Use of Retroviral Vectors for Mapping of Splice Sites in Cottontail Rabbit Papillomavirus

By Nathalie Dostatni,1 Moshe Yaniv,1,* Olivier Danos2 and Richard C. Mulligan2

1Unité des Virus Oncogènes, UA CNRS 014149, Département de Biologie Moléculaire, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France and 2Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, U.S.A.

(Accepted 5 September 1988)

SUMMARY

Cottontail rabbit papillomavirus (CRPV) genomic sequences coding for virus early functions were introduced into a retroviral vector in order to produce cDNAs of the viral early region. Two constructs differing in the length of control sequences preceding the E6 open reading frame were transfected into Psi-2 cells and the released retroviral stock was used to infect NIH3T3 cells. The proviral sequences were rescued from antibiotic G418-resistant virus-infected cells after fusion with Cos cells, amplified as plasmids in Escherichia coli and analysed. Nucleotide sequencing showed that the splicing signals used in the construct containing only the early coding region are the same as in CRPV-expressing tumours, linking the beginning of E1 to the middle of E2. On the other hand, in a construct including most of the long control region a splice donor site located in the 5' end of this region, at position 7810, was very efficiently used, totally excluding the use of the donor site at position 1371. None of the constructs containing CRPV sequences transcribed from Moloney murine leukaemia virus promoter was able to transform mouse fibroblasts after DNA transfection.

The oncogenic potential of several papillomaviruses has been recognized ever since the early studies on cottontail rabbit papillomavirus (CRPV) (Rous & Beard, 1934). This virus is able to induce cutaneous papillomas in cottontail and domestic rabbits. Progression of the benign lesions into invasive carcinomas is observed in up to 25% of infected cottontail rabbits and up to 75% of infected domestic rabbits (Syverton, 1952). Its interest as a model system was recently increased by the observation that certain human papillomaviruses (HPVs) are consistently associated with malignant lesions in epidermodysplasia verruciformis patients and with cervical carcinomas (reviewed in Orth & Favre, 1985). Most of our knowledge about the transforming potential of papillomaviruses is derived from studies of the transformation of rodent fibroblasts by the bovine papillomavirus BPV-1, which is found associated only with benign tumours in cattle. CRPV infection thus provides an interesting example of natural multi-step transformation involving a viral genome associated with genetic and environmental factors, allowing the opportunity to study the oncogenic potential of papillomaviruses in vivo.

The first steps in understanding how CRPV induces transformation have been to characterize the structure of the differentially processed mRNAs produced from the viral genome. Virus-producing lesions contain transcripts from both early and late regions and all tumours studied so far contain two early major RNA species of 1.8 and 1.1 kb (Georges et al., 1984; Danos et al., 1985; Wettstein et al., 1987). These two mRNAs have different initiation sites and they are both processed at the same splicing signals. They cover part of the early region, and most of the E1 coding region is spliced out. The 1.8 kb species contains the E6 and E7 open reading frames and the product of fusion between the N-terminal part of E1 and the C-terminal part of E4 while the
Fig. 1. Construction of the retroviral vectors containing CRPV early sequences. The DraI–NsiI restriction fragment of the CRPV genome (top) was cloned using BclI and BamHI linkers into the BamHI cloning site of the DOL- vector giving the pLJCRI5 plasmid. The EcoRV–NsiI restriction fragment of the CRPV genome was cloned using BclI and BamHI linkers into the BamHI cloning site of the DOL vector giving the DOLCR12 plasmid. DOL- is the DOL vector from which the retroviral splice donor site was deleted. Recombinant plasmid construction and experiments with retrovirus-producing cell lines were done according to the French guidelines for recombinant DNA manipulation.

1.1 kb species contains E7 and the same E1–E4 fused region. It is not yet known whether these transcripts are translated as monocistronic or polycistronic mRNAs. In the first case, only the E6 and E7 open reading frames would be translated, and in the latter situation an E1–E4 fusion protein containing three amino acids encoded by E1 and 100 amino acids encoded by E4 would also be synthesized in the transformed cells.

So far, all direct attempts to detect efficient transforming activity of CRPV in vitro have not been successful. For this reason, we have approached the study of the different early gene products of the virus and of their putative biological activity by producing cDNAs in a retroviral vector system (Cepko et al., 1984; Roberts et al., 1985; Shimotohno & Temin, 1982). This system presents various advantages. First, the CRPV sequences are placed under the transcriptional control of the ubiquitous promoter of the long terminal repeat (LTR) and can be used for transformation studies. Second, the recombinant retroviruses containing processed CRPV genes can be isolated and characterized. They can also be used to introduce viral sequences into mammalian cells of various types including keratinocytes, or into animal tissue, in order to test their activity.

