JOURNAL OF VIROLOGY, Jan. 1994, p. 84–92
0022-538X/94/$04.00+0

Vol. 68, No. 1

Homologs of Vascular Endothelial Growth Factor Are Encoded by the Poxvirus Orf Virus†

DAVID J. LYTTLE, KATE M. FRASER, STEPHEN B. FLEMING, ANDREW A. MERCER,* AND ANTHONY J. ROBINSON‡

Health Research Council Virus Research Unit and University of Otago Centre for Gene Research, University of Otago, Dunedin, New Zealand

Received 23 June 1993/Accepted 24 September 1993

A gene encoding a polypeptide with homology to mammalian vascular endothelial growth factors (VEGFs) has been discovered in the genome of orf virus (OV), a parapoxvirus that affects sheep and goats and, occasionally, humans. The gene is transcribed abundantly early in infection and is found immediately outside the inverted terminal repeat at the right end of the genome. In the NZ2 strain of OV (OV NZ2), the gene encodes a polypeptide with a molecular size of 14.7 kDa, while in another strain, OV NZ7, there is a variant gene that encodes a polypeptide of 16 kDa. The OV NZ2 and OV NZ7 polypeptides show 22 to 27% and 16 to 23% identity, respectively, to the mammalian VEGFs. The viral polypeptides are only 41.1% identical to each other, and there is little homology between the two genes at the nucleotide level. Another unusual feature of these genes is their G+C content, particularly that of OV NZ7. In a genome that is otherwise 63% G+C, the OV NZ2 gene is 57.2% G+C and the OV NZ7 gene is 39.7% G+C. The OV NZ2 gene, but not the OV NZ7 gene, is homologous to the mammalian VEGF genes at the DNA level, suggesting that the gene has been acquired from a mammalian host and is undergoing genetic drift. The lesions induced in sheep and humans after infection with OV show extensive dermal vascular endothelial proliferation and dilatation, and it is likely that this is a direct effect of the expression of the VEGF-like gene.

Orf virus (OV) is a member of the *Parapoxvirus* genus of the *Poxvirus* family. It causes contagious pustular dermatitis in sheep and goats and is transmissible to humans. The biology of the parapoxviruses has been reviewed elsewhere (33, 36).

The genome of OV strain NZ2 (OV NZ2) is a linear double-stranded DNA of 139 kb (25). Sequencing of selected regions has shown that there is a significant degree of conservation of structure and arrangement of genes between OV and vaccinia virus (VAC). VAC, a member of the *Orthopoxvirus* genus, is the most thoroughly characterized poxvirus, and the Copenhagen strain has been completely sequenced (14). Homologs of genes for the VAC 14-kDa fusion protein (A27L) (37), RNA polymerase-associated factor (H4L) (1), a virion envelope antigen (H5R) (15), a type I topoisomerase (H6R) (41), a dUTPase-like polypeptide (F2L) (24, 42), and the DNA polymerase (E9L) have been found in OV (11, 26, 27, 29). Although the two genomes have very different G+C contents, OV being 63% G+C and VAC being 36% G+C (48), and the OV genome is approximately 60 kb smaller than the genome of VAC, their common evolutionary origin is apparent.

The similar locations and orientations of homologous OV and VAC genes on their respective genome maps allow the genomes to be tentatively aligned. The alignment suggests that the OV genome consists of homologs of those VAC genes that lie between the open reading frames (ORFs) F1L and B3R, as described for the Copenhagen strain of VAC (14). However, detailed analysis of 6 kb at the left end of the OV genome has shown that although structural similarities with VAC are apparent, there are significant differences (13).

We have sequenced other areas of the OV genome, including that which contains unique sequences leading up to the inverted terminal repeat (ITR) junction at the right end. Our primary reason for choosing the area at the right end was to determine the nature of a size difference found between the *Kpn*I E fragment of the OV NZ2 genome and the equivalent fragment (*Kpn*I E) in another strain of OV, designated NZ7 (OV NZ7) (34). These fragments span the ITR junction at the right end of the respective genomes, and the 0.8 kb extra DNA found in the OV NZ7 fragment was predicted to lie outside the ITR. It was considered that the variable region might offer sites for the insertion of foreign genes in the construction of OV vectors. We were also interested in finding out what relationship, if any, existed between genes found at this end of the genome with those that might be predicted from the alignment of the OV and VAC genomes.

We report here the sequence of the regions immediately outside the ITR at the right end of both the OV NZ2 and OV NZ7 genomes and describe the nature and location of the extra DNA in OV NZ7. Within this region, we found a gene, not present in VAC, with amino acid sequence homology to mammalian vascular endothelial growth factors (VEGFs) (4, 5, 8, 9, 16, 20, 22, 23, 46). Unexpectedly, the OV NZ2 and OV NZ7 VEGF-like genes show little DNA homology to each other, whereas the flanking sequences are over 98% homologous. Both genes are strongly transcribed early in infection, and it is likely that their products are responsible for the marked capillary endothelial proliferation seen in the dermis following OV infection. In addition, we provide evidence of rearrangement of the OV genome compared with that of VAC.

MATERIALS AND METHODS

Cells and viruses. Primary bovine testis cells were used for the production of OV (2), and virions were purified in sodium

* Corresponding author. Mailing address: HRC Virus Research Unit, PO Box 56, Dunedin, New Zealand. Phone: 64 3 4797730. Fax: 64 3 4792261. Electronic mail address: AAMERCER@OTAGO.AC.NZ.

