertions. *J. Virol.* 83, 6098–6105. doi: 10.1128/JVI.00123-09

Gonzalez-Cao, M., Iduma, P., Karachaliou, N., Santarpia, M., Blanco, J., and Rosell, R. (2016). Human endogenous retroviruses and cancer. *Cancer Biol. Med.* 13, 483–488. doi: 10.20892/j.issn.2095-3941.2016.0080

Gonzalez-Hernandez, M. J., Cavalcoli, J. D., Sartor, M. A., Contreras-Galindo, R., Meng, F., Dai, M., et al. (2014). Regulation of the human endogenous retrovirus K (HML-2) transcriptome by the HIV-1 Tat protein. *J. Virol.* 88, 8924–8935. doi: 10.1128/JVI.00556-14

Ho, B., Baker, P. M., Singh, S., Shih, S.-J., and Vaughan, A. T. (2012). Localized DNA cleavage secondary to genotoxic exposure adjacent to an Alu inverted repeat. *Genes Chromosomes Cancer* 51, 501–509. doi: 10.1002/gcc.21938

Hohn, O., Hanke, K., and Bannert, N. (2013). HERV-K (HML-2), the best preserved family of HERVs: endogenization, expression, and implications in health and disease. *Front. Oncol.* 3:246. doi: 10.3389/fonc.2013.00246

Hughes, J. F., and Coffin, J. M. (2004). Human endogenous retrovirus K solo-LTR formation and insertional polymorphisms: implications for human and viral evolution. *Proc. Natl. Acad. Sci. U.S.A.* 101, 1668–1672. doi: 10.1073/pnas.0307885100

Jones, R. B., Song, H., Xu, Y., Garrison, K. E., Buzdin, A. A., Anwar, N., et al. (2013). LINE-1 retrotransposable element DNA accumulates in HIV-1-infected cells. *J. Virol.* 87, 13307–13320. doi: 10.1128/JVI.02257-13

Kassiotis, G. (2014). Endogenous retroviruses and the development of cancer. *J. Immunol.* 192, 1343–1349. doi: 10.4049/jimmunol.1302972

Kjellman, C., Sjögren, H. O., Salford, L. G., and Widegren, B. (1999). HERV-F (XA34) is a full-length human endogenous retrovirus expressed in placental and fetal tissues. *Gene* 239, 99–107. doi: 10.1016/S0378-1119(99)00372-8

Lamprecht, B., Walter, K., Kreher, S., Kumar, R., Hummel, M., Lenze, D., et al. (2010). Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. *Nat. Med.* 16, 571–579. doi: 10.1038/nm.2129

Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L. J., et al. (2012). Landscape of somatic retrotransposition in human cancers. *Science* 337, 967–971. doi: 10.1126/science.1222077

Li, L., Deng, X., Linsuwanon, P., Bangsberg, D., Bwana, M. B., Hunt, P., et al. (2013). AIDS alters the commensal plasma virome. *J. Virol.* 87, 10912–10915. doi: 10.1128/JVI.01839-13

Li, W., Lee, M. H., Henderson, L., Tyagi, R., Bachani, M., Steiner, J., et al. (2015). Human endogenous retrovirus-K contributes to motor neuron disease. *Sci. Transl. Med.* 7:307ra153. doi: 10.1126/scitranslmed.aac8201

Libbey, J. E., Cusick, M. F., and Fujinami, R. S. (2014). Role of pathogens in multiple sclerosis. *Int. Rev. Immunol.* 33, 266–283. doi: 10.3109/08830185.2013.823422

Maliniemi, P., Vincendeau, M., Mayer, J., Frank, O., Hahtola, S., Karenko, L., et al. (2013). Expression of human endogenous retrovirus-w including syncytin-1 in cutaneous T-cell lymphoma. *PLoS ONE* 8:e76281. doi: 10.1371/journal.pone.0076281

Manghera, M., Ferguson, J., and Douville, R. (2015). ERVK polyprotein processing and reverse transcriptase expression in human cell line models of neurological disease. *Viruses* 7, 320–332. doi: 10.3390/v7010320

