

Identification and Nucleotide Sequence of the Thymidine Kinase Gene of Shope Fibroma Virus

C. UPTON AND G. McFADDEN*

Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

Received 1 August 1986/Accepted 28 August 1986

The thymidine kinase (TK) gene of Shope fibroma virus (SFV), a tumorigenic leporipoxvirus, was localized within the viral genome with degenerate oligonucleotide probes. These probes were constructed to two regions of high sequence conservation between the vaccinia virus TK gene and those of several known eucaryotic cellular TK genes, including human, mouse, hamster, and chicken TK genes. The oligonucleotide probes initially localized the SFV TK gene 50 kilobases (kb) from the right terminus of the 160-kb SFV genome within the 9.5-kb *Bam*HI-*Hind*III fragment E. Fine-mapping analysis indicated that the TK gene was within a 1.2-kb *Ava*I-*Hae*III fragment, and DNA sequencing of this region revealed an open reading frame capable of encoding a polypeptide of 176 amino acids possessing considerable homology to the TK genes of the vaccinia, variola, and monkeypox orthopoxviruses and also to a variety of cellular TK genes. Homology matrix analysis and homology scores suggest that the SFV TK gene has diverged significantly from its counterpart members in the orthopoxvirus genus. Nevertheless, the presence of conserved upstream open reading frames on the 5' side of all of the poxvirus TK genes indicates a similarity of functional organization between the orthopoxviruses and leporipoxviruses. These data suggest a common ancestral origin for at least some of the unique internal regions of the leporipoxviruses and orthopoxviruses as exemplified by SFV and vaccinia virus, respectively.

Thymidine kinase (TK) (EC 2.1.7.21) is an important enzyme in the salvage pathway of procaryotic and eucaryotic cells and is responsible for ATP-dependent phosphorylation of thymidine to thymidine 5'-monophosphate (20). The TK gene is particularly useful in the study of eucaryotic virus genetics because selection procedures exist for both the TK⁺ and TK⁻ phenotypes. Many, if not all, poxviruses encode the TK gene as a dispensable gene for growth in tissue culture, and the vaccinia virus TK gene in particular has been sequenced and extensively characterized (1, 8, 11, 13-20, 34, 35). Although the TK genes of two closely related orthopoxviruses, monkeypox virus and variola virus, have been sequenced and shown to be very similar to that of vaccinia virus (9), nothing is known about their counterparts in other poxvirus genera.

Shope fibroma virus (SFV) is a member of the genus leporipoxvirus and induces characteristic fibromas in rabbits (10, 31). SFV grows well in tissue culture and is amenable to analysis at the molecular level since the physical map of the viral DNA has been recently deduced and the complete genomic library in plasmid vectors is available (5, 7, 36). Members of the poxvirus family replicate exclusively in the cytoplasm of infected cells and probably encode most of the enzymes necessary for the synthesis of viral DNA and RNA (for reviews, see references 6, 28, 29, and 37). It has previously been shown that SFV induces virus-specific TK activity (2), but whether the new activity was virus encoded or an induced cellular activity could not be rigorously ascertained. We therefore surveyed the SFV genome for the viral TK gene with the following objectives in mind. (i) DNA probes to the genomes of orthopoxviruses, such as vaccinia virus, do not cross-hybridize to the DNA genomes from members of the leporipoxvirus genus, but detailed inspection of the amino acid sequences of a counterpart gene such as the TK gene might provide clues as to their evolutionary relationship. (ii) The mapped TK gene would provide a

convenient locus for genetic manipulation of the SFV genome to investigate properties unique to SFV, such as the capacity to induce target cell proliferation. (iii) It has been recently suggested that a subset of the DNA sequences within the SFV terminal inverted repeats (TIRs) may have been acquired from rabbit cells (33) and, since the DNA sequences are known for a variety of eucaryotic TK genes (4, 22-25), specific knowledge of the SFV TK gene sequence might shed light on the possible origin of the SFV genome itself.

Here we report on the use of degenerate oligonucleotide probes to detect the SFV TK gene and provide evidence for a close evolutionary relationship to other poxvirus and eucaryote TK genes.

MATERIALS AND METHODS

Enzymes and reagents. Restriction enzymes were purchased from Boehringer Mannheim, Bethesda Research Laboratories, Pharmacia, and Amersham and used under conditions recommended by the suppliers. T4 DNA ligase, T4 DNA polymerase, and exonucleases III and VII were from Boehringer Mannheim or Bethesda Research Laboratories. T4 polynucleotide kinase was from Pharmacia. [α -³²P]dATP and [γ -³²P]ATP were supplied by New England Nuclear; all of the other reagents used for dideoxy-DNA sequencing were from New England BioLabs.

Oligonucleotides and hybridization conditions for detection of the TK gene. Oligonucleotide pool no. 1 GG(A/G/T/C)CCCATGTT(T/C)TC(A/G/T/C)GG and oligonucleotide pool no. 2 GA(T/C)GA(G/A)GG(G/A)CA(G/A)TT(T/C)TT were synthesized by the Regional DNA Synthesis Laboratory, University of Calgary, Alberta, Canada (see Fig. 1 for strategy details). The procedures for 5' end labeling with T4 polynucleotide kinase and [γ -³²P]ATP were as described elsewhere (27). Southern and dot blots were prehybridized with 0.9 M NaCl-90 mM Tris (pH 7.5)-6 mM EDTA-0.5% sodium dodecyl sulfate-100 μ g of yeast tRNA per ml-10 μ g of calf thymus DNA per ml-5 \times Denhardt solution at 34°C for

* Corresponding author.

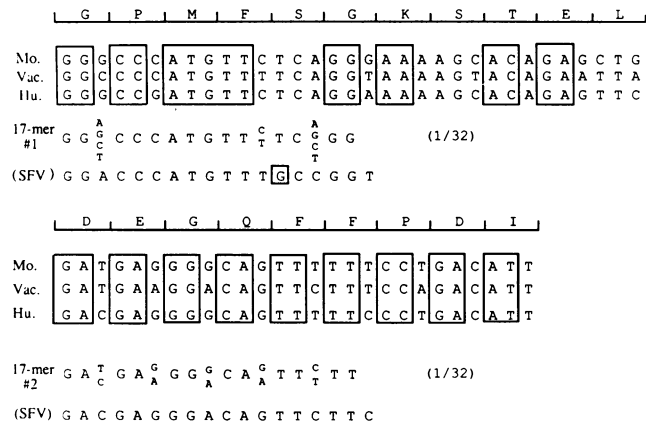


FIG. 1. Oligonucleotide strategy for cloning the SFV TK gene. The amino acid and DNA sequences of two regions conserved among the mouse (Mo), vaccinia virus (Vac) and human (Hu) TK genes are shown. The one-letter amino acid code is used, and the two peptide stretches correspond to residues 26 to 36 and residues 97 to 105 of the human TK gene sequence shown in Fig. 7. The boxed G nucleotide in the SFV sequence (Fig. 7), when compared with pool no. 1, turned out to be the only mismatch observed.

