Elevated levels of human endogenous retrovirus-W transcripts in blood cells from patients with first episode schizophrenia

Y. Yao†, J. Schröder†, C. Nellåker†, C. Bottmer‡, S. Bachmann‡, R. H. Yolken§ and H. Karlsson*†

†Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden. ∗Section of Geriatric Psychiatry, University of Heidelberg, Heidelberg, Germany, and ‡The Stanley Division of Developmental Neurovirology, Johns Hopkins University, School of Medicine, Baltimore, MD, USA
*Corresponding author: H. Karlsson, PhD, Department of Neuroscience, Karolinska Institutet, Retzius väg 8, S-171 77 Stockholm, Sweden. E-mail: hakan.karlsson.2@ki.se

We previously reported on the differential presence of transcripts related to the human endogenous retrovirus (HERV)-W family in cerebrospinal fluid and plasma from patients with first-episode schizophrenia compared with control individuals. Whether this is a consequence of qualitative or quantitative differences in transcription of genomic regions harboring HERV-W elements is not known. The purpose of the present study was therefore to characterize the transcribed HERV-W elements in mononuclear cells obtained from 30 patients first hospitalized for schizophrenia-related psychosis and from 26 healthy control individuals. We observed elevated total levels of HERV-W gag (2.1-fold, \( P < 0.01 \)) but not env transcripts in the cells of patients compared with controls. By using the melting temperatures of the amplicons as a proxy marker for sequence identity, no absolute qualitative differences was detected between the two groups. Mapping of the detected transcripts identified several intronic and intergenic HERV-W elements transcribed in the cells, including elements previously considered transcriptionally silent. Element-specific assays revealed elevated levels of intronic transcripts containing HERV-W gag sequence from the putative gene PTD015 on chromosome 11q13.5 (1.6-fold, \( P < 0.05 \)) in the patients compared with the controls. Thus, studies aiming to further understanding of complex human disease such as schizophrenia may need to be extended beyond the strictly protein-coding fraction of the transcriptome.

Keywords: HERV-W, PTD015, recent-onset, real-time PCR, schizophrenia

Schizophrenia is a neuropsychiatric disorder of unknown etiology. An increasing number of studies suggest that both genetic and environmental factors contribute to the development of the disease. In the cerebrospinal fluids (CSFs) obtained from individuals first hospitalized for schizophrenia or schizoaffective disorder, we previously identified RNA sequences similar to retroviral \( \text{pol} \) genes in 29% of these individuals (Karlsson et al. 2001). All these sequences were related to known human endogenous retroviruses (HERVs), the majority being most closely related to the HERV-W family (Blond et al. 1999). Such sequences were less prevalent in CSF samples obtained from patients with a long-standing history of schizophrenia and were not detected in CSF samples from subjects with noninflammatory neurological diseases or healthy control individuals (Karlsson et al. 2001).

We subsequently reported that HERV-W-related \( \text{gag} \) transcripts were more prevalent also in plasma samples obtained from individuals with recent-onset schizophrenia or schizoaffective disorder compared with those from control individuals (Karlsson et al. 2004).

HERV-W is one of at least 31 different families of HERVs and exists in multiple copies in the human genome (Katzourakis & Tristem 2005). The vast majority of the HERV-W \( \text{gag}, \text{pol} \) and \( \text{env} \) genes have lost their ability to encode the different viral proteins. A HERV-W element on chromosome 7q21, however, contains an intact \( \text{env} \) gene encoding an envelope protein that has been denoted syncytin with a proposed role in placenta biogenesis (Blond et al. 2000; Mi et al. 2000). Aberrant expression of this particular \( \text{env} \) gene has previously been associated with both preeclampsia (Lee et al. 2003) and multiple sclerosis (Antony et al. 2004; Mameli et al. 2007; Perron et al. 2005), whereas post-mortem brains obtained from patients with schizophrenia or bipolar disorders appear to harbor normal levels of this transcript (Frank et al. 2005).

The U3 region of the 5’ retroviral long terminal repeat (LTR) contains the motifs necessary for regulation and initiation of transcription of the retroviral genes: \( \text{gag}, \text{pol} \) and \( \text{env} \). Studies of the levels of transcripts from the retroviral genes and promoter activities of different cloned HERV-W LTRs in various cell lines suggest cell-specific expression of HERV-W transcripts (Mallet et al. 2004; Schon et al. 2001; Yi et al. 2004).