CRPV genomic sequences were inserted into the BamHI cloning site of the DOL or DOL- retroviral vectors (Korman et al., 1987) described in Fig. 1. DOL- is the DOL vector from which the retroviral splice donor site was deleted. The packaging cells used, Psi-2, are NIH3T3 cells containing a defective Moloney murine leukaemia virus (MoMuLV) integrated provirus which constitutively produces the gag, pol and env proteins required in trans for retroviral packaging (Mann et al., 1983). The CRPV fragment containing only the early viral region (open reading
frames E6, E7, E1, E4 and E2) which extends from position 37 to position 4295 was cloned into the BamHI cloning site of the DOL vector, giving the DOLCR12 recombinant plasmid (Fig. 1). A second construct, pLJCR15, was made with a CRPV fragment including 486 additional bp from the 3' end of the long control region (LCR) which was introduced into the BamHI cloning site of DOL− (Fig. 1).

The experimental procedures for virus production and proviral rescue were as described by Roberts et al. (1985). Briefly, the plasmids were transfected into Psi-2 cells by the calcium phosphate coprecipitation technique (Wigler et al., 1977) followed by a glycerol shock 4 h later. The supernatant of the culture was harvested after 24 h, filtered and used to infect NIH3T3 cells. Two days after infection, the antibiotic G418 was added to the culture medium at a final concentration of 800 μg/ml. Ten days later neomycin-resistant colonies were pooled and grown. At this step, titres of retroviral stocks, obtained after transfection of the plasmids DOLCR12 and pLJCR15 into Psi-2 cells, were estimated to be 100 c.f.u./ml and reached 300 or 1000 c.f.u./ml for the DOL or DOL− vectors alone. Since the recombinant proviruses integrated into the genome of the G418-resistant cells contained the simian virus 40 origin of replication, they could be excised from the chromosome after fusion with Cos cells (Botchan et al., 1980; Cepko et al., 1984) and recovered by Hirt extraction (Hirt, 1967). G418-resistant clones (100) were pooled and fused to Cos cells in a 1 : 1 ratio using polyethylene glycol 1000 (Davidson et al., 1976). After 3 days, a Hirt supernatant was prepared and used to transform competent Escherichia coli to kanamycin resistance. More than 800 kanamycin-resistant clones were obtained per plate of fused cells in the first experiment done with the DOLCR12 construct whereas only 30 clones were obtained in a second experiment where pLJCR15 was used. These differences are probably due to variable efficiencies of cell fusion as well of bacterial transformation.

Plasmids from more than 140 kanamycin-resistant clones in experiments using the DOLCR12 construct and from 30 kanamycin-resistant clones using the pLJCR15 construct were further analysed using restriction endonucleases. This analysis allowed us to detect rearrangements in vector sequences. Such rearrangements were found in a large proportion of rescued plasmids from the DOLCR12 construct (45% of the clones analysed) and from the pLJCR15 construct (25%) (Fig. 2). They probably occurred during the excision/recombination process in fused cells and were not further analysed. Among the remaining plasmids, one species was found exclusively in the experiment done with DOLCR12 (21% of the clones analysed) and contained unspliced CRPV sequence (form B1, Fig. 2) whereas three other species carried partially deleted CRPV sequences. Two of these three species were detected with the DOLCR12 construct. The first one retained only the 590 bp 3' exon covering the E2/E4 region (29% of the analysed clones); it was produced by a splicing event between the MoMuLV donor site (form A, Fig. 2), present only in DOLCR12, and the CRPV acceptor site at position 3714. The second species had an altered pattern indicative of a splicing event that removed the E1 coding region (5%), as observed in CRPV-induced tumours that involved a splice donor site at position 1371 and the splice acceptor site at position 3714 (form C1, Fig. 2). The position of this junction was confirmed by sequence analysis (results not shown). With the pLJCR15 construct, only one kind of structure (75% of the clones analysed) showed a deletion in the CRPV sequence. This deletion removes a large sequence between the BglII restriction site (position 7659) in the long control region and a Smal restriction site (position 3981) at the end of the E2/E4 reading frames (form D, Fig. 2).

