

## COMPARATIVE EXPRESSION ANALYSIS OF HUMAN ENDOGENOUS RETROVIRUS ELEMENTS IN PERIPHERAL BLOOD OF CHILDREN WITH SPECIFIC LANGUAGE IMPAIRMENT

Minchev DS<sup>1,2,\*</sup>, Popov NT<sup>3</sup>, Naimov SI<sup>1</sup>, Minkov IN<sup>4</sup>, Vachev TI<sup>1</sup>

**\*Corresponding Author:** Assistant Professor Danail S. Minchev, Department of Medical Biology, Faculty of Medicine, Medical University-Plovdiv, 4000, Plovdiv, Bulgaria. Tel: +359-896-313-627. E-mail: dante17@abv.bg

### ABSTRACT

Specific language impairment (SLI) is a psychiatric condition with a complex etiology and a substantial genetic basis that affects children's verbal communication abilities. In this study, we examined the expression of five different human endogenous retrovirus elements (HERVs) in a cohort of 25 children with SLI and 25 healthy children in the control group. Human endogenous retrovirus elements, a diverse group of repetitive DNA sequences, can potentially cause considerable genetic heterogeneity. They had been integrated in the genome of our ancestors throughout evolution and now consist of about 8.0% of the human genome. Several HERV loci are transcribed in various cell types. Their expression in peripheral blood and in the brain is altered in many neurological and psychiatric diseases. To date, HERV expression profiles have never been studied in patients with SLI. This study aimed to elucidate differentially regulated human endogenous retroelements in peripheral blood of children with SLI, in comparison with healthy controls, through quantitative reverse transcription-polymerase chain reaction (qRT-PCR) methodology. Our results show that two genes: HERV-K (HLM-2) *gag* and HERV-P *env* were expressed at lower levels in the blood samples from SLI children in comparison with those in the control group.

**Keywords:** Blood-based quantitative reverse transcription-polymerase chain reaction (qRT-PCR); Human endogenous retrovirus elements (HERV); expression profiling; Relative quantitation; Specific language impairment (SLI).

### INTRODUCTION

The term "specific language impairment (SLI)" describes a condition affecting children's communication skills that are not caused by any organic damage of the brain, vocal apparatus and sensory systems, or by a pervasive disorder or another type of mental developmental disability [1]. Non-verbal intellectual capacity of the children with SLI is never altered in any way. This condition is relatively common, around 7.0% of the children in pre-school age are diagnosed with SLI [2]. The etiology of SLI has still not been fully elucidated. Many researchers suggest the involvement of genetic and environmental causes. Twin studies have shown that SLI is a condition with a strong genetic basis [3]. Currently, numerous molecular genetic studies, including, gene expression, candidate gene association analyses and genome-wide association studies (GWAS) have identified particular genes and chromosomal loci associated with the disease state, forkhead box P2 (*FOXP2*) and contactin-associated protein 2, (*CNTNAP2*) on chromosome 7; calcium-transporting ATPase 2C2 (*ATP2C2*), and C-MAF inducing protein (*CMIP*) on chromosome 16 [4-8].

From another point of view, human endogenous retrovirus elements (HERVs) are known to be associated with many neuropathological and non-neurological diseases with complex etiologies, such as autism, multiple sclerosis, schizophrenia and different types of cancer. Human endogenous retrovirus elements have derived from ancient in-

<sup>1</sup> Department of Plant Physiology and Molecular Biology, University of Plovdiv "Paisii Hilendarski," Plovdiv, Bulgaria

<sup>2</sup> Department of Medical Biology, Faculty of Medicine, Medical University-Plovdiv, 4000, Plovdiv, Bulgaria

<sup>3</sup> Psychiatric Ward for Active Treatment of Men, State Psychiatry Hospital Pazardzhik, Pazardzhik, Bulgaria

<sup>4</sup> Institute of Molecular Biology and Biotechnologies (IMBB), Plovdiv, Bulgaria

fectious retroviruses that, in many particular evolutionary stages, had integrated into the genome of the germline cells of our ancestors in evolution and started to be conveyed in a vertical fashion through many generations. At present, about 8.0% of the human genome is occupied by endogenous retrovirus elements [9]. The HERVs are composed of two long terminal repeats (LTRs) and four genes, *gag*, *pol*, *pro* and *env*. The *gag* gene encodes a group-specific antigen that serves as a retroviral capsid protein. Another gene, *pro*, codes for the viral protease, while *pol* contains a reverse transcriptase coding region. A HERV element differs from other LTR retrotransposons by the presence of the *env* gene encoding viral membrane proteins. Long terminal repeats contain many regulatory elements such as promoters, enhancers and polyadenylation signals required for retroviral gene expression [10, 11]. All human endogenous retroviruses nowadays are transposition-incompetent and carry various deletions. Despite the absence of apparent transposon activity, HERVs are classified as transposons based on the evidence of transposition during the past human evolution and on the existence of several animal ERVs that have retained their mobility within the host genomes [10, 12, 13].