† This paper is dedicated to the memory of the late Kate Fraser.

‡ Present address: Division of Wildlife and Ecology, Commonwealth Scientific and Industrial Organisation, Lyneham, ACT, Australia.

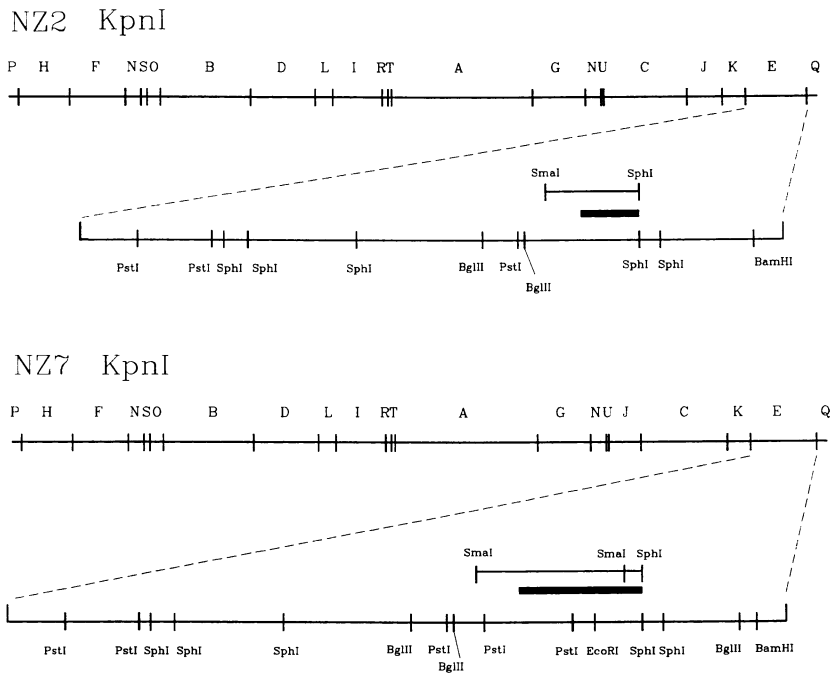


FIG. 1. Restriction endonuclease cleavage site maps of the OV NZ2 and OV NZ7 genomes. Shown for each are the *Kpn*I maps, with sites for *Bam*HI, *Bgl*II, *Eco*RI, *Pst*I, and *Sph*I shown in expanded maps of the *Kpn*I E fragments. The extent of a region of variability flanked by *Sma*I and *Sph*I sites is marked in each strain, and the regions sequenced are indicated by the bars. The bars represent 851 nt in the case of OV NZ2 and 1830 nt in the case of OV NZ7.

diatrizoate gradients (35). OV DNA was isolated from virions in guanidine HCl-CsCl gradients following protease-sodium dodecyl sulfate treatment (25). The origin of OV NZ2 has been reported elsewhere (35). OV NZ7 was obtained from L. Fastier, ICI-Tasman Laboratories, Upper Hutt, New Zealand, as cell culture-adapted virus. It was designated NZ7, as it showed an *Eco*RI restriction endonuclease pattern identical to that of the field strain NZ7 (34).

Enzymes and chemicals. Restriction endonucleases were purchased from either Boehringer GmbH (Mannheim, Germany), Promega Corp. (Madison, Wis.), Amersham International plc (Amersham, England), New England Biolabs (Beverly, Mass.), or Bethesda Research Laboratories Life Technologies Inc. (Gaithersburg, Md.) and were used according to the manufacturers' instructions.

Plasmid constructs. Restriction fragments *Kpn*I E from OV NZ2 and NZ7 were cloned into pUC19 and pTZ18R, respectively, giving plasmids pVU89 and pVU7KE51. Further restriction fragment mapping was carried out on these plasmid clones, and a 1.5-kb *Sma*I-*Sph*I fragment isolated from pVU89 was cloned into the corresponding sites of pTZ18R to give pVUSSE1. A 2.3-kb *Sma*I fragment from pVU7KE51 was cloned into the *Sma*I site of pTZ18R, giving plasmid pVU7KESB3. A second plasmid, pVU7KESA, containing a sequence contiguous to this was isolated following the cloning of a *Sma*I-*Bgl*II fragment from pVU7KE51 into pTZ18R.