4 h and then hybridized under the same conditions with 5'-labeled synthetic oligonucleotides for 16 h. Blots were washed with 6x SSC (1x SSC is 0.15 M NaCl plus 0.015 M sodium citrate, pH 7.0)-0.1% sodium dodecyl sulfate at room temperature for 30 min.

Cloning protocols. The cloning and restriction enzyme mapping of SFV (strain Kasza) *Bam*HI-*Hind*III fragment E (B/H-E) has been described previously (7, 36). The 1.3-kilobase (kb) *Hind*III-*Hpa*II fragment of vaccinia virus WR (34) containing the intact TK gene was cloned into pKCR (36) and used as a positive control for hybridization with the oligonucleotide probe (see Fig. 2).

DNA sequencing. The Sanger dideoxy nucleotide chain termination method was used to sequence unidirectional nested deletions generated by exonuclease III in both insert orientations as described elsewhere (33, 33a).

Analysis of DNA sequences. DNA sequence management, database searches, and alignment of polypeptide sequences were performed with the CORE library programs of BIONET (IntelliGenetics Inc.). The homology matrix analysis used the DNA Inspector II program (Textco) set for strings of 18 nucleotides with up to six mismatches.

RESULTS

Mapping of the SFV TK gene. Preliminary experiments to localize the SFV TK gene by Southern blotting by using cloned vaccinia TK as a probe were negative, even with moderate stringencies of hybridization and washing (unpublished data). As an alternate procedure we investigated using degenerate oligonucleotide probes to known conserved regions of the published TK gene sequences. Alignment of vaccinia virus, mouse, and human TK gene polypeptide sequences revealed three stretches of greater than 10 amino acids conserved among these three proteins. Both the DNA and amino acid sequences of two of these regions (Fig. 1) were used to select the sequence of oligonucleotide probes, but in regions of ambiguity greater emphasis was placed on the vaccinia virus sequence. A 17-mer degenerate oligonucleotide (a pool of 32) was synthesized for each of the two conserved regions. The sequence of these oligonucleotides

(Fig. 1) was derived from the region from glycine 26 to leucine 36 (pool no. 1) and from aspartic acid 97 to isoleucine 105 (pool no. 2) of the human TK gene sequence. The boxed nucleotides within the mouse, vaccinia virus, and human DNA sequences are those fixed by the genetic code if the amino acid sequence is to be maintained. To illustrate the degree of match derived from this approach, we also included the corresponding nucleotide sequences from the SFV TK gene (see below) in Fig. 1. For probe no. 2, a complete 17-base match with the deduced SFV TK gene was contained within the oligonucleotide pool, whereas for probe no. 1 16 of 17 nucleotides was the best possible match.

We used probes made with these oligonucleotide pools to survey a dot blot containing cloned restriction fragments from the complete SFV genome. Since both probe pools produced identical results, only results with probe pool no. 2 are displayed here. In Fig. 2, the cloned *Bam*HI fragments B to T plus the three subclones *Bam*HI-*Hind*III-E, *Bam*HI-*Hind*III-J, and *Hind*III-E) of *Bam*HI fragment A were hybridized with probe pool no. 2. Included on the blot were the different vector DNAs and the positive control, pKTK-15, which contains the 1.3-kb *Hind*III-*Hpa*II fragment encoding the entire vaccinia virus TK gene (34). A positive hybridization signal was seen only with the clone containing the 9.5-kb SFV fragment *Bam*HI-*Hind*III-E, a subclone of the 27.5-kb *Bam*HI fragment A. This particular subclone maps within

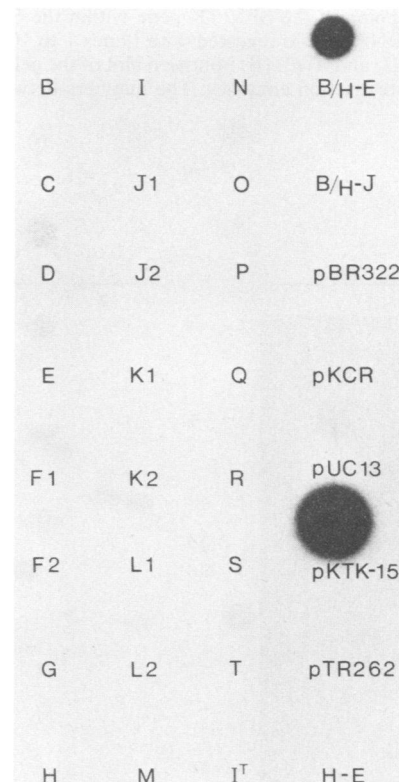


FIG. 2. Dot blot of cloned fragments of the SFV genome probed with ³²P end-labeled oligonucleotide no. 2. Clones in the first three columns contained 10 ng of plasmids containing *Bam*HI fragments B to T of SFV (7, 36). In column 4, *Bam*HI-*Hind*III-E (B/H-E), *Bam*HI-*Hind*III fragment J (B/H-J), and *Hind*III fragment E (H-E) make up the 27.5-kb *Bam*HI A fragment (7). pBR322, pKCR, pUC13, and pTR262 are the cloning vector DNAs. Clone pKTK-15 is a positive control and contains the 1.3-kb *Hind*III-*Hpa*II fragment of vaccinia virus WR encoding the intact vaccinia virus TK gene.