With the exception of the HERV-W element encoding syncytin, little is known regarding the expression and

Received 16 August 2006, revised 26 March 2007, accepted for publication 8 May 2007
transcriptional control of individual HERV-W elements. This gap in knowledge is in part explained by the notion that such noncoding RNAs represent transcriptional noise and therefore lack biological relevance but also by methodological challenges associated with studies of transcripts from large numbers of closely related sequences. Although large-scale studies of transcripts from the different HERV families have been conducted by means of hybridization to arrays of representative sequences (Seifarth et al. 2005) or by use of degenerate primers and probes in real-time polymerase chain reaction (PCR) assays (Forsman et al. 2005), the individual elements are not resolved by such methods.

We recently reported on the use of sequence-dependent variations in melting temperatures (Tms) following amplification in real-time PCR for identification of transcriptionally active HERV-W elements in vitro (Nellaker et al. 2006). By this approach, we observed cell-specific patterns of expression of HERV-W gag and env sequences during baseline conditions. Following influenza A virus infection of different human cell lines, we observed both quantitative and qualitative changes in the expression pattern of such sequences. Mapping revealed transcription of both intronic and intergenic elements in the investigated cell lines. Interestingly, not only elements with intact 5'-LTRs but also those lacking such regulatory regions were expressed and regulated in response to the virus infection (Nellaker et al. 2006).

In peripheral blood mononuclear cells (PBMCs), abnormalities in the expression patterns of protein-coding genes have recently been identified by several investigators in both patients with recent-onset and chronic schizophrenia, suggesting that disease-related alterations in gene expression can be observed in these cells (Bowden et al. 2005; Tsuang et al. 2005; Vavter et al. 2004) and, to some extent, even mirror expression alterations observed in the brain (Glatt et al. 2005). The normal expression pattern of genomic regions harboring HERV-W elements in PBMCs (or in any other cell type) is not known. Neither is it known if abnormalities in the cellular transcription pattern of such regions can be identified in individuals with schizophrenia. The aim of the present study was therefore to conduct a detailed analysis of HERV-W expression in PBMCs obtained from healthy individuals and from patients first hospitalized for schizophrenia, schizophréniform or schizoaffective disorder.

Materials and methods

Patients and control individuals

PBMCs were obtained from 30 individuals (27 ± 8 years [mean ± SD], 22 males and 8 females) presenting at the Department of Psychiatry, University of Heidelberg, with symptoms consistent with schizophrenia-related psychoses including schizophrenia (n = 16), schizophreniform disorder (n = 13) or schizoaffective disorder (n = 1) as defined by the Diagnostic and Statistical Manual of Mental Disorders, 4th edn (Wittchen et al. 1997). All patients experienced their first hospitalization for a psychotic episode and none had a lifetime history of neuroleptic treatment. Subjects were excluded if they had a lifetime history of major head trauma with loss of consciousness, neurological disease, severe substance abuse or serious medical disease. Psychopathological symptoms (positive and negative syndrome scale) (Kay et al. 1987) were assessed on admission (100.2 ± 22.7, mean score ± SD), at the end of the first week of treatment (74.4 ± 21.8) and after remission of acute symptoms before discharge (62.3 ± 22.2). Neurological soft signs (NSS) (Schröder et al. 1991) were examined after remission of florid symptoms (12.8 ± 5.7). Prognostic factors were protocolled on the Strauss–Carpenter scale (Strauss & Carpenter 1974) (54.5 ± 8.1); duration of psychosis was on the Interview for the Retrospective Assessment of the Onset and Course of Schizophrenia and Other Psychoses (Hafner & Löfler 2003) (41.8 ± 49.0 months). The blood samples were obtained 5.0 ± 4.5 days (range 0–14 days) after admission. Neuroleptic treatment was initiated as clinically warranted; on the day blood samples were drawn, six patients received olanzapine or benperidol and one patient received clozapine, while 17 patients were neuroleptic free.

Blood samples were drawn using 5-ml Vacutainer® cell preparation tubes with sodium citrate (BD, Franklin Lakes, NJ, USA) and immediately centrifuged at 1500 × g for 25 min. The PBMC fraction was subsequently aspirated in 0.5–1 ml plasma and stored at −80°C until analysis. PBMCs were collected in an identical manner from 26 controls (31 ± 9 years, 11 males and 15 females) recruited by public advertisement. By screening examination, any evidence of psychiatric, neurological or acute medical illness was ruled out.

All laboratory analyses were carried out simultaneously for all samples with the experimenter being blinded to the clinical diagnoses. Informed consent was obtained from all study participants. The study was approved by the local ethics committee.