Human endogenous retroviruses are divided into families, predominantly depending on the primer binding sites, sequences complementary to the 3' end of a cell tRNA. For example, HERV-K and HERV-W indicate that tRNA<sup>Lys</sup> and tRNA<sup>Trp</sup> are used as reverse transcription primers, respectively. This terminology is not related to the chronology of their discovery or their functions in the human genome [10, 14-16].

To date, HERV expression profiles have never been studied in patients with SLI. This study aims to elucidate differentially regulated human endogenous retroelements in peripheral blood of children with SLI, in comparison with healthy controls, through quantitative reverse transcription-polymerase chain reaction (qRT-PCR) methodology.

## MATERIALS AND METHODS

**Participants.** This study included a group of 25 children with SLI and 25 age- and sex-matched healthy controls. All participants were Bulgarians. None of them had received any medications before blood sampling. The SLI patients met the diagnostic criteria according to Diagnostic and Statistical Manual of Mental disorders, 4th edition [17]. Briefly, these criteria include: 1) language performance, evaluated using standard tests, is two standard deviations below the limit for the respective child's age; 2) verbal skills are at least one standard deviation below the non verbal IQ for the respective child's age; 3) there

are no observable neurological and sensory incapacities or physical impairments that directly affect use of spoken language, nor a pervasive developmental disorder; 4) differential diagnosis is made with respect to the receptive language disorder (comprehensive skills more than two standard deviations below the limit for the child's age) and expressive language disorder (comprehension is within the two standard deviations from the age level limit); 5) the language impairment affects the academic or occupational achievement or with social communication [17-18]. Children were assessed by a multi disciplinary team of child and adolescent psychiatrists and psychologists. Clinical assessment included a standard interview with the parents, free communication with the child and free play. All children who were clinically positive for language delay, vocabulary reduction and limited sentence structure as child language capacity is significantly below what is expected were referred for standard testing with age appropriate and standardized for the country test (HAWIK-R, Manova-Tomova test, Stanford-Binet Intelligence Scale). Clinical data we obtained from the patients were summarized as follows: mean IQ 85.67 ( $\pm$  12.669); verbal IQ 85.27 ( $\pm$  9.376) non verbal IQ 86.93 ( $\pm$  10.498). All patients who have shown significant difference between verbal and non-verbal development, have had no other developmental disorder signs [autism spectrum disorder (ASD), attention deficit hyperactivity disorder (ADHD), learning disorder (LD)], were included in the study. Patients with behavior problems such as oppositional, defiant and/or negative behavior, were not excluded.

From all the patients referred, four were excluded, two due to an IQ below 70 and two due to no difference between verbal IQ and non-verbal IQ. From all tested patients in 20, HAWIK-R was used, in three Manova-Tomova test and in two Stanford-Binet Intelligence Scale. The control group individuals were not tested with specialized developmental tests, just standard clinical testing.

**Ethics Statement.** Ethical admissibility of the study design and the informed consent form was confirmed by the Ethics Committee of Plovdiv Medical University. The Institutional Review Board approved the use of peripheral blood samples in this study. Written informed consent was obtained from the parents of the children in both study groups, with SLI and the typically developing children of the control group. All parents were informed of the aims and the procedures of this study prior to giving their informed consent. All personal information was kept in strict confidence.

**Blood Collection, Sample Storage and RNA Isolation.** A total amount of 2.5 mL peripheral blood from each participant (SLI or healthy) was drawn directly in a PAX-gene blood RNA monovette (PreAnalytiX GmbH, Hom-

brechtikon Switzerland), according to the manufacturer's protocol. All samples were kept at  $-20^{\circ}\text{C}$  and thawed at room temperature for 8 hours, prior to RNA extraction. All procedures required for total RNA isolation were performed using the PAXgene blood miRNA kit (PreAnalytiX GmbH, Feldbachstrasse Switzerland), according to the manual purification protocol. RNA quantity (absorption at 260 nm) and purity (A260/A280 ratio) were measured using the Epoch Micro-Volume Spectrophotometer System (BioTek, Winooski, VT, USA). Only A260/A280 ratios higher than 1.8 were considered acceptable.