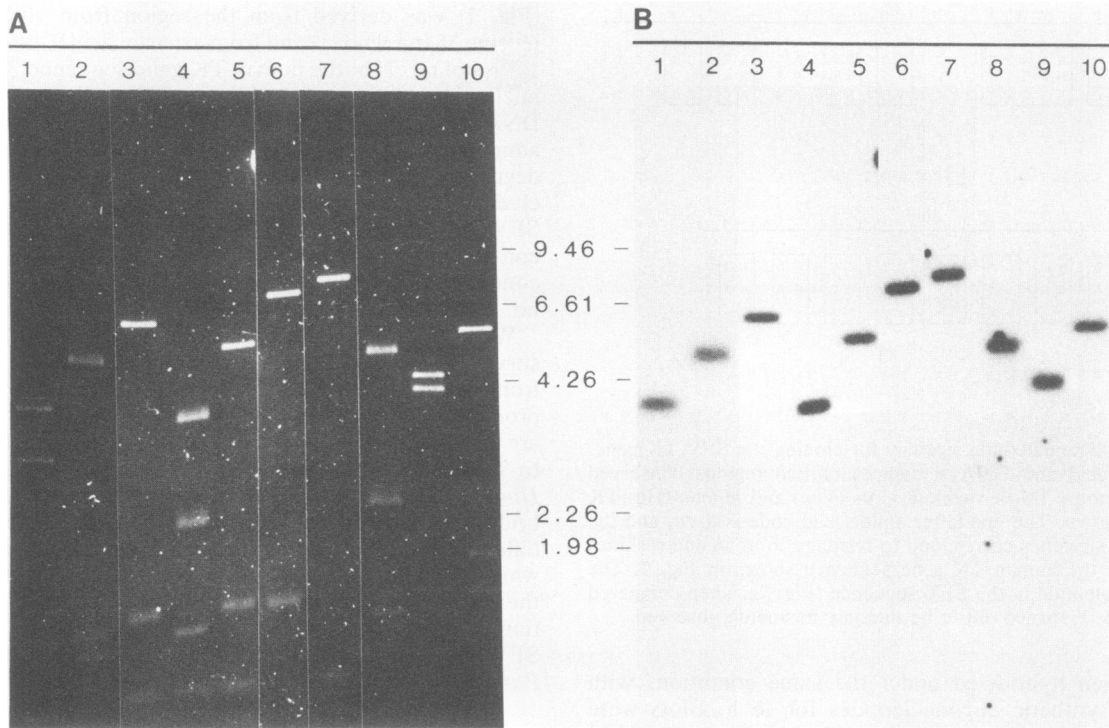


FIG. 3. Mapping of the SFV TK gene within the SFV 9.5-kb *Bam*HI-*Hind*III-E (B/H-E). (A) Ethidium bromide-stained agarose gel of purified SFV DNA B/H-E digested with (lanes 1 to 10, respectively) *Bgl*II+*Xho*I, *Cla*I+*Xho*I, *Ava*I+*Pst*I, *Cla*I+*Pst*I, *Xho*I+*Pst*I, *Xho*II, *Pst*I, *Cla*I, *Bgl*II, and *Ava*I. (B) Southern blot of the gel in A probed with end-labeled oligonucleotide pool no. 2. See Fig. 10 for map positions of the various restriction enzymes. The numbers between the panels indicate the sizes (in kilobases of λ DNA digested with *Hind*III).

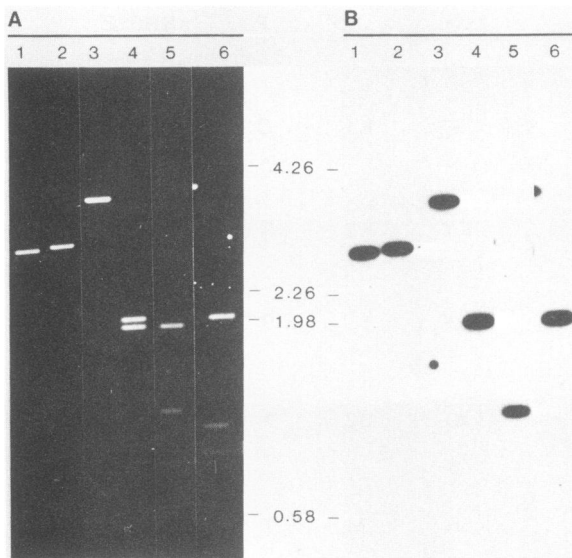


FIG. 4. Fine mapping of the SFV TK gene within the SFV 4.1-kb *Pst*I-*Cla*I fragment. (A) Ethidium bromide-stained agarose gel of the purified 4.1-kb *Pst*I+*Cla*I fragment from SFV B/H-E digested with (lanes 1 to 6, respectively) *Ava*I, *Bgl*II, no enzyme, *Hae*III, *Ava*I+*Hae*III, and *Bgl*II+*Hae*II. (B) Southern blot of the gel in A probed with oligonucleotide no. 2. Enzyme map positions are indicated in Fig. 10. The numbers between the panels are as described in the legend to Fig. 3.

the unique sequences of the SFV genome approximately 50 kb from the right terminal hairpin (7). By using the single *Pst*I site previously mapped at the left end of fragment *Bam*HI-*Hind*III-E for orientation, we used restriction enzyme analysis and Southern blotting to locate the putative SFV TK gene within a 4.1-kb *Pst*I-*Cla*I fragment (Fig. 3, lane 4). This *Pst*I-*Cla*I fragment was purified and subjected to a further round of restriction enzyme digestion and Southern blotting, which localized the region homologous to probe no. 2 within a 1.2-kb *Ava*I-*Hae*III fragment (Fig. 4, lane 5; see also Fig. 10 for the final map position).

