Expression of human endogenous retroviral sequences in peripheral blood mononuclear cells of healthy individuals

Patrik Medstrand,* Mats Lindeskog and Jonas Blomberg

Section of Virology, Department of Medical Microbiology, Sölvegatan 23, S-223 62 Lund, Sweden

The polymerase chain reaction was used to detect expression of retroviral sequences with oligonucleotide primers derived from conserved regions of the retroviral genome. Four primer pairs derived from gag and one from pol were used in amplification of reverse-transcribed total RNA prepared from peripheral blood mononuclear cells of seven blood donors. The amplification pattern was the same from each of the seven samples. Sequencing of cloned amplification products revealed that at least three subclasses of sequences related to the human endogenous retroviruses (HERV) RTVL-H, HERV-E and HERV-K, are expressed in peripheral blood mononuclear cells of healthy individuals. This has not been previously reported.

Human endogenous sequences with similarity to infectious murine, primate and human retroviruses are present in thousands of copies, 0.1 to 0.6%, per genome (Callahan, 1988; Brack-Werner et al., 1989). Recent reports of involvement of presumably endogenous retroviruses in human autoimmune (Garry et al., 1990) and neoplastic (Cianciolo et al., 1984) diseases suggest that their expression may be of considerable consequence. None of the endogenous sequences detected so far have been shown to contain open reading frames (ORFs) for all genes (i.e. gag, pol and env). However, endogenous retrovirus 3 (ERV3) has an ORF which could code for a polypeptide of 650 amino acids of env (Cohen et al., 1985), and human ERV (HERV)-K10 contains an ORF large enough to produce a full-length pol protein (Ono et al., 1986). By using low stringency hybridization, several discrete mRNA retrovirus-like species have been found to be expressed in human neoplastic tissue cells, placenta, transformed cells and cell lines (for a review see Larsson et al., 1989; Leib-Mösch et al., 1990). In this study we used the polymerase chain reaction (PCR) to amplify total RNA after reverse transcription (RT/PCR), with primers derived from evolutionarily conserved regions of gag and pol (Table 1). PCR has been used for detection of both previously unknown and known retroviral sequences (Shih et al., 1989; Bohannon et al., 1991). Here we report that three subfamilies of retroviral endogenous gag- and pol-related sequences are expressed in human peripheral blood mononuclear cells of healthy blood donors.

Lymphocytes were removed from the interface and washed twice (1000 r.p.m., 10 min) with PBS. Lymphocytes were then either stored at −70 °C or directly lysed in a 4 M-guanidinium thiocyanate solution (Merck) and total RNA was prepared according to a CsCl ultracentrifugation protocol (Maniatis et al., 1982). The RNA was treated with RNase-free DNase (Promega), the concentration was determined by u.v. spectrophotometry at 260 nm and it was stored in aliquots of 0.5 μg/μl at −70 °C.

Lymphocyte RNA was analysed by RT/PCR using oligonucleotide primers derived from gag (primer pairs 1, 2, 3 and 4; Table 1) and pol (primer pair 5; Table 1). The primer sequences presented in Table 1 were specified by one of us (M. L.) and synthesized (Scandinavian Gene Synthesis AB) from evolutionarily conserved regions using alignments of gag (P. Medstrand, M. Lindeskog & J. Blomberg, unpublished results) and pol (Doolittle et al., 1989). RT was performed in a volume of 10 μl containing 50 mM-KCl, 20 mM-Tris–HCl pH 8.4, 2.5 mM-MgCl₂, 0.1 mg/ml BSA (Boehringer), 0.5 mM of each dNTP (Boehringer), 100 pmol random hexaoligonucleotides (Boehringer), 20 units RNasin (Promega), 100 units reverse transcriptase (Gibco-BRL) and 1 μg of RNA. After incubation for 1 h at 42 °C the reverse-transcribed samples were incubated at 95 °C for 5 min, transferred to ice and the volume was adjusted to 50 μl with a PCR mix, containing (final concentrations) 50 mM-KCl, 20 mM Tris–HCl pH 8.4, 2.5 mM-MgCl₂, 0.1 mg/ml BSA, 50 ng PCR primers and 1 unit of thermostable DNA polymerase (Perkin-Elmer Cetus). Samples were overlaid with 50 μl of liquid paraffin and amplified using an Intelligent Heating Block (Cambio) for 30 cycles with the following profile: 95 °C for 30 s,
Table 1. Primers used for amplification