**Quantitative RT-PCR Analysis of HERV Expression.** We analyzed the expression of five HERV genes: HERV-K (HLM-2) *gag*, HERV-K *env*, HERV-W *pol*, HERV-P *env*, and HERV-R *env*, using a relative qRT-PCR method. An amount of 1  $\mu\text{g}$  total RNA from each sample was subjected to DNase I treatment using RQ1 RNase-Free DNase (Promega, Madison, WI, USA). These DNase treated aliquots were then used for cDNA synthesis using Maxima First Strand cDNA Synthesis Kit (Thermo Scientific, Waltham, MA, USA) and oligo-dT/random hexamere primers. Reverse transcription reactions were set as follows: 11  $\mu\text{L}$  of total RNA, 4  $\mu\text{L}$  5  $\times$  reaction mix (with the oligo-dT/random hexamere primers), 2  $\mu\text{L}$  Maxima Enzyme Mix (Thermo Scientific), and nuclease free water to a final volume of 20  $\mu\text{L}$ . The qRT-PCR was performed using Maxima SYBR Green qPCR Master Mix (Thermo Scientific) in an ABI PRISM<sup>®</sup> 7500 real-time thermal cycler (Applied Biosystems, Waltham, MA, USA). Each sample was normalized using the *GAPDH* gene as endogenous control. Experiments were performed in duplicates. Each experiment was completed with a melting curve analysis to confirm the specificity of amplification and the lack of non specific amplification (Figure 1). Expression levels of HERV genes were obtained by relative quantification according to the  $2^{-\Delta\Delta\text{Ct}}$  method.

**Primer Sequences.** Individual cDNA samples from SLI (25) and healthy controls (25) were analyzed using a qRT-PCR assay. All primer pairs used for real-time PCR reactions are summarized in Table 1. The two separate HERV-K primer pairs produced amplicons located on different genome sites that are not related to each other.

## RESULTS

**Identification of Differentially Expressed HERV Genes in SLI.** Using the relative quantification real-time PCR approach described by Livak *et al.* [19], we measured the expression levels of all five HERV genes we studied: HERV-K (HLM-2) *gag*, HERV-K *env*, HERV-W *pol*, HERV-P *env*, and HERV-R *env*. Each of the HERV-specific primer pairs used in the procedure was tested on serial dilutions of pooled genomic DNA samples in order that we check the specificity of amplification. We chose the housekeeping gene *GAPDH* as an endogenous control based on its negligible overall variance in the samples. Additionally, we checked the Ct values we obtained for each gene in each cDNA sample. Only Ct differences no greater than 0.2 between the technical replicates of every sample were considered acceptable.

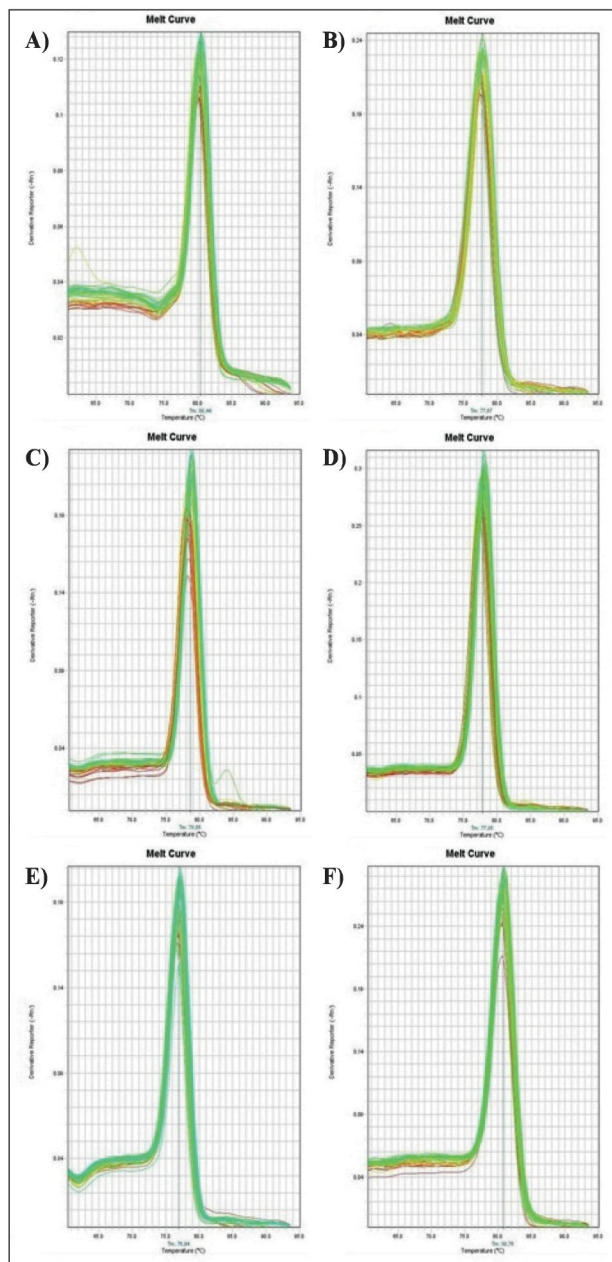
Calculated mean levels for the five HERV genes are shown in Figure 2. The data we obtained is presented as a relative fold difference ( $\log_2$ ) for each gene. Differences in expression between the two groups of individuals were evaluated using the Wilcoxon-Mann-Whitney test. The HERV-K (HML-2) *gag* and HERV-P *env* transcript levels appeared to be significantly lower (Mann-Whitney U test  $p$  value  $<0.05$ ) in the group of children with SLI compared with those in the control group (Figure 3). The other three genes studied: HERV-K *env*, HERV-R *env*, HERV-W *pol*, did not show statistically significant changes between the two groups. Statistical significance is defined as  $p$  value of  $<0.05$  combined with an absolute fold change of  $>1.4$ .