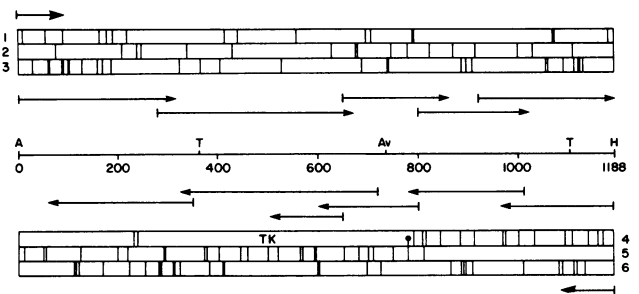


FIG. 5. Sequencing strategy and ORF analysis in the region of the SFV TK gene. The 1,188-base-pair *Ava*I+*Hae*III fragment (see Fig. 10) was sequenced as described in Materials and Methods. Restriction enzymes: A = *Ava*I, Av = *Ava*II, H = *Hae*III, T = *Taq*I. The presumptive initiating AUG of the TK gene is indicated (●). Large arrows at the top and bottom indicate the direction of transcription. The numbers 1 through 6 indicate the six possible open reading frames.

```

10      20      30      40      50      60      70
GGCCGCTCAG GACGATGATG TTGCAATGTC TGACGTACAC ACGATCGTTG AGTATTAAAA CTTTTTACTA

80      90      100     110     120     130     140
CGGTTGTTAA TTAATTCGAA AGATAAACTA GAAGCAGTAG GATATTATTA CGCTCCTCTA TCAGAACAAT

150     160     170     180     190     200     210
ACAAAGCCGT GTTTGATTTT ACAAACACAA AGTCGTTGAA ACAGTTGTTT AACAGACAAC CCGTACACAT

220     230     240     250     260     270     280
TGAGAGTGAG TCTCCCATTT CCGTGGACAA GGGATATTTG GCGGATTTTG TTCTTGCAAC GACTAGATTA

290     300     310     320     330     340     350
AAGAAACAAC TAAATTTGAC GTTAGATAAG GACGTTACGT ACGTAGATCC GTATACGGAT AAAAGGTCCG

360     370     380     390     400     410     420
CGAATATATT GTCTATATTG CATAAAAAC TGAAGTAGTA AAACATTATT TGCAAAATTTA AATCATGACA

435     450     465
ATG TAC GGG GGA CAT ATT CAC CTC ATT ATA GGA CCC ATG TTT GCC CGT AAA AGC
MET Tyr Gly Gly His Ile His Leu Ile Ile Gly Pro Met Phe Ala Gly Lys Ser

480     495     510     525
ACG GAA CTA ATT CGT CTA GTT AGA CGT TAT CAA ATA GCG AAA CAC AAA TGT CTC
Thr Glu Leu Ile Arg Leu Val Arg Arg Tyr Gln Ile Ala Lys His Lys Cys Leu

540     555     570
GTT GTA AAA TAC GAA AAA GAC ATA CGT TAC GGA AAC GGT GTA TGT ACA CAT GAT
Val Val Lys Tyr Glu Lys Asp Ile Arg Tyr Gly Asn Gly Val Cys Thr His Asp

585     600     615     630
AAC ATG AGC ATA ACC GCC GTA TGT ACC CCG TCG TTG GAC AAA ATA GAC TCA GTA
Asn Met Ser Ile Thr Ala Val Cys Thr Pro Ser Leu Asp Lys Ile Asp Ser Val

645     660     675     690
GCC GAA AAC GCC GAA GTT ATT GGG ATA GAC GAG GGA CAG TTC TTC CCC AAT ATA
Ala Glu Asn Ala Glu Val Ile Gly Ile Asp Glu Gly Gln Phe Phe Pro Asn Ile

705     720     735
GCA ACG TTT TGC GAA CGT ATG GCG AAC CGT GGA AAG GTA TTG ATC GTG GCT GCG
Ala Thr Phe Cys Glu Arg Met Ala Asn Arg Gly Lys Val Leu Ile Val Ala Ala

750     765     780     795
TTA GAC GGA ACA TTT CAA CGT AAA CCA TTT AGC AAC ATT TCA GAA CTG ATA CCG
Leu Asp Gly Thr Phe Gln Arg Lys Pro Phe Ser Asn Ile Ser Glu Leu Ile Pro

810     825     840
TTG GCT GAA AAC GTA ACA AAA CTA AAC GCG GTG TGT ATG TAC TGT TAC AAG AAC
Leu Ala Glu Asn Val Thr Lys Leu Asn Ala Val Cys Met Tyr Cys Tyr Lys Asn

855     870     885     900
GGG TCT TTC TCT AAA CGA CTG GGT GAT AAA ATG GAA ATC GAA GTA ATA GGG GGT
Gly Ser Phe Ser Lys Arg Leu Gly Asp Lys Met Glu Ile Glu Val Ile Gly Gly

915     930     945     958     968
AGT GAT AAG TAC AAA TCC GTG TGT AGA AAA TGT TAT TTT TTT TAAAAACAAT GAAAAATAA
Ser Asp Lys Tyr Lys Ser Val Cys Arg Lys Cys Tyr Phe Phe ***

978     988     998     1008     1018     1028     1038
ATTAGATATT TACGAGTATG TGATTGTATT ATCTATCTCG GTGGCGTTTA GGTCACGAC AATCATGGGT

1048     1058     1068     1078     1088     1098     1108
ATACAACACA AATTGGAAGT GTTCATTGTG AGTGAACAACA TCGCTATTAA AGACGCTAAT CTACTCAATG

1118     1128     1138     1148     1158     1168     1178
GAGACAGTTA CGGATGACT ATCAAACATA AATTGAACAC AAAAAAGAGT GTCAGATTCC TTGTCCTGTT

1188
AGAACCCGAG
    
```

FIG. 6. DNA sequence of the 1,188-base-pair *AvaI*+*HaeIII* fragment with the translated amino acid sequence of the SFV TK gene. The asterisks indicate the termination codon.

Cloning and sequencing of the SFV TK gene. The *AvaI*-*HaeIII* 1.2-kb TK gene fragment was blunt ended with T4 DNA polymerase and ligated into the *SmaI* site of M13 mp18, and a series of nested deletions were created by the exonuclease III method for each of the two orientations (12), as described in Materials and Methods. These were sequenced by the dideoxy chain termination method, and the open reading frames (ORFs) for both directions were determined (Fig. 5 and 6). We also performed DNA sequencing reactions by using the two degenerate oligonucleotide pools as primers. Although both were 1:32 mixtures of oligonucle-

otides and one of the 17-mer mixes only had a maximum homology of 16/17, both nevertheless demonstrated correct priming at the expected sites within the viral TK gene sequence (data not shown).