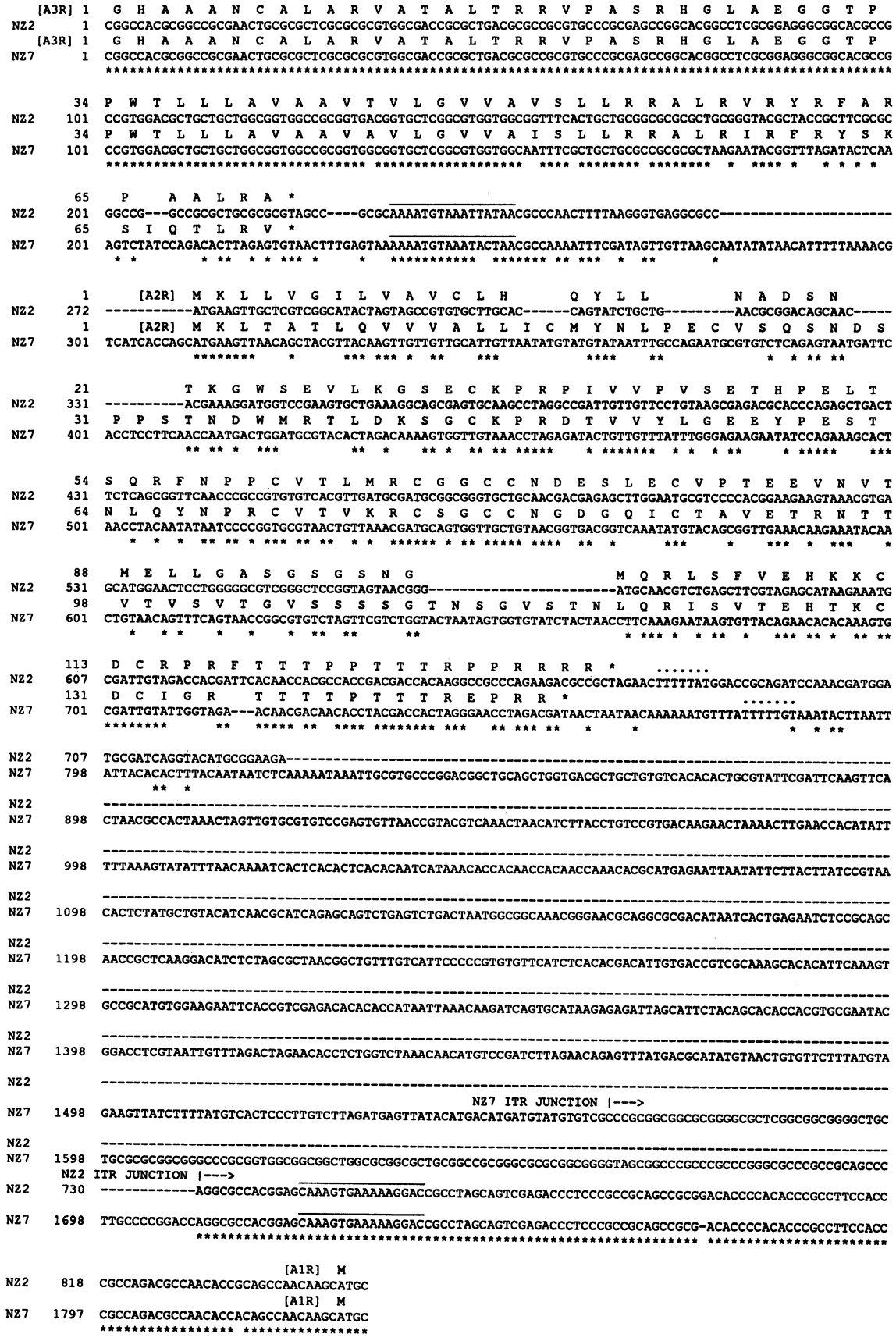
DNA sequencing and analysis. The general methods used have been described previously (13). Templates were prepared by cloning restriction fragments of pVUSSE1 and pVU7KESB3 into replicative-form M13mp18 and M13mp19, and sequencing was conducted by the dideoxynucleotide chain termination method, using commercially available kits. T7 DNA polymerase and *Taq* DNA polymerase multiwell microtiter plate sequencing kits were purchased from Amersham, and the Gene-ATAQ *Taq* DNA polymerase sequencing kit from was purchased from Pharmacia LKB Biotechnology (Uppsala, Sweden).

RESULTS

The *Kpn*I E fragments of OV NZ2 and OV NZ7 and their locations on the respective genomes are shown in Fig. 1. Restriction endonuclease cleavage site maps were deduced for each fragment (Fig. 1), and DNA-DNA hybridizations were carried out between the two strains by using subfragments (not shown). These experiments showed that the size difference could be localized to a region bounded by a *Sma*I site and a *Sph*I site in the two strains (Fig. 1) and that some fragments within this region failed to cross-hybridize. Sequencing of this region was begun from the *Sph*I site. The areas sequenced are shown in Fig. 1, and the sequences are shown in Fig. 2. The OV NZ2 sequence consists of 851 nucleotides (nt) and that of OV NZ7 consists of 1830 nt.

Alignment of the two sequences by using the program CLUSTAL (19) shows that they are 100% homologous at the 5′ ends for a distance of 133 nt. At this point, the sequences diverge until 122 nt from the 3′ end, where they again show almost 100% homology (Fig. 2). The point at which the homology abruptly resumes is at the OV NZ2 ITR junction (Fig. 2), and only two nucleotide differences are seen between the two strains in the 122 nt leading from this point to the *Sph*I site. This region of the ITR in NZ2 contains an early promoter and the initiation codon for a strongly transcribed gene designated ORF 3 (12, 13).