We identified precisely the junction of this deletion by the dideoxy chain termination sequencing method after cloning the Smal-BglII fragment into an M13 vector. The nucleotide sequence of this fragment was also checked by the chemical procedure (Maxam & Gilbert, 1977) using 5'-32P-end-labelled fragments. The sequence (Fig. 3a) shows that the deletion results from a splicing event between the CRPV acceptor site at position 3714 and a donor site localized in the LCR at position 7810. The sequence around this donor site shows homologies with BPV-1 at position 7385 of the late donor site described recently (Fig. 3b) for three different late RNA species (Baker & Howley, 1987). The nucleotide sequence around the splice junction shows that the second exon contains the 3' end of the two overlapping E2/E4 open reading frames with only one ATG codon 99 bp from the 3' end of E2. The first exon is non-coding and contains an ATG
Fig. 2. Structure of the different spliced species. (a) The structure of the LCR and early region of CRPV. Given below are the structure of the DOLCR12 and pLJCR15 recombinant plasmids, the donor or acceptor splice sites described in this work and the expected unspliced and spliced derivatives. Plasmids DOLCR12 gave rise to forms A, B1 and C1, and plasmid pLJCR15 to form D but not to forms B2 and C2. Splice donor site, △; splice acceptor site, ▲: CRPV early region, ■: CRPV long control region, □: retroviral Psi sequence, ▶: retroviral Psi sequence deleted for the retroviral donor site, ◀. (b) Panel A indicates the frequency and the percentage of the different forms observed with each construct relative to the total number of clones analysed. Panel B gives the percentages of the alternative forms observed with each construct relative to the number of clones analysed showing a specific splicing event.

codon at its 3' end. This initiation codon is not in frame with either 5' truncated E2 or E4 reading frames of the second exon, and it is followed by only 11 codons. In the case of CRPV, this cDNA would code for small peptides of 33 and 11 amino acids. This result differs from that observed with the homologous spliced products of BPV-1 where the same ATG codon at the 3' end of the first exon can be used to translate two truncated products of the E2 or E4 open reading frames (Baker & Howley, 1987). However no evidence was presented for the existence of such translation products.
Fig. 3 (a) Sequence of the spliced CRPV cDNA around the splice junction in the form D. The first exon is non-coding with an ATG initiation codon at its 3' end, which is not in frame with either E2 or E4 5'-truncated open reading frames. Stop codon, []; ATG codon, D. (b) Sequence homology following the BPV-1 7385 and the CRPV 7810 splice donor sites.

The G418-resistant NIH3T3 clones that we obtained after infection did not show any characteristics of transformed cells such as focus formation or change in cell morphology. Since the recombinant retroviral vectors place the intact viral early region in a transcriptional unit driven by an efficient promoter they enable testing of the transforming potential of CRPV by direct DNA transfection of different cell lines. We first deleted the entire polyoma virus early region of the pLJCR15 and DOLCR12 vectors giving rise to pLJCR15exc and the DOLCR12exc, respectively. This latter construction had a deletion of the retroviral donor-site sequence, introduced in order to prevent artefactual splicing events such as those described above. The resulting plasmid is designated pLJCR20exc. C127 or BALB/c 3T3 cells were transfected by calcium phosphate coprecipitation with 5 μg of each construct. After 2 days, the cells were separated into four plates and grown with or without G418 selection. Ten days later, neomycin-resistant colonies were counted and the morphology of these in the G418-containing plates was examined whereas the drug-free plates were scored for transformed foci. The DOL− vector and DOL−exc from which the polyoma virus early region had been excised were used
respectively as positive and negative controls for transformation. As shown in Table 1, we only detected foci with the positive control DOL\(^{-}\). In these conditions, the CRPV early genes do not have any detectable transforming activity even if they are placed under the control of a strong heterologous promoter. These experiments were repeated several times and gave identical results.

In the present work, we have introduced either the early region alone or the non-coding plus early regions of the CRPV genome into retroviral vectors. We have shown that specifically spliced copies are rescued after infection of NIH3T3 cells with a stock of these recombinant retroviruses produced in Psi-2 cells. The removal of the E1 coding region located between positions 1371 and 3714 is the major splicing event in the CRPV-expressing tumour VX2 (Danos et al., 1985). We show here that the same splicing signals (form C1, Fig. 2) are used in mouse fibroblasts transfected with the recombinant retroviral plasmid DOLCR12 which contains only the viral early region. The rather low efficiency of these splicing signals (9% of the plasmids resulting from specific maturation) is unexpected if compared with their high efficiency in the VX2 tumours. Unspliced mRNA has not been detected in these cells. Many of the molecules obtained with the DOLCR12 construct (53%) resulted from the artefactual junction of the donor sequence present in the MoMuLV leader to the CRPV acceptor at position 3714 (form A, Fig. 2). However, this rate of utilization of the retroviral donor is certainly underestimated, since such processed RNAs lack the Psi sequence and are packaged at a reduced frequency (Mann & Baltimore, 1985). Therefore, in the context of this particular construct and cell system, the MoMuLV donor is much more efficient than the CRPV donor at position 1371.