Analysis of the SFV TK gene sequence. Examination of the six possible reading frames of the 1.2-kb *AvaI*-*HaeIII* fragment revealed the presence of only one complete long ORF consistent with a viral TK gene (Fig. 5). The DNA sequence of this fragment, from the *HaeIII* site to the *AvaI* site, together with the amino acid sequence of the presumptive TK gene, is presented in Fig. 6. A complete alignment of the SFV, vaccinia virus, human, mouse, and chicken TK gene polypeptides is shown in Fig. 7, for which the first methionine codon of the ORF was used to correctly align the SFV TK gene with the known initiating methionine of the vaccinia virus TK gene. As in the case of genes sequenced from members of the orthopoxvirus genus, the 5' flanking region preceding the SFV TK gene is very A-T rich. The upstream 70 nucleotides are 75.7% A-T and have an A-T-C-G molar base ratio of 42.8:32.9:12.9:11.4. No distinct conserved canonical sequence was located within the promoters of other SFV genes (unpublished data) in part because the A-T

```

1      20      40
MSCINLPTVLPSPSKTRGQIQVILGPMFSGKSTELMRRV HU
MSYINLPTVLPSSPSKTRGQIQVILGPMFSGKSTELMRRV MO
MNLCLTVPGVHPGSPGRPRGQIQVIFGPMFSGKSTELMRRV CH
MYGGHILHLIIGPMFAGKSTELIRLV SFV
MNGGHIQLIIGPMFSGKSTELIRRV VAC
* * * * *

60
RRFQIAQYKCLVIKYAKDTRYSSSF-CTHDRN---TMEAL HU
RRFQIAQYKCLVIKYAKDTRYNSNF-STHDRN---TMDAL MO
RRFRLAQYRCLLVKYAKDTRYCTTGVSTHRRN---TMEAR CH
RRYQIAKHKCLVVYKEDIRYGNV-CTHD-NMSITAVCT SFV
RRYQIAQYKCVTIKYSNDNRYGTGL-WTHDKN-NFEALEA VAC
** * * * *

80      100
PACLRLD-VAQEALGVAVIGIDEGQFFPDIMEFCEAMANA HU
PACMLRD-VTQELLGVAVIGIDEGQFFPDIVDFCEMAMANE MO
PACALQD-VYQEALGSVAVIGIDEGQFFPDIVDFCEKMANE CH
PSLDKIDVAENAE---VIGIDEGQFFPNIAFCEMAMANR SFV
TKLCDVLESITDFS---VIGIDEGQFFPDIVDFCEMAMANE VAC
***** * * * *

120      140
GKTIVIVAALDGTGFRKPFQAIINLNLVPLAESVVK-TAVCME HU
GKTIVIVAALDGTGFRKAFGSIINLNLVPLAESVVKLTAVCME MO
GKTIVIVAALDGTGFRKAFGSIINLNLVPLAESVVKLNVAFCMG CH
GKVLIVAALDGTGFRKPFNSISELIPLAENVTKLNAVCMY SFV
GKIVIVAALDGTGFRKPFNNIINLNLVPLAESVVKLTAVCMK VAC
** * * * *

160      180
CFREAAYTKRLGTEKEVEVIGGADKYHSVCRCLCYFKKASG HU
CFREAAYTKRLGLEKEVEVIGGADKYHSVCRCLCYFKKSSA MO
CYREASYTKRLGAEREVEVIGGADKYHSVCRACRYFQKRPQ CH
CYKNGFSKRLGDKMEIEVIGGSDKYKSVCRKCYFF SFV
CFKEASFKRLGEETEIEIIGNDMYQSVCRKCYIDS VAC
* * * * *

200      220      233
QPAGPDNKENCPVPGKPEAARLFAFQQLQCSPAN HU
QTAGSDNK-NCLVLGQPGQALVVRKLFASQVQLQNSAN MO
QL-GSENKENVPMGVKQLDMPASRKIFAS CH
* * * * *
    
```

FIG. 7. Alignment of amino acid sequences of the human (HU), mouse (MO), chicken (CH), SFV, and vaccinia virus (VAC) TK genes. Amino acid residues conserved among all five polypeptides are indicated by asterisks. (HU and MO are 233 amino acids each; CH is 223 amino acids; SFV is 176 amino acids; VAC is 177 amino acids.)

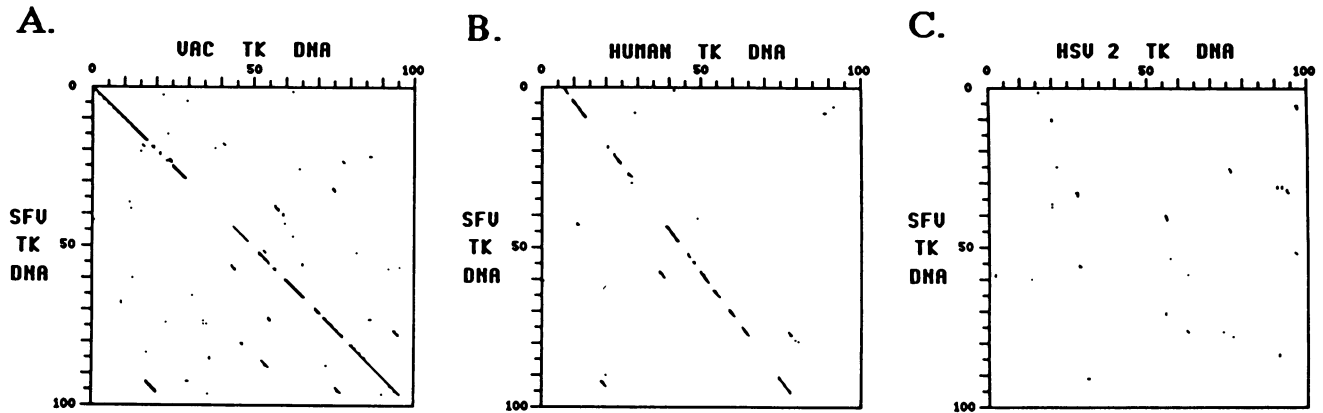


FIG. 8. Homology matrix analysis of the SFV, vaccinia virus, human, and herpes simplex virus type 2 TK gene coding sequences. We compared the nucleotide sequences from the SFV TK gene with those of the (A) vaccinia virus (VAC) (34), (B) human (4), and C herpes simplex virus type 2 (32) TK genes by using a search string length of 18 nucleotides, allowing six possible mismatches. The axis numbers refer to the percentage of the full length for each TK gene. (SFV = 528 nucleotides; vaccinia virus = 531 nucleotides; human = 702 nucleotides; herpes simplex virus type 2 = 1,128 nucleotides.)

richness of these upstream regions makes it difficult to define an accurate consensus sequence.