The homologous sequences at the 5′ end were scanned for ORFs by using the program NLDNA. Very few stop codons are present in this G+C-rich section, and ORFs coding for polypeptides greater than 30 amino acids (aa) are found in all


[A3R] 1 G H A A A N C A L A R V A T A L T R R V P A S R H G L A E G G T P
NZ2 1 CGGCCACGCGGCCGCGAACTGCGCGCTCGCGCGCGTGGCGACCGCGCTGACGCGCCGCGTGCCCGCGAGCCGGCACGGCCTCGCGGAGGGCGGCACGCCG
[A3R] 1 G H A A A N C A L A R V A T A L T R R V P A S R H G L A E G G T P
NZ7 1 CGGCCACGCGGCCGCGAACTGCGCGCTCGCGCGCGTGGCGACCGCGCTGACGCGCCGCGTGCCCGCGAGCCGGCACGGCCTCGCGGAGGGCGGCACGCCG
34 P W T L L L A V A A V T V L G V V A V S L L R R A L R V R Y R F A R
NZ2 101 CCGTGGACGCTGCTGCTGGCGGTGGCCGCGGTGACGGTGCTCGGCGTGGTGGCGGTTTCACTGCTGCGGCGCGCGCTGCGGGTACGCTACCGCTTCGCGC
34 P W T L L L A V A A V A V L G V V A I S L L R R A L R I R F R Y S K
NZ7 101 CCGTGGACGCTGCTGCTGGCGGTGGCCGCGGTGGCGGTGCTCGGCGTGGTGGCAATTTCGCTGCTGCGCCGCGCGCTAAGAATACGGTTTAGATACTCAA
65 P A A L R A *
NZ2 201 GGCCG---GCCGCGCTGCGCGCGTAGCC----GCGCAAAATGTAAATTATAACGCCCAACTTTTAAGGGTGAGGCGCC----------
65 S I Q T L R V *
NZ7 201 AGTCTATCCAGACACTTAGAGTGTAACTTTGAGTAAAAAATGTAAATACTAACGCCAAAATTTCGATAGTTGTTAAGCAATATATAACATTTTTAAAACG
1 [A2R] M K L L V G I L V A V C L H Q Y L L N A D S N
NZ2 272 ----------ATGAAGTTGCTCGTCGGCATACTAGTAGCCGTGTGCTTGCAC------CAGTATCTGCTG---------AACGCGGACAGCAAC-----
1 [A2R] M K L T A T L Q V V V A L L I C M Y N L P E C V S Q S N D S
NZ7 301 TCATCACCAGCATGAAGTTAACAGCTACGTTACAAGTTGTTGTTGCATTGTTAATATGTATGTATAATTTGCCAGAATGCGTGTCTCAGAGTAATGATTC
21 T K G W S E V L K G S E C K P R P I V V P V S E T H P E L T
NZ2 331 ----------ACGAAAGGATGGTCCGAAGTGCTGAAAGGCAGCGAGTGCAAGCCTAGGCCGATTGTTGTTCCTGTAAGCGAGACGCACCCAGAGCTGACT
31 P P S T N D W M R T L D K S G C K P R D T V V Y L G E E Y P E S T
NZ7 401 ACCTCCTTCAACCAATGACTGGATGCGTACACTAGACAAAAGTGGTTGTAAACCTAGAGATACTGTTGTTTATTTGGGAGAAGAATATCCAGAAAGCACT
54 S Q R F N P P C V T L M R C G G C C N D E S L E C V P T E E V N V T
NZ2 431 TCTCAGCGGTTCAACCCGCCGTGTGTCACGTTGATGCGATGCGGCGGGTGCTGCAACGACGAGAGCTTGGAATGCGTCCCCACGGAAGAAGTAAACGTGA
64 N L Q Y N P R C V T V K R C S G C C N G D G Q I C T A V E T R N T T
NZ7 501 AACCTACAATATAATCCCCGGTGCGTAACTGTTAAACGATGCAGTGGTTGCTGTAACGGTGACGGTCAAATATGTACAGCGGTTGAAACAAGAAATACAA
88 M E L L G A S G S G S N G M Q R L S F V E H K K C
NZ2 531 GCATGGAACTCCTGGGGGCGTCGGGCTCCGGTAGTAACGGG----------ATGCAACGTCTGAGCTTCGTAGAGCATAAGAAATG
98 V T V S V T G V S S S S G T N S G V S T N L Q R I S V T E H T K C
NZ7 601 CTGTAACAGTTTCAGTAACCGGCGTGTCTAGTTCGTCTGGTACTAATAGTGGTGTATCTACTAACCTTCAAAGAATAAGTGTTACAGAACACACAAAGTG
113 D C R P R F T T T P P T T T R P P R R R R * .....
NZ2 607 CGATTGTAGACCACGATTCACAACCACGCCACCGACGACCACAAGGCCGCCCAGAAGACGCCGCTAGAACTTTTTATGGACCGCAGATCCAAACGATGGA
131 D C I G R T T T T P T T T R E P R R * .....
NZ7 701 CGATTGTATTGGTAGA---ACAACGACAACACCTACGACCACTAGGGAACCTAGACGATAACTAATAACAAAAAATGTTTATTTTTGTAAATACTTAATT
NZ2 707 TGCGATCAGGTACATGCGGAAGA----------
NZ7 798 ATTACACACTTTACAATAATCTCAAAAATAAATTGCGTGCCCGGACGGCTGCAGCTGGTGACGCTGCTGTGTCACACACTGCGTATTCGATTCAAGTTCA
NZ2 ----------
NZ7 898 CTAACGCCACTAAACTAGTTGTGCGTGTCCGAGTGTTAACCGTACGTCAAACTAACATCTTACCTGTCCGTGACAAGAACTAAAACTTGAACCACATATT
NZ2 ----------
NZ7 998 TTTAAAGTATATTTAACAAAATCACTCACACTCACACAATCATAAACACCACAACCACAACCAAACACGCATGAGAATTAATATTCTTACTTATCCGTAA
NZ2 ----------
NZ7 1098 CACTCTATGCTGTACATCAACGCATCAGAGCAGTCTGAGTCTGACTAATGGCGGCAAACGGGAACGCAGGCGCGACATAATCACTGAGAATCTCCGCAGC
NZ2 ----------
NZ7 1198 AACCGCTCAAGGACATCTCTAGCGCTAACGGCTGTTTGTCATTCCCCCGTGTGTTCATCTCACACGACATTGTGACCGTCGCAAAGCACACATTCAAAGT
NZ2 ----------
NZ7 1298 GCCGCATGTGGAAGAATTCACCGTCGAGACACACACCATAATTAAACAAGATCAGTGCATAAGAGAGATTAGCATTCTACAGCACACCACGTGCGAATAC
NZ2 ----------
NZ7 1398 GGACCTCGTAATTGTTTAGACTAGAACACCTCTGGTCTAAACAACATGTCCGATCTTAGAACAGAGTTTATGACGCATATGTAACTGTGTTCTTTATGTA
NZ2 ----------
NZ7 1498 GAAGTTATCTTTTATGTCACTCCCTTGTCTTAGATGAGTTATACATGACATGATGTATGTGTCGCCCGCGGCGGCGCGGGGCGCTCGGCGGCGGGGGCTGC
NZ2 ----------
NZ7 1598 TGCGCGCGGCGGGCCCGCGGTGGCGGCGGCTGGCGCGGCGCTGCGGCCGCGGGCGCGCGGCGGGGTAGCGGCCCGCCCGCCCGGGCGCCCGCCGCAGCCC
NZ2 730 ----------AGGCGCCACGGAGCAAAGTGAAAAAGGACCGCCTAGCAGTCGAGACCCTCCCGCCGCAGCCGCGGACACCCCACACCCGCCTTCCACC
NZ7 1698 TTGCCCCGGACCAGGCGCCACGGAGCAAAGTGAAAAAGGACCGCCTAGCAGTCGAGACCCTCCCGCCGCAGCCGCG-ACACCCCACACCCGCCTTCCACC
NZ2 818 CGCCAGACGCCAACACCGCAGCCAACAAGCATGC
NZ7 1797 CGCCAGACGCCAACACCACAGCCAACAAGCATGC