In order to determine whether the expression levels of HERV-K (HML-2) *gag* and HERV-P *env* can serve as potential discriminative biomarkers, we performed a receiver operating characteristic (ROC) analysis of the data from the qRT-PCR experiment. The ROC analysis is widely used statistical method for assessing the ability of a potential test marker to distinguish between disease carriers and healthy individuals. It provides a statistical model that presents the sensitivity (or the true positive fraction) of the marker as a function of the false positive fraction (defined as 1-specificity) of the same marker at multiple thresholds. Discriminative performance is

**Table 1.** The HERV-specific qRT-PCR primer pairs.

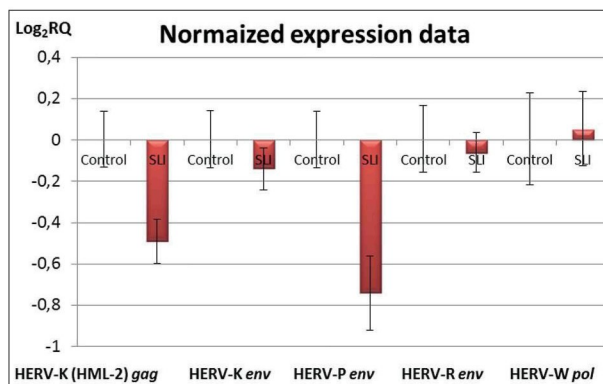
HERV Gene	Forward Primer (5'>3')	Reverse Primer (5'>3')
HERV (HML-2) <i>gag</i>	GGC CAT CAG AGT CTA AAC CAC G	CTG ACT TTC TGG GGG TGG CCG
HERV-K <i>env</i>	GGG TAC CTG GCC CCA TAG AT	CAT CAT CCC TTC TTC CTC AGG TT
HERV-W <i>pol</i>	TGA GTC AAT TCT CAT ACC TG	AGT TAA GAG TTC TTG GGT GG
HERV-P <i>env</i>	CAA GAT TGG GTC CCC TCA C	CCT ATG GGG TCT TTC CCT G
HERV-R <i>env</i>	CAT GGG AAG CAA GGG AAC T	CTT TCC CCA GCG AGC AAT AC

HERV: human endogenous retrovirus elements; qRT-PCR: quantitative reverse-transcription polymerase chain reaction.

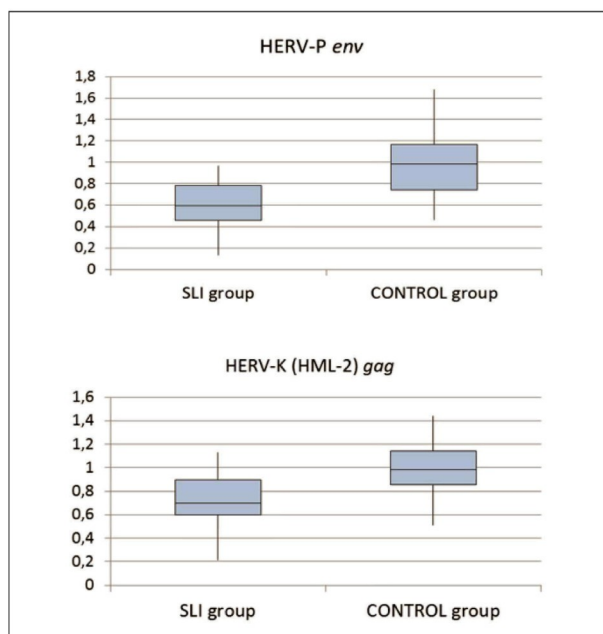


**Figure 1.** Melting (dissociation) curve analysis of the five HERV-encoded genes and the endogenous control used for normalization. A specific single peak was detected in the dissociation curves for all amplicons in every SLI and control sample. **Panel A:** HERV-K (HML-2) *gag* amplicons with a melting temperature  $80.4\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ . **Panel B:** HERV-K *env* with a melting temperature  $77.7\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ . **Panel C:** HERV-P *env* amplicons with a melting temperature  $78.7\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ . **Panel D:** HERV-R *env* with a melting temperature  $77.9\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ . **Panel E:** HERV-W *pol* amplicons with a melting temperature  $76.9\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ . **Panel F:** *GAPDH* with a melting temperature  $80.8\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ .

assessed by the area enclosed between the ROC curve and the X-axis of the graph, and this area is known as area under the curve (AUC). A greater AUC represents a



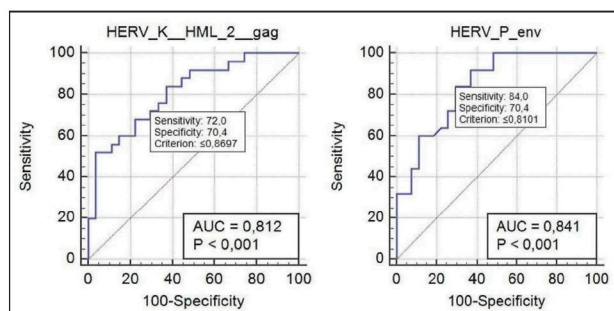
**Figure 2.** Differential expression of HERV in children with SLI (data presented in logarithmic scale). The qRT-PCR analysis of gene expression revealed that levels of HERV-K (HML-2) *gag* and HERV-P *env* were at least 1.4-times lower in the children with SLI compared to those in the control group (1.41- and 1.67-fold change for HERV-K (HML-2) *gag* and HERV-P *env*, respectively). Quantitative evaluation of the other three genes: HERV-K *env*, HERV-R *env* and HERV-W *pol*, found no significant changes in expression between the two groups.



**Figure 3.** Relative quantitation of differentially expressed HERV genes in SLI children and healthy controls. Data is presented in box plot diagrams with linear scale. The HERV levels in the control are normalized to 1.

more accurate discrimination between individuals with the disease and no disease. A single point on the ROC curve represents a particular sensitivity value for a given specificity. Since there is a 50.0% chance to predict the health status of an individual (disease or no disease) only by random guessing, a 50.0% discrimination accuracy of a given marker is possible, but statistically meaningless. Thus, the AUC value can never be lower than 0.5, and





**Figure 4.** Receiver operating characteristic curve analysis on differential expression data for HERV-K (HML-2) *gag* and HERV-P *env*. The ROC curves were calculated using MedCalc (<https://www.mecalc.org>) commercial statistical software. Random predictive accuracy (AUC value of 0.5 that corresponds to a 50.0% chance of discrimination by pure guessing) is denoted by a diagonal line. The ROC curves were plotted for HERV-K (HML-2) *gag* and HERV-P *env* that yielded 0.828 and 0.858 AUC values, respectively.

even the worst test marker can meet a 50.0% predictive accuracy. This random chance is often presented on the ROC curve plot as a dotted diagonal line. The ROC curve was plotted and the AUC values with corresponding 95% confidence intervals (95% CI) were calculated as follows: AUC 0.812 (95% CI: 0.696-0.928) for HERV-K (HML-2) *gag* and 0.841 (95% CI: 0.736-0.946) for HERV-P *env*. The ROC analysis was subsequently used to calculate diagnostic sensitivity and specificity. Calculated sensitivity and specificity of HERV-K (HML-2) *gag* at optimal cutoff were 72.0 and 70.4%, respectively. Moreover, HERV-P *env* showed higher sensitivity 84.0% and 70.4% specificity (Figure 4). Together, these results indicate that the observed statistically significant differences in expression of HERV-K (HML-2) *gag* and HERV-P *env*, can discriminate between SLI cases and healthy controls with considerable accuracy.

## DISCUSSION

To date, there have been no published data on HERV expression in the context of the SLI condition. To the best of our knowledge, this is the first cohort study to announce differences in transcript abundance of specific HERV genes in the blood of children with SLI compared with a control group. We have studied HERV transcript levels in peripheral blood of children with SLI. Throughout our study we have analyzed the expression pattern of five different HERV loci: HERV-K (HML-2) *gag*, HERV-K *env*, HERV-W *pol*, HERV-P *env* and HERV-R *env*. All of the five genes studied appeared to be ubiquitously expressed in the peripheral blood of children in both the SLI and control groups. Two of these, HERV-K (HML-2) *gag*

and HERV-P *env*, appeared to be differentially expressed between the two groups.

Despite the growing evidence of a strong genetic basis for SLI during the last decade, the precise genetic mechanisms of SLI remain elusive. Many researchers suggest that SLI is a multi causation condition with both genetic and environmental factors involved. Numerous molecular studies have established a set of SLI-associated genes whose single nucleotide polymorphism (SNP) profile or differential expression is of particular importance for SLI development [20]. Despite this large amount of research effort, no assessment of the HERV expression has been made in the aforementioned studies.