Cells and human tissue complementary DNA panel

Lymphocytes were isolated from buffy coats from healthy blood donors through Lymphoprep (Nycomed, Oslo, Norway). The following day, cells were washed and resuspended in 5 ml fresh RPMI-1640 at concentration 1 × 10⁶ cells/ml. Total RNA was isolated from the lymphocyte cultures after 24 or 96 h of incubation in the presence or absence of 10 µg/ml of concanavalin A (ConA; Pharmacia, Upsala, Sweden). Total RNA was extracted from the human histiocytic lymphoma cell line U937 (American Type Culture Collection, Manassas, VA, USA), previously reported to express a range of HERV-W elements (Johnston et al. 2001; Nellaker et al. 2006). The human multiple tissue complementary DNA (cDNA) panel was purchased from BD Biosciences (Clontech, Palo Alto, CA, USA).

RNA extraction and cDNA synthesis

Total RNA was extracted from 100–300 µl of the thawed PBMCs suspension using the RNeasy Mini Kit according to the manufacturer’s instructions (Qiagen, Hilden, Germany). The RNA was eluted in 40 µl of ribonuclease-free water and stored at −70°C until further processing. Five microliters of RNA was subsequently treated with 1 unit of deoxyribonuclease (DNase) for 30 min at 37°C to remove contaminating genomic DNA, followed by the addition of 2 µM ethylenediaminetetraacetic acid and inactivation at 65°C for 10 min according to the manufacturer’s instructions (Invitrogen, Groningen, the Netherlands). Oligo-d(T)₁₂–₁₈-primed cDNA was subsequently generated by SuperScript II reagents according to the manufacturer’s instructions (Invitrogen). Strand-specific cDNAs were synthesized using SuperScript III reagents (Invitrogen) at 55°C to increase the specificity. Control reactions in the absence of reverse transcriptase (no-RT controls) were generated in parallel for each sample.

Design of primers and probes

Sequences homologous to HERV-W gag (accession no. AF156961) and env (accession no. AF156963) were identified by a BLAST search against the human genome (16 May 2004) at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST). These sequences were downloaded and aligned using CLUSTALW (http://www.ebi.ac.uk/clustalw/). Assays detecting a number of related HERV-W gag or env sequences using SYBR I Green chemistry, previously described in Nellaker et al. (2006), were designed using PRIMER EXPRESS 2.0 (Applied Biosystems, Palo Alto, CA, USA). The HERV-W gag and env assays detect at least 44 and 11 targets in the
human genome, respectively. Specific assays for different HERV-W elements, using either SYBR Green or TaqMan chemistries, were designed using the PRIMER EXPRESS software (Tables 1 and 2). Primers used for characterization of transcript from chromosome 11q13.5 using regular PCR were also designed using PRIMER EXPRESS (Table 3).

**Real-time PCR**

Real-time PCR was performed using an ABI Prism 7000 sequence detection system (SDS) (Applied Biosystems). One microliter of cDNA template was analyzed in triplicate 25 μl reaction volumes using Platinum SYBR Green qPCR SuperMix (Invitrogen) or TaqMan universal PCR Master Mix (Applied Biosystems) according to instructions from the manufacturers. Initial PCR amplifications were carried out to determine optimal primer and probe concentrations according to the guidelines from Applied Biosystems.

Threshold cycle (Ct) values from the exponential phase of the PCR amplification plot were exported directly into Microsoft Excel work-sheets for further analysis. For each sample, the Ct value for each target transcript was normalized to that of the gene encoding β-actin by calculating ΔCt according to the formula

\[
\Delta Ct = Ct_{\text{target}} - Ct_{\text{β-actin}},
\]

as previously described (Livak & Schmittgen 2001). These ΔCt values were used for the statistical analyses. For each target transcript, the mean ΔCt value obtained in the control group was subsequently subtracted from that of the patient group, obtaining a ΔΔCt value for each target. From these values, fold differences in the levels of transcripts between the two groups were calculated according to the formula

\[
2^{-\Delta\Delta Ct}.
\]

The Tm for each amplicon was determined using the ABI Prism SDS software (Applied Biosystems) by recording the temperatures corresponding to the maximal rate of dissociation of double-stranded DNA. Analysis was performed through the classification of Tms into discrete temperature ranges that could reliably be distinguished as previously described (Nellaker et al. 2006).

