Embryonic stem cell virus, a recombinant murine retrovirus with expression in embryonic stem cells

(retroviral vectors/enhancer/virus host range)

Manuel Grez*1, Ercan Akgün, Frank Hilberg1, and Wolfram Ostertag

Heinrich-Pette-Institut für Experimentelle Virologie und Immunologie, Abteilung Zellbiologie, Martinistrasse 52, 2000 Hamburg 20, Federal Republic of Germany

Communicated by Howard M. Temin, August 6, 1990 (received for review May 21, 1990)

ABSTRACT The expression of Moloney murine leukemia virus and vectors derived from it is restricted in undifferentiated mouse embryonic carcinoma and embryonal stem (ES) cells. We have developed a retroviral vector, the murine embryonic stem cell virus (MESV), that is active in embryonal carcinoma and ES cells. MESV was derived from a retroviral mutant [PCMV(1)-passaged myeloproliferative sarcoma virus (PCMV)] expressed in embryonal carcinoma cells but not in ES cells. The enhancer region of PCMV was shown to be functional in both cell types, but sequences within the 5' untranslated region of PCMV were found to restrict viral expression in ES cells. Replacement of this region by related sequences obtained from the dl-s87rev virus (20) allowed LTR-mediated expression of retroviral genomes in ES cells, independent of the chromosomal site of integration.

MATERIALS AND METHODS

Cell Culture and Viral Infections. The ES cell line CCE was obtained from M. Evans (Department of Genetics, Cambridge University, Cambridge, U.K.) at passage 10, grown on mitomycin-treated primary fibroblast feeder layers, and frozen at passage 11. The undifferentiated state of the cells was monitored by immunofluorescence as described (17). For viral infections ES cells were grown on gelatin-coated plates in buffalo-rat-liver-cell-conditioned medium (21). Infection of fibroblasts and ES cells was performed as described (17). Neomycin-resistant colonies were counted after 12–14 (NIH 3T3 and CCE) or 8–9 days (ES.D3) in selective medium.

For DNA and RNA analysis of neomycin-resistant (neo8) NIH 3T3 and ES cells, cells were infected with virus-containing cell supernatant and selected in G418-containing medium (0.2 mg/ml). The neo8 clones were pooled after 10–12 days in selective medium and expanded. For DNA and RNA analysis of infected and nonselected cells, NIH 3T3 and CCE cells were infected six times with virus supernatants during a 1-week period. Cells were then kept in culture for 3 days before analysis.

Recombinant Plasmids and Viruses. mos−-neo8-PCMV was derived from neo8-PCMV (18) by deletion of a 2030-bp HindIII fragment containing the mos oncogene. Plasmids Mo-MuLV-CAT and PCMV(BX)/MLV-CAT have been described (18). To generate PCMV(BX)/Dde-CAT, a 72-bp Kpn I–Dde I fragment obtained from neo8-PCMV was inserted between the Kpn I site and the Xho I site (which is located 5' from the chloramphenicol acetyltransferase (CAT) gene) in PCMV(BX)/MLV-CAT. PCMV(BX)/BstNI-CAT, PCMV(BX)/Balp-CAT, and PCMV(BX)/PstI-CAT were constructed by insertion of a 143-bp Kpn I–BstI/N fragment, a 178-bp Kpn I–Bgl I fragment, or a 492-bp Kpn I–Pst I fragment.

Abbreviations: EC, embryonal carcinoma; ES, embryonal stem; Mo-MuLV, Moloney murine leukemia virus; MPSV, myeloproliferative sarcoma virus; PCMV, PCC4-cell-passaged MPSV; LTR, long terminal repeat; neo8, neomycin resistant; CAT, chloramphenicol acetyltransferase; MESV, murine embryonic stem cell virus.
*To whom reprint requests should be addressed.

Present address: Institute of Molecular Pathology, A-1030 Vienna, Austria.

Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 9202–9206, December 1990
Cell Biology
DNA and RNA Analysis. Southern blot analysis of genomic DNA was performed as described (18). Total cellular RNA was prepared by the guanidine isothiocyanate method (23). For primer extension analysis, 0.5 pmol of a 32P-labeled CAT-specific oligonucleotide (5'-CCATTTTAGCTTCCTTAGCTTCCT-3') was annealed to 50 μg of total RNA. Reverse transcription and analysis of the reaction products were done as described (24). For the preparation of antisense riboprobes, a 517-bp EcoRV–SpeI I fragment (positions 230 to +287) from mos−-neoR-PCMV and a 522-bp EcoRV–SpeI I fragment (positions 230 to +292) from P5Gneo were cloned into the Bluescript II vector. RNA probes were synthesized from linearized plasmids with T7 RNA polymerase and used without further purification. Hybridization of the riboprobes to total cellular RNA, treatment of the hybrids, and analysis of the radiolabeled RNA were performed according to standard protocols (24).

RESULTS

Expression of PCMV Vectors Is Restricted in ES Cells. To investigate the expression properties of PCMV vectors in ES cells, mos−-neoR-PCMV, a PCMV-derived vector, was used. In this vector, expression of the neoR gene is driven from the viral 5' LTR. Promoter efficiency can thus be estimated from the number of colonies resistant to the antibiotic G418 obtained after viral infection and drug selection.

The efficiency of neoR transfer to fibroblasts with that of mos−-neoR-PCMV was compared with that of mos−-neoR-MPSV, which does not confer G418 resistance to PCC4 cells (17, 18) and, therefore, was not expected to express the neoR gene in ES cells. The efficiency of neoR transfer to fibroblasts was similar for both viruses (Table 1). However, infection of the embryonic stem cell line CCE with mos−-neoR-PCMV resulted in a low number of neoR colonies (Table 1), comparable to that obtained with mos−-neoR-MPSV. The fact that PCMV vectors are efficiently expressed in EC but not in ES cells suggests that other or more stringent restriction mechanisms to viral expression must operate in ES cells. Consequently, we decided to dissect the PCMV genome to define the cis-acting elements involved in the restriction of viral gene expression in ES cells.

The Enhancer Region of PCMV Is Functional in ES Cells. We have shown (18) that a chimeric LTR containing the enhancer region of PCMV (positions –507 to –150) and the promoter region (positions –150 to +34) of the Moloney murine leukemia virus (Mo-MuLV) is functional in F9 and PCC4 cells. This construct [PCMV(BX)/MLV-CAT; Fig. 1A] was used to determine the activity of the PCMV enhancer in ES cells. A similar plasmid containing the Mo-MuLV enhancer and promoter elements (Mo-MuLV-CAT) served as a negative control for CAT expression in ES cells. Both DNAs were transfected into NIH 3T3 and CCE cells with a plasmid conferring resistance against the antibiotic G418 (pMC1neo, ref. 22). From a pool of Geneticin-resistant

<table>
<thead>
<tr>
<th>Virus</th>
<th>neoR, cfu/ml</th>
<th>Rat-1</th>
<th>ES.CCE</th>
<th>ES.CCE/Rat-1 ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>mos−-neoR-MPSV</td>
<td>2.5 × 10^3</td>
<td>1.5</td>
<td>6.0 × 10^-4</td>
<td></td>
</tr>
<tr>
<td>mos−-neoR-PCMV</td>
<td>1.0 × 10^4</td>
<td>0.5</td>
<td>5.0 × 10^-4</td>
<td></td>
</tr>
</tbody>
</table>

Transfer of neoR to fibroblasts and ES cells with MPSV and PCMV vectors. Virus constructs have been described (18). Values represent the total number of neoR colonies per ml of virus calculated from the number of clones obtained at the end-point dilution of the virus supernatants. cfu, Colony-forming units.