six frames. Only one ORF showed significant homology with data base sequences (GenBank version 76), and this was to the carboxy terminus of the VAC ORF F9L. The NZ2 and NZ7 F9L homologs (ORFs A3R in Fig. 2) are virtually identical except for the terminal 10 residues.

In NZ2, the region from the stop codon of the F9L-like gene to the ITR junction is 517 nt and is relatively A+T rich compared with the rest of the genome. Within this region is an ORF flanked by a sequence typical of an early poxvirus promoter (6, 12) and a termination signal for poxvirus early transcription (12, 50) (Fig. 2). The ORF encodes a 133-aa polypeptide with a predicted molecular weight of 14,703 and is the only ORF in this region of the genome.

The OV NZ7 nonhomologous region, extending from the stop codon of the F9L-like gene to the point where homology with the OV NZ2 sequence resumes, consists of an A+T-rich sequence of 1,483 nt. This sequence has a higher A+T content than that seen in OV NZ2 except for 150 nt at its 3′ terminus, which is over 95% G+C. This G+C-rich region was reminiscent of the G+C-rich region found immediately outside the ITR at the left end of the OV NZ2 genome (13), and it was suspected that the junction between the A+T-rich and G+C-rich sequences was the ITR junction in the NZ7 strain. Restriction endonuclease cleavage site mapping of the OV NZ2 *Bam*HI E fragment, which is at the left end of the genome, and the equivalent fragment in OV NZ7, *Bam*HI F (34), has shown that they are closely related if not identical in sequence (data not shown). A comparison of the OV NZ2 *Bam*HI E sequence with the 1,830-nt sequence reported here for OV NZ7 showed that there is almost 100% homology between the two sequences from the *Sph*I site at the right end to a point 148 nt past the OV NZ2 ITR junction. The point of divergence between the two sequences is at the junction of the A+T- and G+C-rich sequences, and we have designated this point as the ITR junction in OV NZ7.

Within the region of OV NZ7 that shows no homology to OV NZ2 is an ORF potentially encoding a polypeptide of 148 aa with a predicted molecular weight of 16,054. An early transcriptional termination sequence, 5′-TTTTTGT, is found 20 nt downstream from the stop codon of this ORF, while upstream is a poxviral early promoter-like sequence, 5′-AAAATGTAAATACTAA (Fig. 2). The sequence from the stop codon to the ITR junction appears to be noncoding, as does the first 148 nt within the ITR. The final 122 nt is homologous to the OV NZ2 ITR and contains an early promoter and initiation codon sequence likely to be the promoter region for a gene equivalent to OV NZ2 ORF 3.