Mayer, J., and Meese, E. U. (2003). Presence of dUTPase in the various human endogenous retrovirus K (HERV-K) families. *J. Mol. Evol.* 57, 642–649. doi: 10.1007/s00239-003-2514-6

Michaud, H. A., de Mulder, M., SenGupta, D., Deeks, S. G., Martin, J. N., Pilcher, C. D., et al. (2014). Trans-activation, post-transcriptional maturation, and induction of antibodies to HERV-K (HML-2) envelope transmembrane protein in HIV-1 infection. *Retrovirology* 11:10. doi: 10.1186/1742-4690-11-10

Nelson, P. N., Roden, D., Nevill, A., Freimanis, G. L., Trela, M., Ejtehadi, H. D., et al. (2014). Rheumatoid arthritis is associated with IgG antibodies to human endogenous retrovirus gag matrix: a potential pathogenic mechanism of disease? *J. Rheumatol.* 41, 1952–1960. doi: 10.3899/jrheum.130502

Nexo, B. A., Hansen, B., Nissen, K. K., Gundestrup, L., Terkelsen, T., Villesen, P., et al. (2013). Restriction genes for retroviruses influence the risk of multiple sclerosis. *PLoS ONE* 8:e74063. doi: 10.1371/journal.pone.0074043

Ohnuki, M., Tanabe, K., Sutou, K., Teramoto, I., Sawamura, Y., Narita, M., et al. (2014). Dynamic regulation of human endogenous retroviruses mediates factor-induced reprogramming and differentiation potential. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12426–12431. doi: 10.1073/pnas.1413299111

Rebollo, R., Romanish, M. T., and Mager, D. L. (2012). Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu. Rev. Genet.* 46, 21–42. doi: 10.1146/annurev-genet-110711-155621

Romano, C. M., Ramalho, R. F., and Zanotto, P. M., de A. (2006). Tempo and mode of ERV-K evolution in human and chimpanzee genomes. *Arch. Virol.* 151, 2215–2228. doi: 10.1007/s00705-006-0792-1

Ruda, V. M., Akopov, S. B., Trubetskoy, D. O., Manuylov, N. L., Vetchinova, A. S., Zavalova, L. L., et al. (2004). Tissue specificity of enhancer and promoter activities of a HERV-K (HML-2) LTR. *Virus Res.* 104, 11–16. doi: 10.1016/j.virusres.2004.02.036

Schlesinger, S., Lee, A. H., Wang, G. Z., Green, L., and Goff, S. P. (2013). Proviral silencing in embryonic cells is regulated by Yin Yang 1. *Cell Rep.* 4, 50–58. doi: 10.1016/j.celrep.2013.06.003

Schmitt, K., Reichrath, J., Roesch, A., Meese, E., and Mayer, J. (2013). Transcriptional profiling of human endogenous retrovirus group HERV-K (HML-2) loci in melanoma. *Genome Biol. Evol.* 5, 307–328. doi: 10.1093/gbe/evt010

Schumann, G. G., Gogvadze, E. V., Osanai-Futahashi, M., Kuroki, A., Münk, C., Fujiwara, H., et al. (2010). Unique functions of repetitive transcriptomes. *Int. Rev. Cell Mol. Biol.* 115–188. doi: 10.1016/B978-0-12-381047-2.00003-7

Sekar, A., Bialas, A. R., de Rivera, H., Davis, A., Hammond, T. R., Kamitaki, N., et al. (2016). Schizophrenia risk from complex variation of complement component 4. *Nature* 530, 177–183. doi: 10.1038/nature16549

Shepelin, D., Korzinkin, M., Vanyushina, A., Aliper, A., Borisov, N., Vasilov, R., et al. (2016). Molecular pathway activation features linked with transition from normal skin to primary and metastatic melanomas in human. *Oncotarget* 7, 656–670. doi: 10.18632/oncotarget.6394