![Fig. 1](image-url) Sequences at the 5' untranslated region of the PCMV genome block viral expression in ES cells. (A) Schematic structure of the plasmid constructs used for functional analysis of the PCMV leader region. Open boxes, viral LTRs; stippled boxes, CAT gene. Fragments from the leader regions of PCMV or dl-587rev were inserted between the viral LTR and the CAT gene. Numbers beneath the constructs show the expected length in nucleotides of the product of the primer-extension reaction. Also indicated are the restriction sites used to generate leader fragments with sequential deletions at the 3' end. The leader sequences obtained from dl-587rev are indicated with a dashed line to reflect nucleotide differences relative to the PCMV sequences. The name of the individual constructs is shown at the right. CAP, cap site. (B) Primer-extension analysis of RNA obtained from stably transfected CCE cells containing the CAT constructs shown in A. ES cells were transfected with pMC1neo (22) and Mo-MuLV-CAT (lane 1), PCMV(BX)/MLV-CAT (lane 2), PCMV(BX)/DdeI-CAT (lane 3), PCMV(BX)/BstN-CAT (lane 4), PCMV(BX)/Balp-CAT (lane 5), or PCMV(BX)/Bal87-CAT (lane 6). Total RNA was extracted from stably transfected ES cells and analyzed for the presence of CAT-specific transcripts. Numbers at the right indicate the expected position of the products of the primer-extension reaction. Lane M contains marker DNA; sizes in bp are to the left.
colonies, RNA was isolated and analyzed for the presence of CAT transcripts by primer extension using a CAT-specific oligonucleotide.

Primer extension of correctly initiated CAT transcripts should result in a 98-nucleotide reverse-transcribed product. Both constructs were equally active in NIH 3T3 cells (data not shown). However, the PCMV(BX)/MLV-CAT construct was 10-fold more active in ES cells than the Mo-MuLV-CAT construct, as judged by the steady-state levels of CAT RNA present in the transfected cells (Fig. 1B, lanes 1 and 2, respectively). These results thus indicate that mutations which allow enhancer activity in EC cells also permit expression in ES cells and that restriction to PCMV gene expression in ES cells is most likely mediated by sequences within the transcribed region of the viral genome.

Intragenic Sequences Block the Expression of PCMV Vectors in ES Cells. To investigate the effect of intragenic sequences on the expression of neo<sup>−</sup>PCMV in ES cells, a 5'→5' deletion analysis of the leader sequences of mos<sup>−</sup>neo<sup>−</sup>PCMV was done. For this, DNA fragments obtained from the 5' untranslated region of the viral genome were inserted between the LTR and the CAT gene in plasmid PCMV(BX)/MLV-CAT (Fig. 1A). All fragments contained a common 5' end at position +34 and convenient restriction sites were used to generate sequential deletions at the 3' end. The CAT constructs were transfected into NIH 3T3 and CCE cells and assayed as mentioned above. All constructs were expressed at similar levels in NIH 3T3 cells (data not shown). The inclusion of leader sequences from positions +34 to +109 into PCMV(BX)/MLV-CAT did not affect expression in ES cells (Fig. 1B, lane 3). However, inclusion of sequences downstream of position +109 completely abolished expression of the CAT gene in ES cells (Fig. 1B, lanes 4 and 5). The inhibitory region was mapped between positions +109 and +180.