For convenience, we named the three ORFs found in each strain as A1R, A2R, and A3R, a system similar to that used for naming ORFs in VAC except that we have chosen the *Bam*HI rather than the *Hin*dIII cleavage site map as the basis for the nomenclature. They are the first three ORFs numbered from right to left in the *Bam*HI A fragment of the OV genome. The letter R indicates that they would be transcribed from left to right. A1R is the gene equivalent to the gene found in the ITR at the left end which had been designated ORF 3 (13), A2R is the ORF found in the region of nonhomology between OV NZ2 and OV NZ7, and A3R is the ORF equivalent to the 3′ end of VAC F9L. Northern (RNA) blot analyses revealed that RNA of the appropriate size and sense is actively transcribed early in the infected-cell cycle from the A2R of both the OV NZ2 and OV NZ7 genomes (Fig. 3).

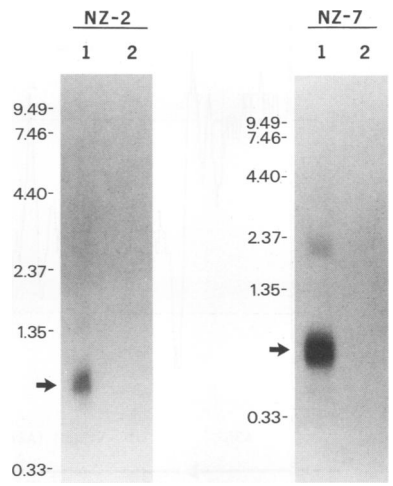


FIG. 3. RNA transcripts from the viral VEGF-like genes detected by Northern blot analysis of total RNA isolated from cells infected with OV NZ2 and NZ7. Details of the methods used have been published elsewhere (12). Briefly, total RNA from bovine testis cells incubated in the presence of cycloheximide and infected with OV at 30 PFU per cell was extracted 6 h postinfection, separated on a 1% agarose gel containing formaldehyde, transferred to a nitrocellulose membrane, and hybridized with single-stranded DNA probes that were either complementary (lanes 1) or noncomplementary (lanes 2) to the predicted viral VEGF transcripts. The positions of RNA size markers in kilobases are shown to the left of each gel.

The low G+C content of the nonhomologous regions of the two strains is shown graphically in Fig. 4, in which the G+C content of the DNA is plotted against distance along the genome. It can be seen that they appear as islands of lower G+C content flanked by conserved G+C-rich sequences. OV NZ2 A2R has a G+C content of 57.2%, while A2R of OV NZ7 is 39.7% G+C. Also shown in Fig. 4 are the locations of ORFs A1R and A3R and of the ITRs.

The polypeptides predicted from the A2R ORFs of OV NZ2 and NZ7 were compared with sequences held in the translated GenBank nucleic acid data base, using FASTA (31). It was found that both polypeptides showed homology to members of the VEGF family (4, 5, 8, 9, 16, 20, 22, 23, 46) and the B chain of the related family of platelet-derived growth factors (PDGFs) (7, 38). When the NZ2 polypeptide (OV-VEGF2) was compared against VEGF, optimized scores ranged from 249 to 277; when it was compared against the PDGF B chain, optimized scores ranged from 128 to 133. The scores obtained when the NZ7 polypeptide (OV-VEGF7) was used were

FIG. 2. Alignment of nucleotide sequences found in a region at the right end of the genomes of OV NZ2 and NZ7. The location of the region sequenced in both strains is shown in Fig. 1. The order of amino acids, represented in the one-letter code, was deduced from the nucleotide sequences. The initial alignment of the nucleotide sequence was done by using CLUSTAL (19), which aligned the homologous sequences at each end. Alignment of the central region was done on the basis of an optimal alignment of the amino acid sequences, using CLUSTAL, predicted from the A2R ORFs (see text). Nucleotide homologies between the two sequences are indicated by asterisks below the aligned sequences. Sequences critical for the initiation of early transcription (6, 12) are indicated by a line above the relevant sequence, and early transcription termination signal sequences (12, 50) are indicated by dots.