In some cases, HERV regulatory regions and HERV encoded products may play an important role in essential physiological processes. For example, syncytin-1 and syncytin-2 proteins that mediate trophoblast cell fusion during placental development, derive from HERV Envelope proteins. In a recent study, Bhat *et al.* [21] revealed that HERV-K Env is able to promote a considerable neuroprotective effect in a case of HIV infection. According to Thompson *et al.* [22], and Sundaram *et al.* [23], about 20.0% of all functionally active transcription factor binding sites in human and murine genomes are LTR-related.

We observed relevant changes in gene expression in terms of a small but statistically significant reduction in HERV-K (HML-2) *gag* (1.41-fold) and HERV-P *env* (1.67-fold) transcript levels. The HERV downregulation presumes an active suppression of these elements. As HERV expression is modulated by various mechanisms, the differential gene expression we observed is probably a result of many factors affecting HERV regulation. Regulation mechanisms of particular importance involve DNA methylation and several *trans*-acting transcription factors [24]. Generally, transcription of active HERV loci is mediated by a 5'-LTR core promoter able to support transcription at a basal level. Transcription is negatively regulated mainly by CpG methylation at the 5'-LTR. According to Wentzensen *et al.* [25], the expression of a specific member of the HERV-H in gastrointestinal cancer cells correlates with demethylation of the 5'-LTR. The presence of enhancer sequences within the 5'-LTRs has also been demonstrated. In addition, several *trans*-acting transcriptional factors, which bind to HERV promoters and enhancers, have been identified: YY1, Sp-1, Oct-1, GATA, GCM1 [24,26-28] Nevertheless, our RNA-seq data from a previous expression study on pooled SLI and ASD samples [29] suggests no changes in the expression levels of the aforementioned transcription factors and the most DNA-methyltransferases. On this basis, we can speculate that the observed differences in the HERV transcript levels can be due to negligibly small but specific changes in various gene regulation processes.

Further efforts are still needed for a complete characterization of HERV regulation.

The most important findings in our study were that HERV expression levels in blood correlate negatively with the SLI development and may specifically discriminate SLI patients from healthy controls. The expression profile we obtained may be useful for the investigation of many diseases of the central nervous system that are suspected to be associated with the retroviral activity.

While the patient cohort was very well characterized and representative for an SLI population, the total number of participants was not large enough to definitively create a model for prediction of SLI or characterize HERV expression differences that may be more highly associated with SLI subtypes. Due to the small sample size, analysis of the data with respect to medication, sex, special diet, race, ethnicity or other potential confounding covariates, was not conclusive.

Our findings are in agreement with several other studies in which HERV downregulation has been clearly established [30-33]. Alterations in HERV expression occur in various pathological conditions, in a wide variety of cancer conditions, in inflammation reactions, and in autoimmunity [30-33]. While a remarkable number of studies support the overexpression of HERVs in various complex conditions, our findings are entirely coherent with those in which significantly lower expression levels of specific HERV loci has also been reported [30-36].

In spite of the fact that a vast majority of human endogenous retrovirus elements are upregulated in malignant cells, in a very few cancer cases, their expression is suppressed. Syncytin-1 is described to be underexpressed in several tumor types, which suggest its potential beneficial effect. Syncytin-1 has been shown to be downregulated in pancreatic adenocarcinomas, while its stable expression in a melanoma cell line correlates with restricted cell growth [30].

Liang *et al.* [31] evaluated transcription of the HERV-related gene psiTPTE22-HERV in kidney, liver, lung, and stomach tumors in comparison with normal tissue samples. The authors prove that psiTPTE22-HERV transcripts were notably lower in all tumor tissues examined. Their results confirm that the expression of the HERV-related gene is suppressed solely in the cancer samples and that this suppression correlates with the malignancy of the cells [31].

According to Kowalczyk *et al.* [32], HERV-W *env* expression in peripheral blood mononuclear cell (PBMC) samples from patients with localized scleroderma, negatively correlates with the severity of the disease. Moreover, HERV-W *env* levels in skin samples from the patients were distinctively lower, compared to those in the control group samples [32].

Lattekivi *et al.* [33] performed a differential expression study of HERVs in psoriasis at a whole transcriptome level. The authors observed high expression of repetitive elements in both patients and healthy controls. Most of the endogenous retrovirus loci with significant changes in expression between the two studied groups appeared to be downregulated in psoriasis skin. The authors set a differential expression fold change thresholds at 0.71 and 1.41 and a false discovery rate at <0.01. They discovered downregulation of 17 HERV families in lesional and three HERV families in non-lesional psoriatic skin in comparison with healthy skin. Their findings clearly demonstrated the global repression of the endogenous retrovirus elements in psoriatic skin. The HERV downregulation has also been confirmed in psoriasis by Gupta *et al.* [34].