Replacement of the Leader Region of PCMV by Functionally Similar Sequences Allows Expression of the CAT Gene in ES Cells. The 5' untranslated region of the viral genome contains elements that are essential for virus replication, packaging, and RNA splicing (25). Any changes within this region must be, therefore, be compatible with these functions. For this reason we searched for naturally occurring replication-competent retroviral mutants containing extensive alterations within the 5' untranslated region. One such mutant is dl-587rev, a revertant of an integration-defective Mo-MuLV mutant (20). The leader region of dl-587rev contains multiple point mutations including five out of six mutations present in the leader region of PCMV but absent from Mo-MuLV (2). To test if the 5' untranslated region of dl-587rev would permit expression from the PCMV LTR in ES cells, a 180-bp DNA fragment (positions +34 to +214) isolated from the mutant virus was inserted within the PCMV LTR and the CAT gene [PCMV(BX)/Bal587-CAT in Fig. 1A]. A plasmid containing a similar fragment obtained from the PCMV leader region was constructed in parallel [PCMV(BX)/Balp-CAT]. Primer-extension analysis of the steady-state levels of CAT RNA in stable transfectants showed the presence of a 260-nucleotide reverse-transcribed product in ES cells transfected with PCMV(BX)/Bal587-CAT recombinants but not in cells transfected with PCMV(BX)/Balp-CAT (Fig. 1B, lanes 5 and 6).

The construct PCMV(BX)/Bal587-CAT contains the viral splice and tRNA primer binding sites. However, it lacks sequences required in cis for packaging of viral RNA. These sequences have been mapped to a 350-bp fragment downstream of position +214 in the Mo-MuLV genome (27). Therefore, it was necessary to test if these sequences are compatible with virus expression in ES cells. Inclusion of dl-587rev sequences up to position +564 in PCMV(BX)/MLV-CAT also did not inhibit expression from the PCMV LTR in ES cells (data not shown).

A PCMV-Derived Virus Containing the Leader Region of dl-587rev Confers G418 Resistance to NIH 3T3 and ES Cells with Similar Efficiencies. The previous findings suggested that the introduction of dl-587rev leader sequences into the mos<sup>−</sup>neo<sup>−</sup>PCMV genome should allow retroviral-mediated expression in ES cells. Therefore, a 530-bp Kpn<sup>+</sup>Int I DNA fragment (positions +34 to +564) including part of the R region, the U5 region, and the entire 5' untranslated region from mos<sup>−</sup>neo<sup>−</sup>PCMV was replaced with the corresponding dl-587rev sequences (Fig. 3). The resulting construct was named P5Gneo. Fibroblasts as well as the ES cell lines CCE

<table>
<thead>
<tr>
<th>Kpn I</th>
<th>R</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mo-MuLV</td>
<td>GTACCCGGTCATTACAAATACCTCCTTCAGCTT-GCATCGACATCTGTGGCTGTCATTCGGG</td>
<td>***</td>
</tr>
<tr>
<td>PCMV</td>
<td>GTACCCGGTCATTACAAATACCTCCTTCAGCTTGACATCTGTGGCTGTCATTCGGG</td>
<td>* * *</td>
</tr>
<tr>
<td>dl587rev</td>
<td>GTACCCGGTCATTACAAATACCTCCTTCAGCTTGACATCTGTGGCTGTCATTCGGG</td>
<td>+40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>US</th>
<th>PBS (Pro)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAGGGCTCTCTCCTGTGATTGATGTCAACCGGTCAGGGGGGCTTTTTCTTTTTTTCGTGGGCTCGGATTGCCAGCC</td>
<td></td>
</tr>
<tr>
<td>GAGGGCTCTCTCCTGTGATTGATGTCAACCGGTCAGGGGGGCTTTTTCTTTTTTTCGTGGGCTCGGATTGCCAGCC</td>
<td>+100</td>
</tr>
<tr>
<td>PBS (Gln)</td>
<td>+160</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bal 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAGGGCTCTCTCCTGTGATTGATGTCAACCGGTCAGGGGGGCTTTTTCTTTTTTTCGTGGGCTCGGATTGCCAGCC</td>
</tr>
</tbody>
</table>