**Cloning and sequencing**

PCR products generated by each primer pair or primer/probe combination, respectively, were purified using the PCR cleanup kit (Qiagen) and ligated into the pCR2.1-TOPO Vector using the TOPO TA Cloning Kit (Invitrogen). Recombinant plasmids found to contain a positive insert by PCR were subsequently sequenced at KISeq (Karolinska Institutet, Stockholm, Sweden). The resulting sequences were mapped into the human genome using a BLAT search at UCSC (Santa Cruz, CA, USA; http://genome.ucsc.edu/cgi-bin/hgBlat, May 2004 assembly). Subsequently, the flanking 5000 bases 5’ of the unambiguously mapped HERV-W gag elements were downloaded and screened against a library of repetitive elements (http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker) for the presence of a HERV-W LTR.

**Statistical analyses**

According to Kolmogorov–Smirnov tests, all variables were normally distributed. Accordingly, unpaired Student’s t-tests (two-tailed) were used to detect significant changes in gene expression. To differentiate potential diagnosis and gender effects, a two-way analysis of variance (ANOVA) was calculated. The potential relation between the levels of transcripts and other variables were analyzed by linear regression and calculating Pearson’s correlation coefficients. All computations were performed using the GRAPH PAD PRISM 3.0 or the SAS software (GraphPad, San Diego, CA, USA and the SAS Institute, Table 1: TaqMan PCR primer and probe sequences

<table>
<thead>
<tr>
<th>Target</th>
<th>Polarity</th>
<th>Sequence 5’ to 3’</th>
<th>Acc. no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3q26.32gag</td>
<td>Sense</td>
<td>TCAGGTCAAACTAGATGACAAACA</td>
<td>AF156961</td>
</tr>
<tr>
<td></td>
<td>Antisense</td>
<td>CAATGAGGTCTACACTGGGAATC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>6-FAM-CCCTGAGGAGTGGTTCG</td>
<td></td>
</tr>
<tr>
<td>ERVWE1 env</td>
<td>Sense</td>
<td>GTCAAACATTCCTCCAGAATGCA</td>
<td>AF156963</td>
</tr>
<tr>
<td></td>
<td>Antisense</td>
<td>CATGAGGTCTACACTGGGAATC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>6-FAM-AAAACATCAAATGGAGGCAGAGAGTGCAGCC-TAMRA</td>
<td></td>
</tr>
</tbody>
</table>

Acc. no., GeneBank accession numbers of sequences used for primer and probe design.

FAM, 6-carboxyfluorescein; MGB, minor groove binder; TAMRA, 6-carboxytetramethylrhodamine.

Table 2: SYBR Green PCR primer sequences

<table>
<thead>
<tr>
<th>Target</th>
<th>Polarity</th>
<th>Sequence 5’ to 3’</th>
<th>Acc. no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>HERV-W gag</td>
<td>Sense</td>
<td>TCAGGTCAAACTAGATGACAAACA</td>
<td>AF156961</td>
</tr>
<tr>
<td></td>
<td>Antisense</td>
<td>CAATGAGGTCTACACTGGGAATC</td>
<td></td>
</tr>
<tr>
<td>HERV-W env</td>
<td>Sense</td>
<td>CCAATGACCATGCTGTTTTAC</td>
<td>AF156963</td>
</tr>
<tr>
<td></td>
<td>Antisense</td>
<td>GAGTTACAGACAAAAAAATTTTC</td>
<td></td>
</tr>
<tr>
<td>5p13.3 gag</td>
<td>Sense</td>
<td>CCTGAGGCTCAGATCAAGGAG</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Antisense</td>
<td>CCCGCTTAGCCAGCATG</td>
<td></td>
</tr>
<tr>
<td>11q13.5 gag</td>
<td>Sense</td>
<td>GTTTGCGCACAATCACTG</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Antisense</td>
<td>CGATCTGGTCTCAGGGAATG</td>
<td></td>
</tr>
<tr>
<td>PTD015</td>
<td>Sense (A)</td>
<td>GGATTTGAGAGAGAAACAGGGAAT</td>
<td>NM_024684</td>
</tr>
<tr>
<td></td>
<td>Antisense (B)</td>
<td>ACCCTTCTCACAACCTTCTCCAC</td>
<td></td>
</tr>
<tr>
<td>β-actin</td>
<td>Sense</td>
<td>CGCAGGAGAGATGACCGGATG</td>
<td>NM_001101</td>
</tr>
<tr>
<td></td>
<td>Antisense</td>
<td>CAGAGGCTAGAGGGATGAC</td>
<td></td>
</tr>
</tbody>
</table>

Acc. no., GeneBank accession numbers of sequences used for primer and probe design; NA, not applicable.
Yao et al.