Weis *et al.* [35] discovered a substantial reduction in the HERV-W *gag* gene expression levels in brain samples from patients with schizophrenia and bipolar disorder. Along with a noticeable overexpression of HERV-H, Balastrieri *et al.* [36] discovered downregulation of HERV-W in peripheral blood samples of children with autistic spectrum disorder. In contrast to HERV-H, HERV-W was found to be downregulated in the autistic group compared to the healthy controls (with less than a 2-fold change). These results clearly suggest that HERV transcript levels can specifically characterize a complex psychiatric condition such as autism and serve as putative molecular markers for it. It is noteworthy to mention that despite the small absolute change, the reported HERV-W downregulation is still considered important in terms of statistical significance.

Our results together with the vast scientific data obtained by other research groups support that HERV genes play an important role in normal cell processes. Human endogenous retroviruses are not only “junk” or “selfish” DNA, but also the key factors and markers for many pathological conditions. Further clarification of the molecular mechanisms and pathways in which HERVs participate is needed for a complete understanding of their functions in norm and pathology.

In general, despite the huge body of evidence supporting overexpression of HERVs in many diseases, pathologies with significantly lower expression levels of HERVs have also been reported [30-36]. In this context, our study presents evidence of minor (>1.4-fold), but statistically reliable downregulation of two HERV-encoded genes. Our finding is consistent with previous studies in which low HERV expression has been described. Of particular interest is the fact that differential expression of HERV-K (HML-2) *gag* and HERV-P *env* is a sensitive marker that can discriminate children with SLI from healthy controls.

## ACKNOWLEDGMENTS

The authors are most grateful to the families of the probands for their collaboration. We thank Professor V. Stoyanova, Assistant Professor H. Ivanov and Professor I. Ivanov (Department of Pediatrics and Medical Genetics, Medical University of Plovdiv, Plovdiv, Bulgaria) for their technical insights and support of this study.

**Declaration of Interest.** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

**Funding.** This project received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 739582 (Project PlantaSYST).

## REFERENCES

- Leonard LB. Children with specific language impairment and their contribution to the study of language development. *J Child Lang.* 2014; 41(Suppl 1): 38-47.
- Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. *J Speech Lang Hear Res.* 1997; 40(6): 1245-1260.
- Bishop DV. What causes specific language impairment in children? *Curr Dir Psychol Sci.* 2006; 15(5): 217-221.
- Newbury DF, Monaco AP. Genetic advances in the study of speech and language disorders. *Neuron.* 2010; 68(2): 309-20.
- Pariani MJ, Spencer A, Graham JM, Rimoin DL A 785kb deletion of 3p14.1p13, including the FOXP1 gene, associated with speech delay, contractures, hypertonia and blepharophimosis. *Eur J Med Genet.* 2009; 52(2-3): 123-127.
- Carr CW, Moreno-De-Luca D, Parker C, Zimmerman HH, Ledbetter N, Martin CL, *et al.* Chiari I malformation, delayed gross motor skills, severe speech delay, and epileptiform discharges in a child with FOXP1 haplo-insufficiency. *Eur J Hum.* 2010; 18(11): 1216-1220.
- Horn D, Kapeller J, Rivera-Brugués N, Moog U, Lorenz-Depiereux B, Eck S, *et al.* Identification of FOXP1 deletions in three unrelated patients with mental retardation and significant speech and language deficits. *Hum Mutat.* 2010; 31(11): E1851-E1860.
- Vernes SC, MacDermot KD, Monaco AP, Fisher SE. Assessing the impact of FOXP1 mutations on developmental verbal dyspraxia. *Eur J Hum Genet.* 2009; 17(10):1354-1358.
- Gifford R, Tristem M. The evolution, distribution and diversity of endogenous retroviruses. *Virus Genes.* 2003; 26(3): 291-315.
- Shin W, Lee J, Son SY, Ahn K, Kim HS, Han K. Human-specific HERV-K insertion causes genomic variations in the human genome. *PLoS One.* 2013; 8(4): e60605.
- Flockerzi A, Ruggieri A, Frank O, Sauter M, Maldener E, Kopper B, *et al.* Expression patterns of transcribed human endogenous retrovirus HERV-K(HML-2) loci in human tissues and the need for a HERV Transcriptome Project. *BMC Genomics.* 2008; 9(4): 354 9: 354-370.
- Gifford WD, Pfaff SL, Macfarlan TS. Transposable elements as genetic regulatory substrates in early development. *Trends Cell Biol.* 2013; 23(5): 218-226.
- Ribet D, Harper F, Dupressoir A, Dewannieux M, Pierron G, Heidmann T. An infectious progenitor for the murine IAP retrotransposon: emergence of an intracellular genetic parasite from an ancient retrovirus. *Genome Res.* 2008; 18(4): 597-609.
- Jurka J. Repbase update: A data-base and an electronic journal of repetitive elements. *Trends Genet.* 2000; 16(9): 418-420.
- Tristem M. Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database. *J Virol.* 2000; 74(8): 3715-3730.
- Mager DL, Medstrand P. Retroviral repeat sequences. In: Cooper D, Editor. *Nature Encyclopedia of the Human Genome.* Hampshire, UK: MacMillan Publishers. 2003: 291-315.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (4th ed., text rev.).* Washington, DC, USA: American Psychiatric Association, 2000
- Asikainen M. *Diagnosing specific language impairment (Academic Dissertation).* University of Tampere, Finland, 2005. Available from: <https://trepo.tuni.fi/bitstream/handle/10024/67552/951-44-6455-9.pdf>.
- Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_t}$  method. *Methods.* 2001; 25(4): 402-408.