Fig. 2. Nucleotide sequence comparison between the Kpn I-Bal 1 leader fragments of PCMV, Mo-MuLV, and dl587-rev. Stars denote nucleotide differences between the sequences. The boundaries of the R and US region are indicated. The binding sites for the tRNA primer (tRNA<sub>Pro</sub> for Mo-MuLV and PCMV and tRNA<sub>Gln</sub> for dl-587rev) are indicated by lines above or beneath the sequences. The nucleotide numbering refers to the cap site as position +1. Position +160 contains the G→A transition present in the EC host range mutant B2 (26).
and D3 were infected with virus stocks prepared from P5Gneo. neoR transfer to NIH 3T3 cells by either P5Gneo or mos−neoR-PCMV was similarly efficient. However, P5Gneo conferred G418 resistance to ES cells at a frequency that was about 1000 times higher than that of the parental virus and only 1.5–10 times lower than to NIH 3T3 cells (Table 2). P5Gneo thus confers neomycin resistance to fibroblasts and ES cells with similar efficiencies and, therefore, was renamed "murine embryonic stem cell virus" (MESV).

Expression of MESV in ES Cells Is Mediated by the Viral Enhancer Elements. Although the site of virus integration can have profound effects on the transcription of integrated proviruses (26), the titer of MESV on ES cells suggested that expression of MESV in these cells was mediated by the viral transcriptional control regions and was not significantly influenced by sequences flanking the proviral integration site.

To test this assumption, NIH 3T3 and CCE cells were infected repeatedly with MESV or mos−neoR-PCMV and expanded without selection prior to the preparation of DNA and RNA. A G418-selected population of infected NIH 3T3 and CCE cells was assayed in parallel. RNA was analyzed for the presence of correctly initiated viral transcripts by the RNase protection assay (24). For this, riboprobes were prepared from linearized plasmid templates containing viral sequences between the EcoRV (position −230) and the Spe I (position +287/+292) restriction sites. Accordingly, hybridization of these probes to viral-specific RNA followed by RNase treatment of the hybrids should protect 287 and 292 nucleotides in samples prepared from infected cells

Table 2. P5Gneo confers G418 resistance to ES cells with high efficiency

<table>
<thead>
<tr>
<th>Virus</th>
<th>NIH 3T3</th>
<th>CCE</th>
<th>D3*</th>
<th>ES/NIH 3T3 ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>P5Gneo</td>
<td>1.5 × 10^5</td>
<td>1.0 × 10^5</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6 × 10^6</td>
<td>9.4 × 10^4</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5 × 10^6</td>
<td>2.1 × 10^4</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.2 × 10^3</td>
<td>1.0 × 10^3</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>mos−neoR-PCMV</td>
<td>3.3 × 10^3</td>
<td>25</td>
<td>7.5 × 10^-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.1 × 10^3</td>
<td>83</td>
<td>2.7 × 10^-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.3 × 10^4</td>
<td>41</td>
<td>3.2 × 10^-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.4 × 10^3</td>
<td>1.7</td>
<td>3.1 × 10^-4</td>
<td></td>
</tr>
</tbody>
</table>

Virus titrations were performed as described in the legend to Table 1. The ES cell lines used for infection had been maintained for 12–14 (CCE) and 8–12 (D3) passages in culture.

RNA preparation from mos−neoR-PCMV infected and non-selected CCE cells did not protect the riboprobe against RNase digestion (Fig. 4A, lane 7). A 287-nucleotide protected fragment was observed in RNA samples prepared from mos−neoR-PCMV-infected and G418-selected as well as in infected and nonselected NIH 3T3 cells (Fig. 4A, lanes 4 and 6) but was barely visible in infected and drug-selected CCE cells (Fig. 4A, lane 5). In contrast, a 292-nucleotide fragment was protected against RNase digestion by RNA obtained from MESV-infected and nonselected NIH 3T3 and CCE cells (Fig. 4A, lanes 13 and 14, respectively) as well as in samples obtained from neoR cells (Fig. 4A, lanes 11 and 12). Besides the 292-nucleotide fragment, a 280-nucleotide fragment was also protected by RNA prepared from MESV-
infected cells. However, this fragment was also observed when total RNA obtained from noninfected cells was used for the protection assay (Fig. 4A, lane 10). Most likely, this fragment reflects the presence of endogenous transcripts with partial homology to the probe used.