Table 3: 11q13.5 HERV-W-gag-specific oligonucleotides

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence 5’ to 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>425</td>
<td>CCCTCTCAAGTGGGAG</td>
</tr>
<tr>
<td>665</td>
<td>TCTCGGCTGAGTATAG</td>
</tr>
<tr>
<td>765</td>
<td>CACTGGGACACAGAAATCAGAATAC</td>
</tr>
<tr>
<td>807</td>
<td>GTTTGGGCAACAACTGT</td>
</tr>
<tr>
<td>830</td>
<td>AGGACTAAGGAAACTAGGAAAGCT</td>
</tr>
<tr>
<td>993</td>
<td>TTAAGATTAGTCGGCCTTCAGTACAG</td>
</tr>
</tbody>
</table>

For the orientation and use of the primers, see Fig. 4.

Results

No traces of contaminating genomic DNA could be detected in any of the control samples generated in the absence of reverse transcriptase by means of the HERV-W gag SYBR I Green PCR assay (data not shown). In the samples (30 cases vs. 26 controls) included in each assay, transcripts encoding β-actin were detected in all samples and the Ct values did not differ between the two groups (data not shown). The levels of transcripts encoding β-actin were thus used as an endogenous control. However, not all samples contained sufficient starting amounts of RNA to allow detection of all target transcripts (as indicated by a high Ct value obtained in the assay for transcripts encoding β-actin) in the subsequent assays and such samples were therefore excluded from the quantitative analyses. In the remaining samples, the levels of β-actin encoding transcripts remained at similar levels in the two groups. No significant differences in the incidence of any given transcript between two groups was observed.

Quantitative analyses of HERV-W-related transcripts

As determined by SYBR Green assays, the relative levels of transcripts from HERV-W gag elements were significantly elevated by 2.1-fold in PBMCs from 28 recent-onset schizophrenia patients compared with PBMCs from 25 control individuals. Although this difference was more pronounced in the female patients and the controls, the ANOVA yielded a significant main effect ‘diagnosis’ (df = 1, 51, F = 4.91, P < 0.05) but no significant main effect ‘sex’ nor interaction ‘diagnosis × sex’. However, the levels of transcripts from HERV-W env elements among the 24 patients and 21 controls with detectable levels of such transcripts were present at similar levels in the two groups (Fig. 1a). Expression levels of HERV-W gag in the patients were not correlated to the duration of initial hospitalization/exposure to medication, indicating that the differential expression of gag sequences is not an effect of medication.

Qualitative analyses of HERV-W-related transcripts

As was previously noted in vitro (Nellaker et al. 2006), amplicons with different Tms were generated in the HERV-W gag and HERV-W env SYBR Green assays. Sequencing of amplicons from multiple wells from both patients and controls, representative of each of the Tms, verified their relatedness to HERV-W. Based on the sequencing results and interassay Tm variations, the amplicons generated in the HERV-W gag and env assays were classified into discrete Tm ranges. These Tm ranges represented sets of sequences indistinguishable from each other based on their Tms. The frequency distribution of the HERV-W gag and env amplicons in these categories, i.e. the detectable pattern of transcribed HERV-W elements, did not differ between the two groups (data not shown). The relative levels of gag-related transcripts in each of the different Tm ranges were elevated in the patients although only those in the 80.2 ± 0.2°C Tm range reached statistical significance (2.8-fold elevation, P < 0.05). We identified three different sequences in this Tm range. A similar analysis of the different Tm ranges generated in the env assay indicated that all were expressed at similar levels in the two groups (Fig. 1b).

Identification of transcribed genomic elements

Of 11 different sequences detected in the HERV-W gag assay, nine could be mapped to a unique genomic position...
Similarly, four of the six different env sequences could be mapped to exact positions in the genome (Fig. 2b). In this assay, transcripts encoding syncytin could not be distinguished from two other env genes. We therefore measured its levels using a specific assay. Transcripts encoding syncytin, similarly as for other HERV-W env elements, were detected at similar levels in the PBMC samples obtained from first-hospitalized patients and control individuals (data not shown).

As shown in Fig. 2, three gag sequences on 3q26.32, 5p13.3 and 11q13.5, respectively, were identified in the Tm range (80.2 ± 0.2°C) detected at significantly elevated levels in the patients. We subsequently measured the levels of each of these transcripts by specific assays. The relative levels of transcripts from the HERV-W gag on 11q13.5 were significantly elevated by 1.6-fold \( (P < 0.05) \) in PBMCs from 28 patients compared with PBMCs from 25 normal controls, whereas the levels of transcripts from those on 3q26.32 and 5p13.3 did not differ significantly between the two groups (Fig. 3).

