Only those genes of the KIAA1245 gene subfamily that contain HERV(K) LTRs in their introns are transcriptionally active

A.E. Illarionova, T.V. Vinogradova *, E.D. Sverdlov

Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 16/10 Miklukho-Maklaya, Moscow 117997, Russia

Received 15 February 2006; returned to author for revision 13 April 2006; accepted 21 June 2006
Available online 25 September 2006

Abstract

Insertion of LTRs into some genome locations might seriously affect regulation of the neighboring genes expression. This hypothesis is widely accepted but, however, not confirmed directly. Earlier, we have identified a family of closely related genes highly similar to the KIAA1245 mRNA counterpart. This family included a subfamily of genes some of which contained and the others lacked an LTR in their structure. We compared transcription of several closely related genes of the subfamily differing in the presence or absence of LTRs. Only LTR-containing genes were transcribed in transformed cell lines, tumorous and embryonic human tissues, whereas LTR-lacking genes remained silent. Since the genes were in the same intracellular microenvironment, we suggested that this effect was most probably due to intrinsic cis-characteristics of integrated LTRs and confirmed this by demonstrating high enhancer activity of KIAA1245 LTRs. The expression of the LTR-containing genes in embryonic tissues might suggest their involvement in evolutionary events during primate speciation.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Human endogenous retrovirus; LTR; Gene families; Regulation of gene expression; Human genome; Evolution

Introduction

Retroelements (REs) that replicate and transpose through RNA intermediates are the only known transpositionally active group of transposable elements in mammals. In vertebrates, REs occupy up to 30–40% of genetic information (Weiner et al., 1986). REs can affect genome properties in many ways including numerous host DNA rearrangements due to recombination events (Kazazian and Moran, 1998; Lei and Vorechovsky, 2005), transductions of 5′ or 3′ RE flanking sequences into new genomic loci, by creating pseudogenes (Boeke and Stoye, 1997; Brosius, 1999; Elrouby and Bureau, 2001; Goodier et al., 2000) or by causing RNA recombinations (Lei and Vorechovsky, 2005). Being mobile carriers of transcription regulatory elements, REs can affect expression of many host genes by providing transcriptional regulatory signals. Furthermore, recently expanded gene classes, such as those involved in immunity or response to external stimuli, have transcripts enriched in REs, whereas REs are excluded from mRNAs of highly conserved genes with basic functions in development or metabolism. These data, although indirectly, support the view that REs played a significant role in the diversification and evolution of mammalian genes (van de Lagemaat et al., 2003).

The potential ability to affect gene regulation, in particular those genes involved in embryo development, makes REs likely candidates for taking part in speciation processes (Hedges and Batzer, 2005; Kazazian, 2004; Sverdlov, 1998, 2000). In particular, REs might well be at least partly responsible for phenotypic differences between Homo sapiens and its closest relatives, Pan paniscus and Pan troglodytes chimpanzee species (Kazazian, 2004; Mayer and Meese, 2005; Sverdlov, 2000). These differences can be suggested to arise not only from changes in gene content but also from variations in regulation of some genes common to the related species (Han and Boeke, 2005; King and Wilson, 1975; Sverdlov, 2000). Perhaps, the most likely candidates for such a role are endogenous retroviral
long terminal repeats (LTRs) that flank retroviral cores in genomic DNA (Bannert and Kurth, 2004; Leib-Mosch and Seifarth, 1995). Their structure harbors functional enhancers, promoters, and polyadenylation signals required for retroviral gene expression (Gifford and Tristem, 2003; Urnovitz and Murphy, 1996). It was recently suggested that LTRs may drive transcription of unique host non-repetitive sequences because 5′-termini of various mRNAs were found to harbor LTR-related sequences (van de Lagemaat et al., 2003).