Shuvarikov, A., Campbell, I. M., Dittwald, P., Neill, N. J., Bialer, M. G., Moore, C., et al. (2013). Recurrent HERV-H-mediated 3q13.2–q13.31 deletions cause a syndrome of hypotonia and motor, language, and cognitive delays. *Hum. Mutat.* 34, 1415–1423. doi: 10.1002/humu.22384

Suntsova, M., Garazha, A., Ivanova, A., Kaminsky, D., Zhavoronkov, A., and Buzdin, A. (2015). Molecular functions of human endogenous retroviruses in health and disease. *Cell. Mol. Life Sci.* 72, 3653–3675. doi: 10.1007/s00018-015-1947-6

Suntsova, M., Gogvadze, E. V., Salozhin, S., Gaifullin, N., Eroshkin, F., Dmitriev, S. E., et al. (2013). Human-specific endogenous retroviral insert serves as an enhancer for the schizophrenia-linked gene PRODH. *Proc. Natl. Acad. Sci. U.S.A.* 110, 19472–19477. doi: 10.1073/pnas.1318172110

Sverdlov, E. D. (2000). Retroviruses and primate evolution. *Bioessays* 22, 161–171. doi: 10.1002/(SICI)1521-1878(200002)22:2<161::AID-BIES7>3.0.CO;2-X

Ting, C. N., Rosenberg, M. P., Snow, C. M., Samuelson, L. C., and Meisler, M. H. (1992). Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. *Genes Dev.* 6, 1457–1465. doi: 10.1101/gad.6.8.1457

Trombetta, B., Fantini, G., D'Atanasio, E., Sellitto, D., and Cruciani, F. (2016). Evidence of extensive non-allelic gene conversion among LTR elements in the human genome. *Sci. Rep.* 6:28710. doi: 10.1038/srep28710

Trubetskoy, D. O., Zavalova, L. L., Akopov, S. B., and Nikolaev, L. G. (2002). Purification of proteins specifically binding human endogenous retrovirus K long terminal repeat by affinity elution chromatography. *J. Chromatogr. A* 976, 95–101. doi: 10.1016/S0021-9673(02)01236-0

Tuan, D., and Pi, W. (2014). In human β-globin gene locus, ERV-9 LTR retrotransposon interacts with and activates β-but not γ-globin gene. *Blood.* 124:2686.

Turelli, P., Castro-Diaz, N., Marzetta, F., Kapopoulou, A., Raclot, C., Duc, J., et al. (2014). Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements. *Genome Res.* 24, 1260–1270. doi: 10.1101/gr.172833.114

Uleri, E., Mei, A., Mameli, G., Poddighe, L., Serra, C., and Dolei, A. (2014). HIV Tat acts on endogenous retroviruses of the W family and this occurs via Toll-like receptor 4: inference for neuroAIDS. *AIDS* 28, 2659–2670. doi: 10.1097/QAD.0000000000000477

van Horssen, J., van der Pol, S., Nijland, P., Amor, S., and Perron, H. (2016). Human endogenous retrovirus W in brain lesions: rationale for targeted therapy in multiple sclerosis. *Mult. Scler. Relat. Disord.* 8, 11–18. doi: 10.1016/j.msard.2016.04.006

Wang, T., Zeng, J., Lowe, C. B., Sellers, R. G., Salama, S. R., Yang, M., et al. (2007). Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18613–18618. doi: 10.1073/pnas.0703637104

Wu, Z., Mei, X., Zhao, D., Sun, Y., Song, J., Pan, W., et al. (2015). DNA methylation modulates HERV-E expression in CD4+ T cells from systemic lupus erythematosus patients. *J. Dermatol. Sci.* 77, 110–116. doi: 10.1016/j.jdermsci.2014.12.004

Young, J. M., Whiddon, J. L., Yao, Z., Kasinathan, B., Snider, L., Geng, L. N., et al. (2013). DUX4 binding to retroelements creates promoters that are active in FSHD muscle and testis. *PLoS Genet.* 9:e1003947. doi: 10.1371/journal.pgen.1003947