Based on our previous in vitro findings of transcripts from HERV-W elements lacking the regulatory regions in the LTR (Nellaker et al. 2006), we examined if such elements were expressed in these in vivo samples. According to RepeatMasker analyses of the mapped elements (Fig. 2a), a proviral element with an intact 5'-LTR (5q12.1) was identified, as well as elements lacking the regulatory U3 region of the 5'-LTR (the pseudoelements: 1p34.2, 2p16.2, 3q26.32, 4q33, 11q13.5 and 12p13.31) and elements lacking recognizable LTR sequences (5p13.3 and 11p14.3) (Costas 2002; Pavlicek et al. 2002). The provirus and four out of the six pseudoelements were located in introns of annotated genes, all opposite to the transcriptional direction of the respective gene. The two truncated elements lacking LTR sequences were located in intergenic regions. The env transcripts all mapped to intergenic elements, including that on 1p34.2, which was also detected by the gag assay.

### Characterization of the HERV-W element on 11q13

The HERV-W pseudoelement on 11q13.5 is inserted into the negative strand of the second intron of the gene PTD015 (Strausberg et al. 2002), putatively encoding a conserved protein of 122 amino acids belonging to a large family of uncharacterized proteins. Alignment of the 5.1 kb retroviral element to the consensus HERV-W sequence in RepBase (HERV17) (Jurka 2000) shows that it contains 1191 and 2349 bases of the gag and pol genes, respectively, whereas only 320 bases in this element show homology to the consensus env gene (Fig. 4). Although expressed sequence tag (EST) support for transcripts containing the gag sequence is missing, the GenBank messenger RNA sequences AK022306 and AB063619 map to the pol/env regions located some 1400 bases 3' of our detected gag sequence, supporting transcription from this region (Fig. 4). The element harbors 12 open reading frames longer than 50 bases in the viral orientation, (−) strand, and 6 in the host gene orientation, (+/−) strand. These range from 150–474 bases, none of them is, however, in a context favorable for translation (Kozak 1989).

### Characterization of HERV-W gag transcripts from 11q13

All analyses were performed on total cellular RNA. We therefore investigated if the elevated levels of transcripts from the HERV-W gag on 11q13.5 stemmed from unprocessed primary transcripts from the PTD015 gene and thus served as a marker of an elevated transcriptional activity from the

---

**Figure 2:** Mapping of HERV-W elements transcribed in PBMCs. Sequences and genomic positions of mapped HERV-W gag (a) and env (b) elements. Arrows indicate the primer annealing sites. A detected sequence was mapped to a genomic element only if this displayed identity also in the underlined 3' regions of the primers. Tm indicates melting temperature ranges of the amplicons. Dashes indicate nucleotides identical to those in the prototypical sequences (indicated at the top). Open circles indicate gaps in the sequence introduced to optimize the alignment.
this gene. This did however not appear to be the case because the levels of PTD015 mRNAs as determined by primers A/B (Fig. 4, Table 2) spanning the second intron indicated similar levels in patients and control individuals. In addition, no correlation between spliced PTD015 transcripts and gag transcripts was observed in the clinical samples (data not shown). Taken together, these findings indicate independent regulated expression of the retroviral element. This was further investigated in a commercially available panel of cDNAs generated from adult human tissues. Relative levels of transcripts encoding PTD015 and the intronic HERV-W cDNAs generated from adult human tissues. Relative levels of transcripts were calculated as described in the Material and methods. n, number of samples with detectable levels of targets in each group. *P < 0.05, unpaired t-test. Results are expressed as mean ± SEM.

Using DNase-I-treated RNA isolated from U937 cells and freshly isolated human lymphocytes, we subsequently determined the polarity of the gag transcripts from this element. Total RNA from these cells was reverse primed using either oligo 765 or 807 (Fig. 4, Table 3). Subsequent PCRs using the primer pairs 993/830 or 665/425 (for detection of sense or antisense transcripts, respectively) resulted in amplification of products primed by both oligo 765 and 807 (Fig. 4). Sequencing of these two products verified that they were both generated from the HERV-W element on 11q13.5. Thus, in addition to intronic sense transcripts, potentially encompassing incompletely processed RNAs, the gag sequence of the HERV-W element on 11q13.5 appears to be part of intronic antisense transcripts in U937 cells as well as in primary lymphocytes. This finding implies promoter activity upstream of the integration site on the genomic (−) strand (Fig. 4).