HERV LTRs were reported to be involved in transcription regulation of cellular downstream genes (Abrink et al., 1998; Di Cristofano et al., 1995; Kowalski and Mager, 1998; Larsson and Andersson, 1998; Samuelson et al., 1996; Ting et al., 1992). There were also reports on HERV germ line insertions that changed tissue specificity of expression of human genes (Pi et al., 2004; Schulte et al., 1996), although this kind of data should be considered with caution. As an example, it was reported that an integration of an endogenous retrovirus into the 5′-flanking region of the human amylase gene led to its expression in the human salivary gland, apart from the pancreas where this gene is normally expressed (Ting et al., 1992). However, a more detailed analysis (Samuelson et al., 1996) revealed that the retroelement was not required for amylase transcription in the primate salivary gland. Of course, this example in no way excludes the possibility of strong direct effect of HERVs on nearby gene expression.

Despite many efforts, the potential of LTRs to change the mode of human gene expression remains mostly hypothetical. It is highly difficult to prove that a gene acquired a new regulatory feature due to an RE integration rather than to accompanying events that escaped the investigator’s attention. One of the ways that could help to resolve the problem might be comparative analysis of expression modes of members of one and the same gene family that were recently duplicated in evolution. Being structurally similar, such genes can be different in retroelement integrations. In this case, genes in the same cells with the same set of transcription factors and other possible modulators of gene activity could differ in their functional behavior depending on the presence of REs.

In this report, we describe a study of such a subfamily, part of the KIAA1245 family described by us earlier (Vinogradova et al., 2002). The members of this subfamily are highly similar but some of them carry HERV(K) LTRs in their second intron, whereas the others do not. We demonstrated that only those genes that harbor LTR integrations were transcriptionally active. We also report high enhancer activity of an intronic LTR in a transient expression system.

Results and discussion

Evolutionary tree of the KIAA1245 subfamily members

When studying transcription of HERV(K) LTRs in normal and tumorous tissues, we discovered a transcript that contained an LTR and was highly similar to a fragment of the KIAA1245 mRNA (Nagase et al., 1999). Earlier, an analysis of genomic sequences deposited in GenBank revealed 10 sequences very similar to this mRNA. The sequences were shown to differ in both the degree of identity between the exons and the presence (or absence) of an HERV(K) LTR in the second intron (Vinogradova et al., 2002).

Further analysis revealed quite a number of other sequences in GenBank also similar to different fragments of the KIAA1245 mRNA. For our purposes, we identified two groups of the sequences closely similar but clearly distinct in the region of interest, characterized by the presence or absence of an LTR in the intron between exons 2 and 3. The structure of the region of interest of a typical member of the family is schematized in Fig. 1A.

We demonstrated that this LTR insertion had occurred after the divergence of orangutan from other hominoids but before the divergence of the gorilla lineage, i.e. between 8 and 13 Mya (Vinogradova et al., 2002). An evolutionary analysis of the groups above (see Fig. 1) revealed the existence of at least three extant LTR-containing and two LTR-lacking sequences closely related to the KIAA1245 mRNA (Fig. 1B). The age of the LTR-lacking sequences (AL356957 and AL592309) exceeds 48.3 and 48.8 Mya, respectively. The subgroup of LTR-containing sequences emerged between 8.6 and 48.3 Mya. The youngest AL954711 sequence diverged from the other subgroup sequences not more than 5.8 million years ago.

Differential expression of the LTR-containing and LTR-lacking genes

Below, LTR-containing and LTR-lacking genes should be understood as the members of the KIAA1245 subfamily throughout the text.

We were interested in finding differences among LTR-containing and LTR-lacking genes in their transcriptional activity. To this end, we designed RT-PCR primers of three types (see Fig. 2A and Table 2):

- Pr1 and Pr4 primers to amplify mRNAs of all genes of the family, LTR+(+).
- Pr1 and Pr2 to amplify only LTR-containing members, LTR+(−).
- Pr1 and Pr3 to selectively amplify only LTR-lacking genes of the family, LTR(−).

The differential primers were designed based on differences in the sequences of the 2nd and 3rd exons that, although conserved, still contain enough differences to allow the design of the required primers.

The primers were targeted at the exon termini bordering the intron as shown in Fig. 2A. Such a choice allowed us to test the specificity of the discriminating primers using total genomic DNA (Fig. 2B). One can see a clear difference in the length of amplicons produced with discriminating primers Pr1–Pr2 and Pr1–Pr3 aimed at distinguishing LTR-containing from LTR-lacking genes, whereas the non-discriminating Pr1–Pr4 primer pair provides the amplification of both types of genes.