To assess the extent to which the SFV TK gene is related to the other known eucaryotic TK genes, we performed homology matrix analysis with the nucleic acid sequences of several representative TK genes. Figure 8A displays the observed homology between the SFV and vaccinia virus TK genes. Clear and extensive homology could be detected throughout the lengths of the two poxvirus TK genes, indicating a common origin. The example of the human TK gene versus the SFV TK gene (Fig. 8B) illustrates that the extent of homology score matches with the cellular TK gene also extends discontinuously through the full length of the coding domains. When the SFV TK gene was compared with all other known cellular TK genes (mouse, hamster, and chicken), homology patterns similar to that with the human TK gene were observed (data not shown). As first noticed in the case of the vaccinia virus TK gene by Kwoh and Engler (22), there is no obvious relationship of poxvirus TK genes to the herpesvirus TK gene (Fig. 8C). To further assess the degree to which the SFV TK gene is related to the other TK genes at the amino acid level, we calculated the percentage of identical amino acids for all of the pairwise matches (Table 1). Interestingly, the extent of identity between the SFV and vaccinia virus TK genes (65.5%) was rather less than the range of variations between the pairwise combinations of the human, mouse, chicken, and hamster TK genes (74.1 to 88.5%), suggesting that the SFV and vaccinia virus TK genes are more diverged than are the known cellular TK genes.

Analysis of SFV TK gene flanking sequences. Although there was clearly a close relationship between the SFV and vaccinia virus TK genes, very little evidence existed to suggest that these two viral genomes were similarly organized. Probes made to the two viral genomes did not cross-hybridize even under conditions of moderate stringency, and DNA sequencing of the SFV TIR (33, 33a; Upton et al., manuscript in preparation) indicated that their viral TIRs are unrelated in terms of sequence organization and encoded gene products. Nevertheless, computer analysis indicated two stretches of 80% identity in the 5' flanking sequences of the SFV and vaccinia virus TK genes (SFV, TGXXTATATT-22-AACXATXATT; vaccinia virus, TGXXTATATT-16-AACXATXATT). We therefore decided to translate all of the available upstream sequences of SFV, vaccinia virus, monkeypox, and variola virus TK genes (9, 34). Surprisingly, the long ORFs terminating immediately upstream from the TK gene in each of these four poxviruses were also found to be closely related to each other (Fig. 9). In the case of SFV, the last nucleotide of the termination codon of this upstream ORF lies 38 bases upstream of the first nucleotide of the TK gene sequence. In the vaccinia and monkeypox viruses, these distances are 18 and 20 nucleotides, respectively, whereas in variola virus the last A nucleotide of the termination codon of the upstream ORF is also the first nucleotide in the initiating ATG of the TK coding sequence. Only three small gaps are necessary to align the upstream SFV polypeptide sequence with the upstream ORFs in the other three poxvirus genomes, and although the vaccinia, variola, and monkeypox virus up-

TABLE 1. Extent of homologies among poxvirus TK gene amino acid sequences and those of eucaryotic cellular TK genes

Source of TK gene sequence	% Homology with sequences of:						
	SFV	Vaccinia virus	Variola virus	MPV ^a	Human	Mouse	Hamster
Vaccinia virus	65.5						
Variola virus	64.3	97.2					
MPV ^a	64.3	96.6	97.2				
Human	61.0	69.6	68.4	67.8			
Mouse	59.3	68.4	66.6	66.6	85.9		
Hamster	58.7	66.7	65.4	65.5	88.5	87.6	
Chicken	61.2	67.4	65.7	65.1	75.4	79.3	74.1

^a MPV, Monkeypox virus.

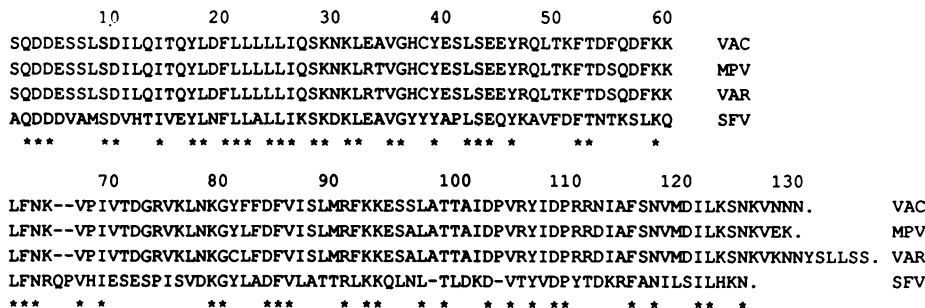


FIG. 9. Alignment of the C terminus of polypeptides translated from partial ORFs present immediately upstream of the poxvirus TK genes. Amino acid residues conserved among all four sequences are indicated by asterisks. Only the C-terminal amino acid sequences from vaccinia (VAC), monkeypox (MPV), and variola (VAR) viruses which align with the 125-amino-acid stretch from the *AvaI-HaeIII* fragment of SFV (Fig. 6) are displayed.

stream ORFs are much more closely related to each other than to SFV, 42.4% (53/125) of the C-terminal 125 residues are conserved throughout all four sequences. The level of homology between the complete upstream ORF proteins must await further sequencing studies in the four viral genomes.

DISCUSSION

One of the drawbacks of using the SFV model to analyze virus-cell interactions involving the tumorigenic poxviruses is the absence of defined genetic mutants. The TK marker has proven to be valuable in the establishment of vaccinia virus as a cloning vector and for the development of surrogate genetics in poxviruses (26, 30). Therefore it is of particular interest to identify and characterize the SFV TK gene as a potential target site for genetic manipulations. Here we report that, although DNA probes to the vaccinia virus TK gene were unable to detect the counterpart TK gene in the SFV genome under hybridization and washing conditions of very moderate stringency, degenerate oligonucleotide probes were successful in localizing and identifying the SFV TK gene. A similar protocol using degenerate oligonucleotide probes has recently been successful in identifying an SFV growth factor gene related to EGF, TGF α , and the vaccinia virus growth factor gene (W. Chang, C. Upton, S. Hu, A. F. Purchior, and G. McFadden, Mol. Cell. Biol., in press). Figure 10 shows the expanded physical map of the 160-kb SFV genome (7) and the location of the TK gene within the 9.5-kb *B/H-E*. No other related SFV se-

quences were detected, indicating that the TK gene is a single-copy gene, as it is in vaccinia virus.