<table>
<thead>
<tr>
<th>Pair</th>
<th>Name†</th>
<th>Sequence‡</th>
<th>Length</th>
<th>Target (T_d)§</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RHGAL</td>
<td>CTTTATTACCAACTCAGTGCCTCACCAGYAT</td>
<td>380</td>
<td>RTVL-H (72/62)</td>
</tr>
<tr>
<td></td>
<td>HTGAR</td>
<td>TTTAGTTGGGTGACATCTCAGTTTCCARTG</td>
<td>309-</td>
<td>MoMuLV (60/72), FLV (54/70), SSAV (52/72), BaEV (50/72)</td>
</tr>
<tr>
<td>2</td>
<td>RHGAL</td>
<td>CTTTATTACCAACTCAGTGCCTACCCAGYAT</td>
<td>354</td>
<td>HERV-E (42/64), S-71 (42/54), RaLV (50/60)</td>
</tr>
<tr>
<td></td>
<td>MLGAR</td>
<td>CAATGACCTTTTTTACATAGGCAGCRA</td>
<td>712</td>
<td>MoMuLV (74/72), FLV (72/76), SSAV (74/72), BaEV (76/72)</td>
</tr>
<tr>
<td>3</td>
<td>MLGAR</td>
<td>CAATGACCTTTTTTACATAGGCAGCRA</td>
<td>751</td>
<td>HERV-E (58/64), S-71 (76/54), RaLV (70/60)</td>
</tr>
<tr>
<td></td>
<td>MLGAR</td>
<td>CAATGACCTTTTTTACATAGGCAGCRA</td>
<td>709</td>
<td>HERV-E (82/64)</td>
</tr>
<tr>
<td>4</td>
<td>41GAL</td>
<td>AACCCCACTTGGGCTAGTGGCACCACCA</td>
<td>294</td>
<td>MPMV (58/72), SRV1 (64/72), SRV2 (64/66), MMTV (66/66)</td>
</tr>
<tr>
<td></td>
<td>MLGAR</td>
<td>CAATGACCTTTTTTACATAGGCAGCRA</td>
<td></td>
<td>HERV-K (66/70), IAPh (64/64), IAPm (64/68), RSV (52/64)</td>
</tr>
<tr>
<td>5</td>
<td>ABDPOL</td>
<td>TCCCTCTGGATCTCTCTGTTTTTGYT</td>
<td>198</td>
<td>Human (72/72), chicken (70/72)</td>
</tr>
<tr>
<td></td>
<td>ABDPOR</td>
<td>TGGCTGGATCTCTCTGTTTTTGYT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>H33L</td>
<td>CCCCCTGAGGAGGGGTGAAAGAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H33R</td>
<td>CTTGCGCTCTGCAAACGACCAGAT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Primer pairs 1 and 2 were derived from the capsid protein (CA), primer pairs 3 and 4 from the CA and the nucleocapsid protein of gag, and primer pair 5 from the reverse transcriptase region of pol.
† The suffixes L or R indicate left or right primers.
‡ Primer sequences are oriented in sense (5' to 3') direction.
§ T_d, Calculated dissociation temperature for hybrids between primers and respective sequences; T_d(°C) left primer/T_d(°C) right primer.