Sequencing of the PCR amplification products confirmed that they corresponded to the genes of the KIAA1245
A similar analysis was performed with the amplification products obtained with the same primers but with the genomes of chimpanzee, gorilla and gibbon. Also in this case the sequences were closely similar to those of the KIAA1245 subfamily genes in the human genome (data not shown).

No visible PCR amplification products were detected with the primers above and mRNAs isolated from different adult tissues including forebrain, cerebellum, testicles, lymphocytes, liver and small intestine (Fig. 3).

In contrast, in all human embryonic tissues tested, such as basal ganglion, frontal lobe, occipital cortex, hypothalamus, thalamus, and hippocampus, the LTR-containing genes produced detectable levels of mRNA (Fig. 2C, Fig. 3). No PCR products of the LTR-lacking genes were detected (Table 1 and Fig. 2C, Fig. 3). The amplification products for the LTR-containing genes in all embryonic tissues studied were visible mostly after 33 PCR cycles. Considering that this number for actin is 27 cycles, it means that the corresponding mRNAs were present in more than 10 copies per cell. It is essential that the same number of PCR cycles was required to detect products also with the non-discriminating primers, while the LTR(–) specific primers failed to produce any visible product amounts even by the 39th cycle (Table 1). It suggests that only the LTR-containing genes produce amplifiable quantities of mRNA, and the non-discriminating primers amplify only this mRNA. This conclusion is supported by sequencing the amplicons obtained: all of 12 randomly chosen independent clones corresponded to the LTR-containing mRNAs.

We tested also several transformed human cell lines (Jurkat, Tera1, NT2/D1, RMS and NGP) and a testicular tumor, seminoma, for expression of various members of the KIAA1245 subfamily (see Table 1, Fig. 3). The result was similar to that of embryonic tissues. In the tumors and cell cultures tested, only the products from mRNAs of the LTR-containing genes were observed. A PCR analysis with the LTR(–) specific primers revealed no amplification products.

We were also interested in particular contributions of individual mRNAs produced by different LTR-containing genes in the amplification product. To this end, we used primers PrA/PrC and PrB/PrD to discriminate among three different mRNAs designated here as AL954711, AL592284, and AL049742 (Figs. 1B, 4A, B and Table 2). The discriminative capacity of the primers was proven using human genomic DNA. The PrA/PrC primer pair was designed to discriminate AL049742 sequence from AL954711, AL592284, whereas the pair of PrB/PrD allowed discriminating AL954711 and AL592284 sequences from AL049742. In spite of some diversity in sequence between AL954711 and AL592284, these two members of the family could be distinguished only by sequencing the PCR amplification product of the PrB/PrD pair. The discriminating primers were used to determine the

Fig. 1. (A) Scheme of the partial exon–intron structure common for the members of the KIAA1245 gene subfamily closely similar in the exon structure but distinct in the presence/absence of an LTR in the intron between exons 2 and 3. Exons are marked by black rectangles; LTR, a long terminal repeat of the HERV(K) family comprising the U3, US, and R regions. (B) A nearest neighbor dendrogram of genomic sequences homologous to the KIAA1245 mRNA. The dendrogram is plotted based on the comparison of intron sequences between the first and second, second and third, third and fourth, and the fourth and fifth exons. GenBank accession numbers are indicated. LTR(+)—the LTR-containing group of sequences, LTR(–)—the LTR-lacking group of sequences. The time scale (on the right) in Myr (million years).
content of particular mRNAs in embryonic and tumor tissues, as well as in cell cultures (Fig. 4C).

The products obtained by PCR-amplification of cDNA from tumorous (seminoma) and embryonic (frontal lobe) tissues, as well as from the Jurkat cell line were cloned. For each tissue/cell line sample, 20 clones were randomly selected and sequenced. The results obtained showed that in tumor and embryonic tissues mRNAs of all three LTR-containing genes were amplified. In transformed cell cultures, only two genes of three (AL592284 and AL049742) were apparently transcribed.