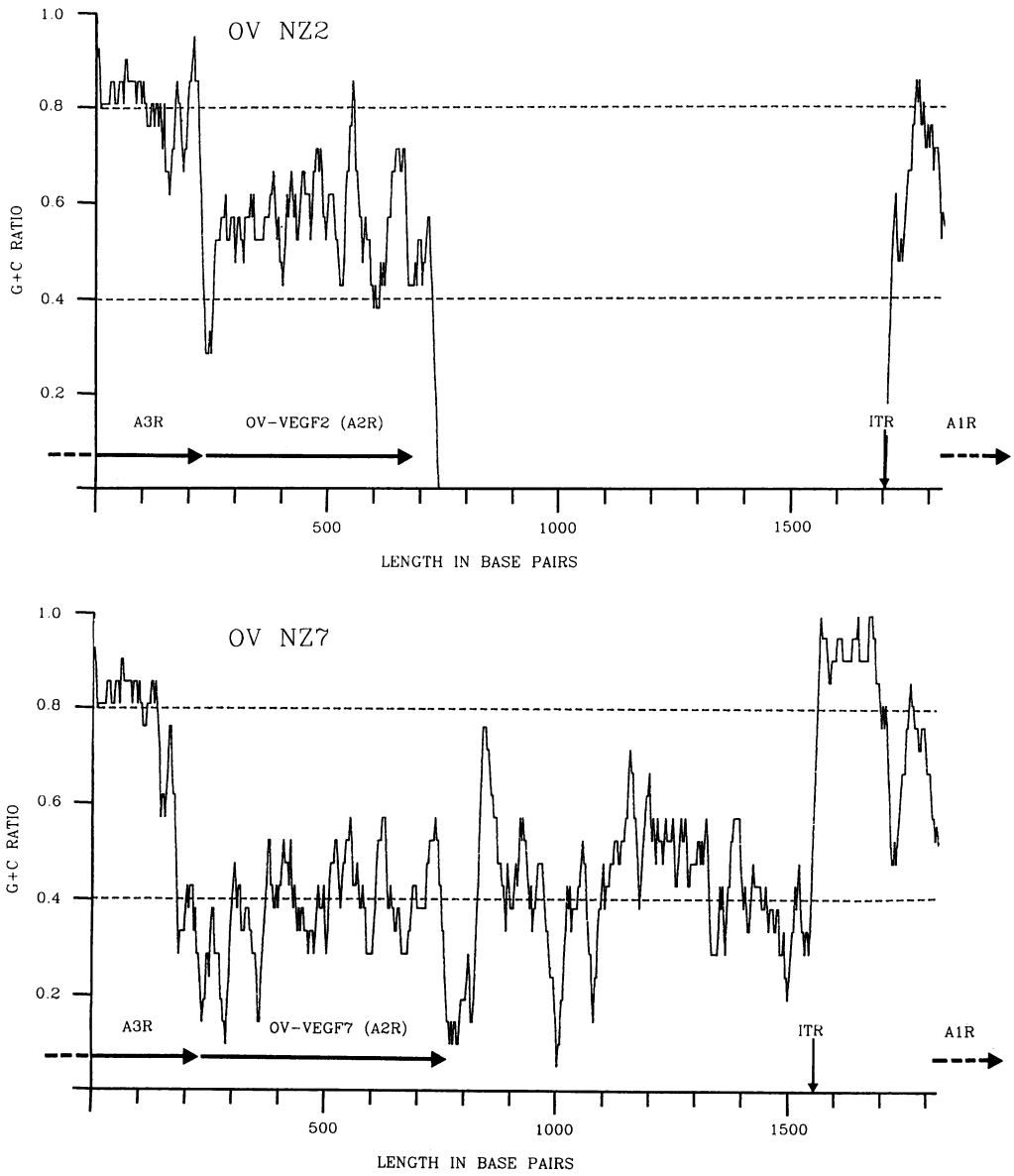


FIG. 4. Comparison of the G+C content of the OV NZ2 and OV NZ7 genomes in the regions containing the VEGF-like genes. The G+C content of each sequence was plotted by using the program COMPLOT (43) (window length of 21 nt, plotting every 3 nt). The positions and orientations of ORFs and the ITR junctions are indicated. The OV NZ2 sequence was plotted with one gap of 979 nt placed between the last nucleotide of unique sequence and the ITR.

somewhat lower: 172 to 184 against VEGF and 117 to 124 against PDGF. Bovine VEGF gave the highest score with both viral polypeptides; the optimized score with OV-VEGF2 was 277, with 43.7% identity over 103 aa, and that with OV-VEGF7 was 184, with 23.9% identity over 134 aa. When OV-VEGF2 and OV-VEGF7 were compared by using FASTA, the optimal score was 257, with 42.1% identity over 140 aa.

VEGFs have been identified in the genomes of cattle (8, 16, 22), humans (20, 22, 23), rats (4, 5), mice (32), and guinea pigs (2a). In a model for the structure of the human VEGF gene (9), it is proposed that there are at least four introns and that alternative splicing pathways produce four mature peptides, $VEGF_{121}$, $VEGF_{165}$, $VEGF_{189}$, and $VEGF_{206}$. A comparison of these alternative forms of VEGF with the viral polypeptides reveals that both are most closely related to $VEGF_{121}$. In common with all other poxviral genes examined (28), the two viral VEGF-like genes do not appear to contain introns.

Pairwise identity and similarity scores between the OV sequences and the VEGFs were calculated by using the program HOMED (44) following optimal alignment with CLUSTAL (19). The identity and similarity scores are shown in Table 1. Although the scores give some measure of the relatedness between the various VEGFs and their splicing isoforms, the presence or absence of exons or signal sequences in the mammalian genes complicates the interpretation of the relatedness. However, it does show that the viral polypeptides are as different from one another as the human $VEGF_{121}$ is

TABLE 1. VEGF amino acid sequence comparisons[a]