Homology analysis of the deduced SFV TK gene protein sequence indicates a close relationship with the vaccinia, variola, and monkeypox orthopoxvirus TK genes and the human, mouse, hamster, and chicken TK genes. There is no obvious homology among any of the poxvirus TK genes, including the SFV, and herpes simplex virus TK gene (32). If one simply considers the percentage of identical amino acids among the various TK genes, it can be concluded that the closest relatives to the SFV TK gene are from the orthopoxvirus members vaccinia virus, variola virus, and monkeypox virus. Nevertheless, the SFV TK gene is clearly distinct and is further diverged from that of vaccinia virus than any of the cellular TK genes (human, hamster, chicken, and mouse) are from each other (Table 1). Since the natural host of SFV is the rabbit and there is some evidence that SFV may have at some point during its evolution acquired cellular sequences through cDNA intermediates found in covalently closed circular DNA species (33), it will be of interest to compare the SFV TK gene sequence with those of the rabbit TK gene and other eucaryotic TK genes as more are discovered and sequenced. In this regard, the two degenerate oligonucleotide probe pools described in this communication may be of use in identifying and isolating related TK genes from other eucaryotes.

Recent evidence suggests that the SFV genes implicated in the generation of fibromas in infected rabbits lie within a 5- to 7-kb region near the junction of the viral TIRs with unique internal sequences (3). In particular, SFV and malignant

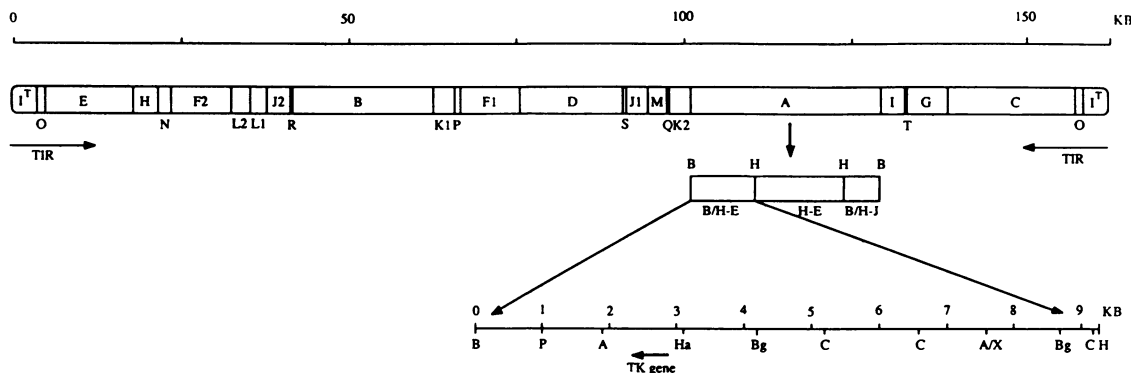


FIG. 10. Summary of mapping and DNA sequencing data for the SFV TK gene. *BamHI* fragments are shown on the genomic map of SFV. Restriction enzymes: A = *AvaI*, B = *BamHI*, Bg = *BgIII*, C = *Clal*, H = *HindIII*, Ha = *HaeIII*, P = *PstI*, X = *XhoI*. Abbreviations: TIR, terminal inverted repeat; B/H-E, *BamHI-HindIII-E*; H-E, *HindIII-E*; B/H-J, *BamHI-HindIII-J*.

Downloaded from http://jvi.asm.org/ on July 10, 2015 by UNIV OF VICTORIA

rabbit virus, a related recombinant leporipoxvirus capable of inducing fibromas in rabbits that are histologically very similar to those induced by SFV, have been shown to share less than a half dozen expressed ORFs in common (3, 33a). The identification of the SFV TK gene will now permit direct construction of TK⁻ mutants for use as recipients for reinsertion of the cloned viral TK gene into various sites of the viral TIR so as to inactivate the relevant SFV genes in a systematic fashion.

The discovery of related and highly conserved long ORFs immediately upstream to the poxvirus TK genes was unexpected. The entire 12,397-base-pair TIR of SFV has been sequenced (33, 33a; C. Upton, A. M. Delange, and G. McFadden, submitted for publication), and virtually no DNA sequence homology with the published TIR sequences of the orthopoxvirus members could be detected. Nine large ORFs have been observed in the SFV TIR, but all of them encode polypeptides possessing no known related counterparts in the published protein database (33a). Furthermore, all DNA probes to the SFV genome tested to date have uniformly been unable to elicit cross-hybridization with the vaccinia virus genome, suggesting considerable evolutionary divergence of the leporipoxvirus and orthopoxvirus genomes. Nevertheless, the sequence data presented here suggest that, whereas the TIRs of SFV and the orthopoxviruses were generated by independent events, the unique internal regions manifest sufficient organizational similarity in the region of the viral TK gene to suggest a common origin from an ancestral virus.

ACKNOWLEDGMENTS

G.M. is a scholar of the Alberta Heritage Foundation for Medical Research. This work was supported by operating grants from the National Cancer Institute of Canada and the Alberta Cancer Board. The computer resource BIONET is funded by a Public Health Service grant from the National Institutes of Health, Bethesda, Md.

We thank R. A. Maranchuk and A. Wills for technical assistance and B. Bellamy for preparing the manuscript.