We previously reported that serum deprivation of U937 cells induced both quantitative and qualitative changes in transcripts containing HERV-W sequences (Nellaker et al. 2006). Potential transcriptional regulation of the HERV-W element on 11q13.5 was therefore further investigated in U937 cells cultured in the presence or absence of 10% fetal calf serum for 24 h. In serum-deprived cells, the levels of polyadenylated 11q13.5 HERV-W gag transcripts were significantly elevated by 3.1-fold (P < 0.05) compared with cells grown in complete medium (Fig. 5b). As determined by strand-specific reverse transcription followed by real-time PCR, the levels of 11q13.5 HERV-W gag sense and antisense transcripts were elevated by 5.1-fold (P < 0.01) and 3.1-fold (P < 0.01), respectively, in starved cells (Fig. 5b). Taken together, these findings indicate regulated expression of both sense and antisense strands from this genomic region.

**State of activation of PBMCs in patients and controls**

Activation of cells of the immune system has been reported to influence the relative levels of transcripts from HERV elements, including HERV-W (Johnston et al. 2001). A number of reports also suggest subtle immune system aberrancies in patients with schizophrenia compared with control individuals, reviewed in Strous and Shoenfeld (2006). We therefore investigated if the state of activation of the PBMCs differed between patients and control individuals by comparing the levels of transcripts from two genes: malate dehydrogenase 1 (MDH1) and hexokinase 3 (HK3). These genes encode enzymes involved in the tricarboxylic acid cycle and glycolytic pathways, respectively, known to be affected by activation of both myeloid and lymphoid cells (Aidoo et al. 1996; Keast & Newsholme 1990; Suzuki et al. 2000). In addition, transcripts from MDH1 were previously reported to be differentially expressed in B-cell blasts derived from patients with schizophrenia (Vawter et al. 2004). Both these genes appeared to be transcribed at comparable levels in PBMCs from patients and from control individuals (data not shown) suggesting comparable levels of metabolic activities in cells present in our preparations from patients and control individuals. To further investigate if our findings could potentially reflect lymphocyte activation in the patients not detected by our assays, we investigated if the levels of HERV-W gag transcripts were at all affected by mitogen stimulation in lymphocyte cultures from three blood donors. As shown in Fig. 5c, con A stimulation for up to 96 h, however, did not affect the overall levels of HERV-W gag transcripts or those from the gag gene on 11q13.5. Successful stimulation of these cultures was indicated by the induced transcription from the MDH1 gene to meet the increased metabolic demands of proliferation, whereas the levels of transcripts from the HK3 gene were downregulated in the stimulated cells (Fig. 5c). Taken together, these findings do not support differential activation of PBMCs in the patients as the underlying cause of the elevated levels of transcripts containing HERV-W gag sequences.
Correlations with clinical measures

Finally, we investigated if the levels of HERV-W transcripts were correlated to any clinical measure in the patients. The ΔCt values obtained (high numbers indicating a low level of transcripts and vice versa) from the analyses of overall HERV-W gag elements were significantly correlated to the duration of untreated illness (r = 0.57, P < 0.005) in the patients (Fig. 6). Thus, there appears to be an inverse correlation between the relative levels of HERV-W gag transcripts and the duration of untreated illness. Other clinical measures, such as severity of psychopathological symptoms on admission, 1 week after admission and after remission of acute psychosis or NSS scores, showed no significant correlations with the levels of HERV-W gag transcripts. No correlations between the relative expression levels of gag element on 11q13.5 with any of these clinical measures was detected in the patients (data not shown).

Discussion

From the present study, it is evident that human PBMCs in vivo transcribe a number of different genomic HERV-W elements. These transcripts can be mapped to elements containing an intact regulatory LTR region as well as to those lacking this region (or the entire LTR). The transcribed HERV-W elements detected were located to both intronic and intergenic regions. Because the current analysis was conducted on total cellular RNA, we cannot exclude that some of the detected transcripts from intronic elements may therefore represent partially or unprocessed primary transcripts. Our present findings are thus in agreement with our previous findings in human cell lines (Nellaker et al. 2006). Taken together, our data support transcription of not only proviral elements but also elements lacking 5’ regulatory U3 regions and suggest that the potential for regulation of transcription from regions harboring HERV-W elements is more complex than can be appreciated solely from studies of HERV-W LTR-related promoter activities.