In a second experiment using the pLJCR15 construct in which the retroviral donor site is deleted and which contains additional sequences from the non-coding region, we were able to identify a specific splicing event between the CRPV acceptor site at position 3714 and a CRPV donor site localized in the LCR at position 7810, available in RNAs transcribed from an upstream promoter (Wettstein et al., 1987). This splicing event could well be related to the formation of a double spliced mRNA with a third exon containing the late L1 or L2 reading frames as is indeed the case for HPV-6 and HPV-11 (Chow et al., 1987) and for BPV-1 (Baker & Howley, 1987). Despite the presence in the pLJCR15 construct of the CRPV donor site at position 1371, we were not able to monitor its utilization (form C2, Fig. 2). This result indicates that the CRPV 7810 donor site is more efficient than the CRPV 1371 site in this system. Likewise, it should be noted that the CRPV sequence in this region contains many potential acceptor sites next to the active one at position 3714 (positions 3724, 3735, 3746, 3762 and 3774). They are used at very low levels in VX2 cells (Danos et al., 1985) and are not used at detectable levels in mouse fibroblasts in the two experiments described here.

### Table 1. Test for the transforming activity of retroviral vectors containing CRPV or polyomavirus sequences

<table>
<thead>
<tr>
<th>Transfected cells†</th>
<th>Plasmid</th>
<th>G418-resistant clones</th>
<th>Foci</th>
<th>Without selection*</th>
<th>Foci</th>
</tr>
</thead>
<tbody>
<tr>
<td>C127</td>
<td>DOL(^{-})</td>
<td>28, 96</td>
<td>14, 32</td>
<td></td>
<td>96, 90</td>
</tr>
<tr>
<td></td>
<td>DOL(^{-})-exc</td>
<td>51, 63</td>
<td>0, 0</td>
<td></td>
<td>0, 0</td>
</tr>
<tr>
<td></td>
<td>plJCR15exc</td>
<td>44, 48</td>
<td>0, 0</td>
<td></td>
<td>0, 0</td>
</tr>
<tr>
<td></td>
<td>plJCR20exc</td>
<td>12, 24</td>
<td>0, 0</td>
<td></td>
<td>0, 0</td>
</tr>
<tr>
<td>BALB/c 3T3</td>
<td>DOL(^{-})</td>
<td>52, 105</td>
<td>27, 20</td>
<td></td>
<td>41, 52</td>
</tr>
<tr>
<td></td>
<td>DOL(^{-})-exc</td>
<td>300, 450</td>
<td>0, 0</td>
<td></td>
<td>0, 0</td>
</tr>
<tr>
<td></td>
<td>plJCR15exc</td>
<td>300, 450</td>
<td>0, 0</td>
<td></td>
<td>0, 0</td>
</tr>
<tr>
<td></td>
<td>plJCR20exc</td>
<td>300, 450</td>
<td>0, 0</td>
<td></td>
<td>0, 0</td>
</tr>
</tbody>
</table>

* The number of G418-resistant clones and foci with or without selection are given for two separate plates.

† Cells were transfected by calcium phosphate coprecipitation with 5 μg of the plasmid.
Consequently it appears that a strong selection occurs in both natural and heterologous systems for the utilization of the different splicing signals. However, we have analysed only the detailed structure of the major splicing events and we cannot exclude the possibility that other minor events occurring in mouse fibroblasts are not represented in our sample. Furthermore, naturally used splicing signals could be inefficient in this cell system, as observed for instance in similar experiments done with the E1A gene of adenovirus 5 (Roberts et al., 1985).

Finally, the present experiments demonstrate that cloning of rabbit papillomavirus sequences beyond a strong promoter does not reveal their transforming potential; similar observations have been made with certain human papillomaviruses (Matlashewski et al., 1987). This can be explained by the formation of RNA species with a deleted viral E6 and E7 region but containing most of the LCR. However, the experiments with plasmid pLJCR20exc which efficiently synthesizes the E6-E7 mRNA strongly suggest that the mouse fibroblasts do not easily respond to the putative transforming genes of CRPV.

We are grateful to L. V. Crawford for the gift of BALB/c3T3 cells and to F. Thierry for valuable discussions. This work was supported by grants from the Centre National de la Recherche Scientifique, the Institut National pour la Sante et la Recherche Medicale, the Association pour la Recherche sur le Cancer and the Ligue Nationale Francaise contre le Cancer.

REFERENCES


*(Received 9 May 1988)*