20. Rice ML. Language growth and genetics of specific language impairment. *Int J Speech Lang Pathol.* 2013; 15(3): 223-233.
21. Bhat RK, Rudnick W, Antony JM, Maingat F, Ellstad KK, Wheatley BM, *et al.* Human endogenous retrovirus-K(II) envelope induction protects neurons during HIV/AIDS. *PLoS One.* 2014; 9(7): e97984.
22. Thompson PJ, Macfarlan TS, Lorincz MC. Long terminal repeats: From parasitic elements to building blocks of the transcriptional regulatory repertoire. *Mol Cell.* 2016; 62(5): 766-776.
23. Sundaram V, Cheng Y, Ma Z, Li D, Xing X, Edge P, *et al.* Widespread contribution of transposable elements to the innovation of gene regulatory networks. *Genome Res.* 2014; 24(12): 1963-1976.
24. Sverdlov ED. *Retroviruses and Primate Genome Evolution*, 1st ed. Boca Raton, FL, USA: CRC Press, 2005.
25. Wentzensen N, Coy JF, Knaebel HP, Linnebacher M, Wilz B, Gebert J, *et al.* Expression of an endogenous retroviral sequence from the HERV H group in gastrointestinal cancers. *Int J Cancer.* 2007; 121(7): 1417-1423.
26. Li F, Karlsson H. Expression and regulation of human endogenous retrovirus W elements. *APMIS.* 2016; 124(1-2): 52-66.
27. Lee WJ, Kwun HJ, Kim HS, Jang KL. Activation of the human endogenous retrovirus W long terminal repeat by herpes simplex virus type 1 immediate early protein 1. *Mol Cells.* 2003; 15(1): 75-80.
28. Li F, Nellåker C, Sabunciyan S, Yolken RH, Jones-Brando L, Johansson AS, *et al.* Transcriptional derepression of the ERVWE1 locus following influenza A virus infection. *J Virol.* 2014; 88(8): 4328-4337.
29. Ivanov H, Popov N, Vachev T, Yordanova R, Pacheva I, Galabova F, *et al.* Differential expression of protein-coding genes in children with specific language impairment and autistic spectrum disorders. *Sci Res USB-Plovdiv Ser G.* 2017; 20(1): 13-16.
30. Grandi N, Tramontano E. HERV envelope proteins: Physiological role and pathogenic potential in cancer and autoimmunity. *Front Microbiol.* 2018; 9(1): 462-487.
31. Liang Q, Ding J, Xu R, Xu Z, Zheng S. The novel human endogenous retrovirus-related gene, psiTPTE22-HERV, is silenced by DNA methylation in cancers. *Int J Cancer.* 2010; 127(8): 1833-1843.
32. Kowalczyk MJ, Dańczak-Pazdrowska A, Szramka-Pawlak B, Żaba R, Osmola-Mańkowska A, Silny W. Human endogenous retroviruses and chosen disease parameters in morphea. *Postepy Dermatol Alergol.* 2017; 34(1): 47-51.
33. Lättekivi F, Köks S, Keermann M, Reimann E, Prans E, Abram K, *et al.* Transcriptional landscape of human endogenous retroviruses (HERVs) and other repetitive elements in psoriatic skin. *Sci Rep.* 2018; 8(1): 4358-4365.
34. Gupta R, Michaud HA, Zeng X, Debbaneh M, Arron ST, Jones RB, *et al.* Diminished humoral responses against and reduced gene expression levels of human endogenous retrovirus-K (HERV-K) in psoriasis. *J Transl Med.* 2014; 12(3): 256-263.
35. Weis S, Llenos IC, Sabunciyan S, Dulay JR, Isler L, Yolken R, *et al.* Reduced expression of human endogenous retrovirus (HERV)-W GAG protein in the cingulate gyrus and hippocampus in schizophrenia, bipolar disorder, and depression. *J Neural Transm (Vienna).* 2007; 114(5): 645-655.
36. Balestrieri E, Arpino C, Matteucci C, Sorrentino R, Pica F, Alessandrelli R, *et al.* HERVs expression in autism spectrum disorders. *PLoS One.* 2012; 7(11): e48831.