Quantitative analysis of the steady-state levels of MESV RNA in infected cells showed that the amounts of viral RNA per proviral copy were equal in G418-selected cells and in nonselected cells, indicating that expression of MESV in ES cells is mediated by the viral enhancer elements and thus independent of the chromosomal site of integration.

**DISCUSSION**

Retroviral-mediated expression is blocked in embryonic cell lines (EC and ES cells). By using a genetic approach, retroviral mutants that are expressed in EC cells have been isolated (7, 17, 26). However, attempts to isolate retroviral mutants with an extended host range to ES cells have failed (unpublished results). It was, therefore, necessary to define retroviral regions involved in the restriction of viral expression in ES cells and then replace these regions by elements compatible with viral expression in these cells. For this we used one of our EC-host range mutant viruses, the PCMV virus, which is expressed in F9 and PCC4 cells (17, 18). Transcription analysis of recombinants containing various regions of the viral genome showed that the PCMV LTR is functional in ES cells and that the 5′ untranslated region of the virus was the major determinant in the restriction of viral expression in ES cells (Fig. 1). Replacement of these sequences by the leader sequences of dl-587rev (20) resulted in MESV, a recombinant virus that confers neomycin resistance to fibroblasts and ES cells with similar efficiencies (Table 2).

One crucial aspect in the development of MESV was the observation that the LTR of the parental virus (PCMV), in contrast to that of Mo-MuLV, is functional in ES cells. However, expression of the CAT gene in ES cells was completely blocked by the inclusion of leader sequences between the PCMV LTR and the reporter gene (Fig. 1). The requirement of this region for conserved functional features that are indispensable for the viral life cycle makes it necessary to replace the 5′ untranslated region of PCMV with functionally equivalent sequences. The retroviral mutant dl-587rev differs extensively in the leader sequences from the wild-type virus and shows some sequence homologies with PCMV that may be important for expression in ES cells (Fig. 2). The reassembled PCMV virus with the leader sequences of dl-587rev (MESV) was expressed efficiently in ES cells, as estimated from the number of G418-resistant colonies obtained (Table 2) and the steady-state levels of viral RNA in infected and nonselected cells (Fig. 4A).

The amount of viral RNA per viral copy in ES cells was estimated to be 60–80 times lower than in NIH 3T3 cells (Fig. 4A). In contrast, the number of G418-resistant colonies observed after MESV infection of fibroblasts and ES cells is similar (Table 2). One possible explanation for this is that the level of expression observed in ES cells is above a threshold required to render cells G418 resistant. Increased levels of expression thus should not result in an increase in the number of G418-resistant cells.

Although infection of EC or ES cells with Mo-MuLV-based vectors containing the neo5 gene results in efficient viral integration, only a very small fraction of the infected cells express the neo5 gene. Mutations within regulatory regions or promotion of transcription by sequences adjacent to the viral integration site accounted for the rare cases of activation of the neo5 gene (5, 6, 26, 28–30). MESV thus differs from Mo-MuLV-based vectors in that expression of MESV in EC and ES cells is mediated by the 5′ LTR and independent of the chromosomal site of integration.

We are especially indebted to S. Goff for providing the dl-587rev virus and to A. Gossler for performing the viral titrations on the D3 cell line. We thank C. Stocking, J. Novok, and K. Harbers for critical comments on the manuscript. This work represents part of the doctoral thesis of E.A. (Faculty of Biology, University of Hamburg). This work was supported by grants from the Deutsche Forschungsgemeinschaft and the Stiftung Volkswagenwerk. The Heinrich-Pette-Institut is financially supported by the Freie und Hansestadt Hamburg and the Bundesministerium für Jugend, Familie, Frauen und Gesundheit.