Very few is known about transcription of the KIAA1245 subfamily genes. Full-length RNAs of the KIAA1245 gene subfamily were detected in several tissues and tumors of adult humans. LTR(−) genes were expressed in human mammary gland, testis, and lung cancer. The predicted sizes of proteins for the transcript of AL356957 from human mammary gland and transcripts of AL592309 from lung cancer and testis are 1214, 633, and 633 amino acid residues, respectively. Those for transcripts of LTR(+) genes from human spinal cord, brain and skin are 839, 1377 and 670 amino acids, respectively.

All the genes studied here are located on human chromosome 1. LTR(−) AL592309 is localized on the short arm (1p36.12), whereas all others on the long arm. LTR(−) AL356957 is located on 1q21.2, and three LTR(+) genes on 1q21.1, 0.6 Mb from the first gene. The distances between these three genes vary within 1.6–2 Mb. Such long distances allow one to suggest that negative and positive co-regulation of the LTR(−) and LTR(+) genes, respectively, are not a consequence of genes’ close proximity, but rather caused by some intrinsic structural
features. We suggest that these features are conferred by the presence of LTRs in the gene sequences.

The promoter regions of the genes in question are poorly studied. However, using the Genomatix software, we tried to identify and compare promoters of the LTR(+) and LTR(−) genes. All promoters were characterized by many retroelement insertions. In particular, a 1.3-kb fragment upstream of the first exon, twice as long as the promoter, was shown to contain two Alu, two Line2 and one MaLR elements. The locations and class of the REs were identical for all genes except one LTR(+) gene that had an extra Alu element. The promoter regions of the KIAA1245 subfamily genes apparently have no systematic structural differences that could explain the systematic differential transcription of LTR(+) and LTR(−) genes.

Since differential transcription may be due to different methylation of promoter regulatory sequences, we compared the transcription levels of the KIAA1245 subfamily genes in line 293 cells and in the same cells but treated with 5-aza-2′-deoxycytidine, a DNA demethylating agent. PCRs with pairs of primers specific to (i) all genes of the KIAA1245 subfamily (Pr1–Pr4), (ii) only LTR-containing genes (Pr1–Pr2), and (iii) only LTR-lacking genes (Pr1–Pr3) showed that the transcription

![Fig. 3. Transcription of LTR-containing and LTR-lacking genes of the KIAA1245 subfamily in different tissues and cell lines. A, tumors and adult brain; B, embryonic brain; C, cell lines; D, controls used in each PCR reaction. (+)C, plus control, PCR amplification with discriminating primers and genomic DNA as template; (−)C, minus control, PCR amplification with (-RT) cDNA as template. Pr1–Pr2, amplification of only LTR(+) genes; Pr1–Pr3, only LTR(−) genes; Pr1–Pr4, both LTR(+) and LTR(−) genes. The number of amplification cycles is indicated under each lane.](image-url)
level of the genes, in particular LTR(+) genes, was the same both in normal and demethylated cells (unpublished). It seems that the DNA methylation level did not affect the tissue specificity of transcription of the genes under study.

We used transient expression assay to test the enhancer activity of KIAA1245 LTRs cloned in both orientations in a pGL3-Promoter Vector. The results are illustrated in Fig. 5. It can be seen that in Tera1 cells, where the transcription of the LTR(+) genes is markedly higher than of the LTR(−) genes (see Table 1), the enhancer activity of the used KIAA1245 LTR in the forward orientation is approximately 20-fold higher than that of the SV40 promoter, while in the reverse orientation it is comparable to the enhancer activity of SV40 enhancer. These data substantiate our conclusion that differential expression of LTR(+) and LTR(−) KIAA1245 gene subfamily members is due to LTRs rather than to other factors.

Summarizing the data above, it can be concluded that the presence of LTRs plays an important role in regulation of transcriptional activity of the KIAA1245 subfamily genes. Although the number of genes in the subfamily is limited, the same selectivity was observed in all studied tissues: the LTR(+) genes were active in contrast to the LTR(−) genes. This enhances statistical significance of the result.

The data obtained demonstrate an interesting phenomenon of a specific LTR effect on gene expression. This effect takes place in the same intracellular microenvironment and is, therefore, most probably due to intrinsic cis-characteristics of integrated LTRs. An important difference with known cases of the involvement of LTRs in transcription regulation of neighboring genes is that the LTRs here reside in a gene intron and most probably do not function as promoters.