LITERATURE CITED

- Bajszár, G., R. Wittek, J. P. Weir, and B. Moss. 1983. Vaccinia virus thymidine kinase and neighboring genes: mRNAs and polypeptides of wild-type virus and putative nonsense mutants. *J. Virol.* **45**:62-72.
- Barbanti-Brodano, G., M. Portolani, A. Bernardini, F. Stirpe, A. Mannini-Palenzona, and M. La Placa. 1968. Thymidine kinase activity in human amnion cell cultures infected with Shope fibroma virus. *J. Gen. Virol.* **3**:471-474.
- Block, W., C. Upton, and G. McFadden. 1985. Tumorigenic poxviruses: genomic organization of malignant rabbit virus, a recombinant between Shope fibroma virus and myxoma virus. *Virology* **140**:113-124.
- Bradshaw, H. D., Jr., and P. L. Deininger. 1984. Human thymidine kinase gene: molecular cloning and nucleotide sequence of a cDNA expressible in mammalian cells. *Mol. Cell. Biol.* **4**:2316-2320.
- Cabirac, G. F., D. S. Strayer, S. Sell, and J. L. Leibowitz. 1985. Characterization, molecular cloning and physical mapping of the Shope fibroma virus genome. *Virology* **143**:163-170.
- Dales, S., and B. G. T. Pogo. 1982. *The Biology of Poxviruses*. Springer-Verlag, New York.
- Delange, A. M., C. Macaulay, W. Block, T. Mueller, and G. McFadden. 1984. Tumorigenic poxviruses: construction of the composite physical map of the Shope fibroma virus genome. *J. Virol.* **50**:408-416.
- Dubbs, D. R., H. Otsuka, H. Qavi, and S. Kit. 1983. Mapping thymidine kinase-deficient mutants of vaccinia virus by marker rescue with hybrid plasmid DNAs containing portions of the HindIII-J fragment of virus DNA. *Virology* **126**:408-411.
- Esposito, J. J., and J. C. Knight. 1984. Nucleotide sequence of the thymidine kinase gene region of monkeypox and variola viruses. *Virology* **135**:561-567.
- Febvre, H. 1962. The Shope fibroma virus of rabbits, p. 79-111. A. J. Dalton and F. Haguenu, (ed.), *Ultrastructure of tumors induced by viruses: Charles Oberling memorial volume*. Academic Press, Inc., New York.
- Francke, C. A., N. A. Roseman, and D. E. Hraby. 1985. Expression and regulation of the vaccinia virus thymidine kinase gene in non-permissive cells. *Virus Res.* **3**:13-17.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**:351-359.
- Hraby, D. E. 1985. Inhibition of vaccinia virus thymidine kinase by the distal products of its own metabolic pathway. *Virus Res.* **2**:151-156.
- Hraby, D. E., and L. A. Ball. 1981. Cell-free synthesis of enzymatically active vaccinia virus thymidine kinase. *Virology* **113**:594-601.
- Hraby, D. E., and L. A. Ball. 1981. Control of expression of the vaccinia virus thymidine kinase gene. *J. Virol.* **40**:456-464.
- Hraby, D. E., and L. A. Ball. 1982. Mapping and identification of the vaccinia virus thymidine kinase gene. *J. Virol.* **43**:403-409.
- Hraby, D. E., and L. A. Guarino. 1984. Novel codon utilization within the vaccinia virus thymidine kinase. *Virus Res.* **1**:315-320.
- Hraby, D. E., R. A. Maki, D. B. Miller, and L. A. Ball. 1983. Fine structure analysis and nucleotide sequence of the vaccinia virus thymidine kinase gene. *Proc. Natl. Acad. Sci. USA* **80**:3411-3415.
- Hraby, D. E., D. B. Miller, and L. A. Ball. 1982. Synthesis of vaccinia virus thymidine kinase in microinjected *Xenopus* oocytes. *Virology* **123**:470-473.
- Kit, S. 1976. Thymidine kinase, DNA synthesis, and cancer. *Mol. Cell. Biochem.* **11**:161-182.
- Kit, S., G. N. Jorgensen, A. Liau, and V. Zaslusky. 1977. Purification of vaccinia virus-induced thymidine kinase activity from [³⁵S]methionine-labeled cells. *Virology* **77**:661-676.
- Kwoh, T. J., and A. Engler. 1984. The nucleotide sequence of chicken thymidine kinase and the relationship of its predicted polypeptide to that of vaccinia virus thymidine kinase. *Nucleic Acids Res.* **12**:3959-3971.
- Lau, Y. F., and Y. W. Kan. 1984. Direct isolation of the functional human thymidine kinase gene with a cosmid shuttle vector. *Proc. Natl. Acad. Sci. USA* **81**:414-418.
- Lewis, J. A. 1986. Structure and expression of the Chinese hamster thymidine kinase gene. *Mol. Cell. Biol.* **6**:1998-2010.
- Lin, P. F., H. B. Lieberman, D. B. Yeh, T. Xu, S. Y. Zhao, and F. H. Ruddle. 1985. Molecular cloning and structural analysis of murine thymidine kinase genomic and cDNA sequences. *Mol. Cell. Biol.* **5**:3149-3156.
- Mackett, M., G. L. Smith, and B. Moss. 1982. Vaccinia virus: a selectable eukaryotic cloning and expression vector. *Proc. Natl. Acad. Sci. USA* **79**:7415-7419.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- McFadden, G., and S. Dales. 1982. Organization and replication of poxvirus DNA, p. 173-190. In A. S. Kaplan (ed.), *Organization and replication of viral DNA*. CRC Press, Inc., Boca Raton, Fla.
- Moss, B. 1985. Replication of poxviruses, p. 685-703. In B. N. Fields (ed.), *Virology*. Raven Press, New York.
- Panicalai, D., and E. Paoletti. 1982. Construction of poxviruses as cloning vectors: insertion of the thymidine kinase gene from herpes simplex virus into the DNA of infectious vaccinia virus. *Proc. Natl. Acad. Sci. USA* **79**:4927-4931.
- Shope, R. E. 1932. A transmissible tumor-like condition in rabbits. *J. Exp. Med.* **56**:793-802.
- Swain, M. A., and D. A. Galloway. 1983. Nucleotide sequence of the herpes simplex virus type 2 thymidine kinase gene. *J. Virol.*

- 46:1045–1050.
33. **Upton, C., and G. McFadden.** 1986. DNA sequence homology between the terminal inverted repeats of Shope fibroma virus and an endogenous cellular plasmid species. *Mol. Cell. Biol.* **6**:265–276.
- 33a. **Upton, C., and G. McFadden.** 1986. Tumorigenic poxviruses: analysis of viral DNA sequences implicated in the tumorigenicity of Shope fibroma virus and malignant rabbit virus. *Virology* **152**:308–321.
34. **Weir, J. P., and B. Moss.** 1983. Nucleotide sequence of the vaccinia virus thymidine kinase gene and the nature of spontaneous frameshift mutations. *J. Virol.* **46**:530–537.
35. **Weir, J. P., G. Bajszár, and B. Moss.** 1982. Mapping of the vaccinia virus thymidine kinase gene by marker rescue and by cell-free translation of selected mRNA. *Proc. Natl. Acad. Sci. USA* **79**:1210–1214.
36. **Wills, A., A. M. Delange, C. Gregson, C. Macaulay, and G. McFadden.** 1983. Physical characterization and molecular cloning of the Shope fibroma virus DNA genome. *Virology* **130**:403–414.
37. **Wittek, R.** 1982. Organization and expression of the poxvirus genome. *Experientia* **38**:285–310.