The following abbreviations of retroviruses are used: RTVL-H, human endogenous retrovirus RTVL-H; MoMuLV, Moloney murine leukemia virus; FLV, feline leukemia virus; SSAV, simian sarcoma-associated virus; BaEV, baboon endogenous retrovirus; HERV-E, HERV clone 4-1; S-71, HERV-S-71; RaLV, rat leukemia virus; MPMV, Mason-Pfizer monkey virus; SRV1 and SRV2, simian AIDS retrovirus, serotype 1 and 2; MMTV, mouse mammary tumor virus; HERV-K, human endogenous retrovirus HERV-K; IAP, intracisternal type A particle, hamster and mouse; RSV, Rous sarcoma virus (Coffin, 1990). For the nomenclature of HERVs, see Larsson et al. (1989).

50 °C for 30 s and 72 °C for 1 min. The final extension step was 5 min. Fractions of 15 µl were loaded on a Tris-borate–EDTA 2% agarose gel. For each sample histone primers were used in parallel as a positive primer control of the RNA preparation, RT and cDNA amplification. A control without reverse transcriptase was added prior to reverse transcription was made to detect possible traces of genomic DNA in the DNase-treated RNA preparation. Precautions were taken to avoid any kind of contamination with extraneous nucleic acid. In both the RNA preparation, RT and amplification steps parallel negative water controls were used. These steps were performed in separate locations in the laboratory. In no case was there evidence of contamination.

Seven lymphocyte RNA preparations from seven blood donors were analysed. The same amplification pattern with minor intensity differences, shown in Fig. 1, was obtained with all seven. Primer pairs 1, 3, 4 and 5 and the histone control primer pair gave bands of expected sizes (Table 1). A double band indicating amplification products of different lengths was obtained with primer pair 1. No amplification was detected with primer pair 2. Besides the expected amplification product of around 700 bp of primer pair 4, a fragment of 170 bp was obtained. The control without reverse transcriptase was always negative, proving that only RNA was present prior to RT.

To determine the nucleotide sequence, 1 µl (about 50 ng) of the amplification products of primer pairs 1, 3, 4 and 5 was used for ligation into the plasmid pCR using the TA Cloning kit (Invitrogen) according to the manufacturer’s instructions. Bacteria were screened for inserts by PCR. In this case M13 forward and reverse primers were used. Plasmid DNAs with inserts of the sizes expected were isolated from bacteria using the alkaline lysis method (Maniatis et al., 1982). The sequencing reactions were performed by the dideoxy-nucleotide chain termination method (Sanger et al., 1977) as described in the Sequenase Version 2.0 kit (United States Biochemical). M13 forward and reverse primers were used. All sequence comparisons were made using the PC-GENE program package (IntelliGenetics).
The sequencing results (Fig. 2) indicated the following. (i) Primer pair 1 amplified sequences related to RTVL-H2, which is a member of a multicopy family related to human T cell lymphotropic viruses and type C retroviruses of approximately 1000 copies per cell (Mager & Freeman, 1987). Clones of two different lengths in a ratio of about 1:1 were obtained. Both the long, represented by clone 1:1, and short, represented by clone 1:2, sequence were 90% similar to RTVL-H2. In the shorter, a deletion of 17 nucleotides, compared to the genomic homologue, was present (Fig. 2a). (ii) Primer pairs 3 and 4 detected sequences similar to the clone 4-1 of HERV-E to which they were 80% (clone 3:3) and 76% (clone 4:2) similar, respectively. HERV-E is a multicopy sequence with similarity to type C retroviruses (Repaske et al., 1983). Primer pairs 1, 3 and 4 detected sequences in gag (Table 1). The 170 bp amplification product of primer pair 4 was not sequenced. (iii) Primer pair 5 amplified pol sequences. One (clone 5:30) was almost identical (98% similar), and a second (clone 5:28) only 76% similar to HERV-K10, which is a member of a group of type B-related sequences (Ono et al., 1986) (see also Larsson et al., 1989; Leib-Mösch et al., 1990).