More plausible seems the suggestion that the inserted LTRs in our case function as tissue specific enhancers acting at embryonic stages of development. Such insertions may thus confer new traits to the organisms that could be the subject of selection in primate evolution. Future studies will hopefully shed light on causes and consequences of the phenomenon observed.

### Materials and methods

#### Cell lines

Jurkat (TIB-152; acute T cell leukemia), NT2/D1 (CRL-1973; pluripotent human testicular embryonal carcinoma) and Tera 1 (HTB-105; human testicular malignant embryonal carcinoma) cell lines were cultured in media recommended by ATCC, NGP-127 (neuroblastoma cell line) and RMS13 (rhabdomyosarcoma cell line) were kindly provided by Paul S. Meltzer and grown as described previously (Elkahloum et al., 1996). All cell lines were cultured with 10% fetal calf serum (FCS) at 37 °C and 5% CO2.

#### Tissue samples

Samples of human tumors (seminoma) were provided by Blokhin Cancer Research Center, Russian Academy of Medical Sciences, Moscow. Tumor was sampled from orchidectomy specimens with Testicular Germ Cell Tumors under non-neoplastic conditions. Representative samples were divided into two parts, one of them being immediately frozen in liquid nitrogen, and the other formalin-fixed and paraffin-embedded for histological analysis. Fetal brain samples were autopsy material obtained after a therapeutic abortion at gestation age 24 weeks. The sampling was made by a written consent of the patient and approved by the ethical committees of the Institute of Bioorganic Chemistry RAS and the Research Center for Obstetrics, Gynecology and Perinatology RAMS.

#### Oligonucleotide primers

Oligonucleotide primers for PCR amplification (Table 2) were synthesized using an ASM-102U DNA synthesizer (Laboratory of Structure and Functions of Human Genes, Institute of Bioorganic Chemistry, Moscow, Russia).

#### RNA isolation and RT-PCR

Total RNA was isolated from frozen tissues pulverized in liquid nitrogen and from cell lines using an RNeasy Mini RNA purification kit (Qiagen). All RNA samples were further treated with DNase I to remove residual DNA. cDNA synthesis was performed according to the manufacturer’s protocol using random hexamer primers (Perkin Elmer), with (+RT) or without (−RT) addition of PowerScript II reverse transcriptase (Clontech). The efficiency of cDNA synthesis was equal in all preparations, as verified by RT-PCR with 3′-actin gene specific primers (Gene Checker Kit, Invitrogen).

A cDNA equivalent to 20 ng total RNA was used as template in each PCR reaction. PCR was performed in a final volume of 50 μl using unique primers (Table 2). 10 μl of the reaction

---

### Table 1

Transcription of the KIAA1245 subfamily genes in different embryonic tissues and cell lines

<table>
<thead>
<tr>
<th>Embryonic tissues and cell lines</th>
<th>PCR cycles&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Pr1/Pr4</th>
<th>Pr1/Pr2</th>
<th>Pr1/Pr3</th>
<th>N/D&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basal ganglion</td>
<td>36</td>
<td>36</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal lobe</td>
<td>36</td>
<td>36</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occipital cortex</td>
<td>33</td>
<td>33</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothalamus</td>
<td>36</td>
<td>36</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalamus</td>
<td>33</td>
<td>33</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hippocampus</td>
<td>36</td>
<td>36</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seminoma No. 1</td>
<td>33</td>
<td>33</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seminoma No. 2</td>
<td>33</td>
<td>33</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jurkat</td>
<td>33</td>
<td>33</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tera 1</td>
<td>33</td>
<td>33</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT2/D1</td>
<td>36</td>
<td>36</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMS</td>
<td>33</td>
<td>33</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGP</td>
<td>33</td>
<td>33</td>
<td>N/D</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pr1/Pr4, Pr1/Pr2, and Pr1/Pr3, primer pairs used.

<sup>a</sup> The number of PCR amplification cycles sufficient to make the product detectable.

<sup>b</sup> N/D—the product was not detected after 39 PCR amplification cycles.
mixture after 24, 27, 30, 33, 36 and 39 cycles of amplification was analyzed by electrophoresis in 0.8–2.0% agarose gels. Gel images obtained by charge-coupled device (CCD) camera systems were quantitated using Gel-Pro analysis software (Media Cybernetics). All RT-PCR experiments were reproduced at least twice using independent cDNA preparations.