Finding polyadenylated transcripts containing intronic or intergenic HERV-W elements with limited protein-coding capacities raises questions regarding the nature of these transcripts and their potential biological significance. It has, during recent years, become evident that a large proportion of transcription occurs outside of the boundaries of known genes, reviewed in Johnson et al. (2005). In a recent study of the evolution of intergenic transcription, Khaitovich et al. (2006) recently reported that such transcription appears to have been conserved to the same extent as that of traditional protein-coding genes, which suggests that intergenic transcripts have functional significance. If this applies also to regions containing repetitive HERV-W elements remains to be investigated.

In the process of looking for disease-associated changes in HERV-W transcripts in the PBMCs obtained from patients first hospitalized for schizophrenia-related psychoses, we did not detect any absolute qualitative differences (i.e. absence
or presence) in HERV-W gag or env transcripts compared with cells from control individuals. Quantitatively, however, our analyses indicated increased levels of transcripts from a subset of HERV-W elements in the patient group samples. Element-specific analyses detected elevated levels of transcripts containing the gag sequence of an intronic HERV-W element on 11q13.5 in the patients compared with the control individuals.

Variations detected in the relative amounts of transcripts from the element on 11q13.5 in various normal tissues indicate cell-type-specific expression and regulation of levels of transcripts containing this element. Furthermore, changes in transcription of both sense and antisense transcripts in response to serum deprivation in vitro suggest that this particular transcript is not merely representing unprocessed primary transcripts from the surrounding putative gene PTD015.

The levels HERV-W-gag-related transcripts were inversely correlated to the duration of untreated psychosis but not to other clinical measures including disease severity in these patients. This is in agreement with our previous findings where detectable retroviral RNA was more prevalent in CSF or plasma from recent-onset patients than from chronic patients (Karlsson et al. 2001, 2004). Taken together, these findings suggest that the levels of such transcripts are related to biological processes involved in the transition from health to disease in these individuals.

Age, sex, geographic or socioeconomic differences and neuroleptic therapy have to be considered as potential confounding variables. Mean age did not significantly differ between groups. While the patients’ group comprised a higher proportion of males, the levels of transcripts showed only marginal, nonsignificant sex differences. All patients came from the same geographical area in the Palatine and received a uniform therapy with atypical neuroleptics.

With the exception of the HERV-W element on chromosome 7q21 containing the intact env gene encoding syncytin,
few studies have been undertaken to experimentally address the transcriptional regulation of transcripts containing HERV-W elements. The finding that the levels of transcripts containing HERV-W elements (including that on 11q13.5) were affected by serum deprivation suggests that environmental stressors can influence the transcriptional activity of such regions in vivo. This notion is supported by our previous observations regarding transcription of HERV-W elements, including that on 11q13.5, in human cell lines infected with influenza A virus (Nellaker et al. 2006). Furthermore, herpes simplex virus type 1 has previously been reported to transactivate HERV-W sequences, although the individual responsive elements were not identified (Lee et al. 2003). Whether such or other environmental stressors are responsible for the present findings, however, remains to be determined.

A number of recent array studies have identified gene expression changes in peripheral blood cells obtained from individuals suffering from schizophrenia, which suggests that this cell population is, to some extent, useful for the identification of markers of the disease (Bowden et al. 2006; Glatt et al. 2005; Tsuang et al. 2005; Vawter et al. 2004). Compared with these, the present study is limited by the fact that we only have data on expression of a restricted number of genomic regions. Thus, differences in expression from many coding and noncoding regions in the present sample remain undetected. The lack of data from a clinical follow-up evaluation of these patients prevents us from drawing any conclusions regarding the prognostic value of our observations.

In conclusion, we have identified transcripts containing HERV-W elements in PBMCs obtained from healthy controls and from patients with recent-onset schizophrenia or schizoaffective disorder. Some of these elements, including an intrinsic element on the noncoding strand of PTD015 on 11q13.5, appear to be transcribed at a higher rate in the patients during the transition from susceptibility to manifestations of symptoms. These observations leave a number of unresolved questions that need to be addressed in future studies, particularly regarding the biological processes underlying the transcriptional activation, the characteristics of the full-length transcripts and the functional significance of such transcripts.

It is the number of noncoding transcripts as opposed to the number of genes that appear to most closely correlate to the complexity of an organism (Taft et al. 2007). The present study supports the idea that it may be worthwhile to extend gene expression profiling of complex human diseases to include noncoding transcripts, as previously proposed by Perkins et al. (2005).

References


Genes, Brain and Behavior (2008) 7: 103–112

HERV-W and schizophrenia


Acknowledgments

This study was supported by the Stanley Medical Research Institute, Stiftelsen Bror Gadelius Minnesfond and the Swedish Research Council (21X-20047).