The sequences presented here were all of RNA origin. All were similar to published genomic sequences. Considering an error frequency of Taq DNA polymerase

Fig. 2. Sequence alignment of the nucleotide sequences obtained from cloned cDNA PCR products of primer pairs 1, 3, 4 and 5 (see Fig. 1) and previously reported similar genomic sequences. (a) The derived nucleotide sequences of a long and a short amplification product of primer pair 1 aligned with RTVL-H2. (b) Alignment of nucleotide sequences detected by primer pairs 3 and 4, and HERV-E. (c) Alignment between the nucleotide sequence obtained with primer pair 5 and HERV-K10. The complete nucleotide sequences between the primers of primer pairs 1 and 5 were sequenced in both directions. Only 207 bp of the 700 bp amplificate of primer pair 3 was aligned with RTVL-H2. (b) Alignment of nucleotide sequences detected by primer pairs 3 and 4, and HERV-E.

The sequencing results (Fig. 2) indicated the following. (i) Primer pair 1 amplified sequences related to RTVL-H2, which is a member of a multicopy family related to human T cell lymphotropic viruses and type C retroviruses of approximately 1000 copies per cell (Mager & Freeman, 1987). Clones of two different lengths in a ratio of about 1:1 were obtained. Both the long, represented by clone 1:1, and short, represented by clone 1:2, sequence were 90% similar to RTVL-H2. In the shorter, a deletion of 17 nucleotides, compared to the genomic homologue, was present (Fig. 2a). (ii) Primer pairs 3 and 4 detected sequences similar to the clone 4-1 of HERV-E to which they were 80% (clone 3:3) and 76% (clone 4:2) similar, respectively. HERV-E is a multicopy sequence with similarity to type C retroviruses (Repaske et al., 1983). Primer pairs 1, 3 and 4 detected sequences in gag (Table 1). The 170 bp amplification product of primer pair 4 was not sequenced. (iii) Primer pair 5 amplified pol sequences. One (clone 5:30) was almost identical (98% similar), and a second (clone 5:28) only 76% similar to HERV-K10, which is a member of a group of type B-related sequences (Ono et al., 1986) (see also Larsson et al., 1989; Leib-Mösch et al., 1990).

The sequences presented here were all of RNA origin. All were similar to published genomic sequences. Considering an error frequency of Taq DNA polymerase
of 0.25% after 30 PCR cycles (Saiki et al., 1988; Gelfand & White, 1990), the sequence variability of the lymphocyte clones compared to the genomic ones, presented in Fig. 2, cannot merely be explained as misincorporation artefacts. The wide span (22% dissimilarity) between the two HERV-K10-like clones indicates that several rather different type B-related reverse transcriptase-coding sequences are expressed in normal cells. These two sequences have ORFs corresponding to that of HERV-K10. The translated sequences of clones 1, 3 and 4 were interrupted by stop codons. With allowance for minor nucleic acid replacements during PCR, the pattern of frameshifts and stop codons may have functional importance.

The primer pairs used in this study were designed to detect many different retroviral sequences. However, further studies are required to determine whether classes of sequences other than those presented here are expressed. Previous reports of retroviral sequences expressed in human cells were mainly based on nucleic acid hybridization and there is a lack of precise data regarding the extent of expression and degree of variation of individual retroviral sequences. Their physiological role and involvement during tumorigenesis is unknown (Leib-Mösch et al., 1990). Primer pairs presented here can be used for comparison of expression at a detailed level in normal and tumour cells of the three retroviral classes, which will increase our understanding of the role of these transcripts.

We thank Alistair Kidd for excellent technical advice and support, Elzbieta Vincic for sample collection and preparation, and Anna Blixt for help in the initial stages. This work was supported by the Medical Faculty, University of Lund, Sweden, grant no. 236, Replico AB, Lomma, and the Österlund foundation.

References


(Received 3 February 1992; Accepted 11 April 1992)