**Genomic PCR**

10 ng of genomic DNA purified from tissues or cell cultures was PCR-amplified with specific primers (Table 2) under the same conditions as RT-PCR.

**Plasmid construction**

Routine genetic engineering manipulations were carried out according to the published protocols (Sambrook et al., 1989). PrLTR For and PrLTR Rev primers (Table 2), corresponding to the HERV(K) LTR flanking regions within the AL954711 gene locus, were used for PCR amplification of the full-sized LTR from human genome DNA as template. Amplification ofLTRs was performed using native Pfu DNA polymerase (Stratagene).
E. coli pGL3-PV LTR For and pGL3-PV LTR Rev, were propagated in promoter sequence. The resulting constructs, designated as pGL3-Promoter Vector (Promega) upstream of the SV40 Gel purified PCR fragments were cloned in a SmaI digested in standard PCR conditions recommended by the manufacturer. experiments (error bars) is presented.

promoter (pGL3-PV) and lacking LTR. Standard error of the mean for three reverse orientation (pGL3-PV LTR Rev) or the SV40 promoter and the LTR in either forward (pGL3-PV LTR For) or is shown as a ratio of the luciferase activity expressed by the plasmids containing gene locus) in transiently transfected Tera 1 cells. The relative enhancer activity Transfection and luciferase assay

in standard PCR conditions recommended by the manufacturer. Gel purified PCR fragments were cloned in a Smal digested pGL3-Promoter Vector (Promega) upstream of the SV40 promoter sequence. The resulting constructs, designated as pGL3-PV LTR For and pGL3-PV LTR Rev, were propagated in E. coli DH5α, isolated using a Wizard Plus Midiprep DNA Purification System (Promega, USA), and checked for mutations by sequencing.

The following plasmids were used for transient transfection: pGL3-PV, a plasmid containing only the SV40 promoter; pGL3-CV, a plasmid containing both the SV40 promoter and enhancer; pGL3-PV LTR For and pGL3-PV LTR Rev, plasmids containing the SV40 promoter and the LTR in the forward and reverse orientation, respectively. Transient transfections of Tera 1 cells were performed using Lipofectin (Invitrogen) as recommended by the manufacturer. The cells (2×10⁶ per 25 cm² flask) were washed with DMEM/RPMI (1:1), 10 mM HEPES, and then 1 ml of the same medium containing 3 μg of the luciferase expression plasmid, 1 μg of a pEGFP-N1 (Clontech) cotransfected plasmid and 10 μl of Lipofectin per flask was added. The cells were incubated for 5–6 h and then cultured for 48 h in 5 ml of the corresponding growth medium, harvested and lysed. The cell lysates obtained were analyzed for luciferase activity using a Luciferase Assay System (Promega) followed by normalization to this content. Transient transfection experiments were repeated three times.

Homologous sequences were searched against available databases using the BLAST program on the National Center for Biotechnology Information (NCBI) web site [http://ncbi.nlm.nih.gov/blast]. The search was done using Non-redundant and HTGS (contains incomplete sequences of the human genome) Databases of GenBank+EMBL+DDBJ (EST Division).

The genomic context of sequence hits was determined with the help of the RepeatMasker program [http://ftp.genome. washington.edu/cgi-bin/RepeatMasker]. Cellular genes neighboring LTRs in question were searched using the Human Genome BLAST [http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html] and Human Genome Browser Draft [http://genome.ucsc.edu] programs.

Distance templates were computed with the DNADIST program and the Kimura matrix [http://www.bic.nus.edu.sg:8887/phylip/dnadist.html]. A tree of the closest homologs was constructed using the FITCH program of the PHYLIP package. Divergence timing estimations for two LTR-containing groups were done with the help of the Clustal_w win program modified by T. Gorodentseva.

A search for promoter regions and binding sites of transcription factors was performed using the Genomatix project (http://www.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Promoter software for search and analysis of promoter regions (htp://www.genomatix.genomatix.de) that includes the Gene2Pro...