	NZ7 VEGF	Bovine $VEGF_{120}$	Bovine $VEGF_{164}$	Human $VEGF_{121}$	Human $VEGF_{165}$	Human $VEGF_{189}$	Guinea pig VEGF	Rat VEGF	Human placental VEGF
NZ2 VEGF	73.3[b] (41.1)[c]	49.0 (26.9)	50.0 (25.6)	56.6 (28.5)	49.7 (25.0)	44.4 (22.3)	43.8 (24.6)	50.3 (24.6)	55.3 (27.5)
NZ7 VEGF		51.3 (19.6)	51.7 (20.5)	60.1 (23.0)	51.5 (19.9)	46.1 (17.8)	43.7 (16.1)	51.7 (19.0)	57.9 (19.6)
Bovine $VEGF_{120}$			62.5 (62.6)	79.7 (76.2)	61.2 (58.6)	54.6 (52.6)	70.6 (65.2)	60.1 (52.6)	64.3 (35.9)
Bovine $VEGF_{164}$				75.4 (72.8)	98.6 (95.8)	87.7 (84.7)	83.8 (79.5)	97.5 (88.4)	60.9 (30.8)
Human $VEGF_{121}$					76.8 (77.0)	68.3 (67.9)	59.5 (53.9)	73.9 (65.4)	76.9 (37.3)
Human $VEGF_{165}$						89.0 (88.4)	82.6 (76.4)	96.9 (87.4)	61.1 (29.7)
Human $VEGF_{189}$							73.5 (67.4)	86.2 (77.2)	54.6 (26.5)
Guinea pig VEGF								83.5 (74.7)	51.7 (30.3)
Rat VEGF									62.2 (33.3)

[a] Scores calculated by using the program HOMED (44) after the CLUSTAL alignment shown in Fig. 5.
[b] Percentage similarities.
[c] Percentage identities.

from the human placental VEGF. The optimal alignment (Fig. 5) provides another means of comparison and shows that the VEGFs and the viral polypeptides show strong homology to each other over a region of 90 aa. When alignments are done with the simian sarcoma virus v-*sis* gene product (not shown), the 90-aa region can be shown to be homologous to an 88-aa region defined in v-*sis* as the minimal transforming region (18, 40).

The mammalian VEGF genes code for a signal sequence, and in the case of the bovine, human, and rat genes, a consensus cleavage recognition site has been identified. In the bovine VEGF, the cleavage recognition site is Ser-Gln-Ala at residues 24 to 26 (22). The amino-terminal end of both the OV-VEGF2 and OV-VEGF7 contains a sequence of hydrophobic residues indicative of a signal sequence. Each such sequence is preceded by a Lys at residue 2, and each contains a putative cleavage site at residues 20 to 22 (Ala-Asp-Ser) in OV-VEGF2 and residues 25 to 27 (Ser-Gln-Ser) in OV-VEGF7 (Fig. 5). The remaining sequences of both viral gene products contain the region homologous to the v-*sis* minimal transforming region, conserved in all members of the VEGF family, and a short carboxy-terminal region containing Arg residues. The sequence equivalent to the minimal transforming region in both retains the eight cysteines conserved in the VEGFs, two of which are essential for dimerization and function of the v-*sis* gene product (40) (residues 112 and 114 in OV-VEGF2 and 130 and 132 in OV-VEGF7) (Fig. 5). Each sequence also includes a potential N-linked glycosylation site, residues 85 to 87 in OV-VEGF2 and residues 95 to 97 in OV-VEGF7, which align with the glycosylation site in the VEGFs (Fig. 5). Lacking in both viral genes are four blocks of sequence corresponding to the proposed central four exons in the human VEGF gene, including the highly basic sequence of 24 residues shown to be responsible for the cell-associated forms of human VEGF, $VEGF_{189}$ and $VEGF_{206}$ (9) (residues 142 to 165 in human $VEGF_{189}$; Fig. 5).

Significant homology between OV-VEGF2 and the mammalian VEGFs can also be detected at the DNA level. FASTA gives an optimally aligned score of 221 to 283 over regions of 308 to 321 nt. This is not the case with OV-VEGF7, where there is considerable divergence at the DNA level from the mammalian VEGFs and the OV-VEGF2 sequence, nor is there detectable hybridization between the two viral genes under conditions of moderate stringency. To eliminate the possibility that each strain contains a copy of both gene variants but in different genomic locations, we hybridized the OV NZ2 gene to the total genome of OV NZ7 and vice versa. Neither gene showed cross-hybridization (not shown).

A CLUSTAL alignment of the polypeptides encoded by the two viral genes was used to align the DNA sequences in the nonhomologous regions (Fig. 2). The alignment suggests that compared with OV-VEGF7, two deletions, each of 24 nt, have occurred in that portion of the OV-VEGF2 gene that encodes hydrophilic regions in the OV NZ7 polypeptide and that a 6-nt deletion has occurred in the signal sequence. There is also a 3-nt deletion at the carboxy-terminal end of the OV-VEGF7 gene compared with OV-VEGF2. The placement of the deletions in the noncoding regions between the stop codon of the F9L gene and the initiation codons of the VEGF-like genes were made arbitrarily, using the principle of maximum parsimony. The junction between the ITR and unique sequence was chosen as the point at which to place the 889-nt gap in the OV NZ2 sequence.

DISCUSSION

We have compared the DNA sequence adjacent to the ITR junctions in two strains of OV, NZ2 and NZ7. Immediately adjacent to the ITR of OV NZ7 is an 889-nt noncoding region, not found in OV NZ2, which is AT rich and which can account for the major part of the size difference detected between the *Kpn*I E fragments of the two strains. Internal to this region in OV NZ7 is a gene potentially encoding a 16-kDa protein with homology to the mammalian VEGFs and with a G+C content of 39.7%. In the OV NZ2 genome, adjacent to the ITR, we found a gene with a G+C content of 57.2% potentially encoding a 14.7-kDa protein also with homology to mammalian VEGFs. Both viral genes are strongly expressed early in infection.

Adjacent and internal to the VEGF-like genes in both

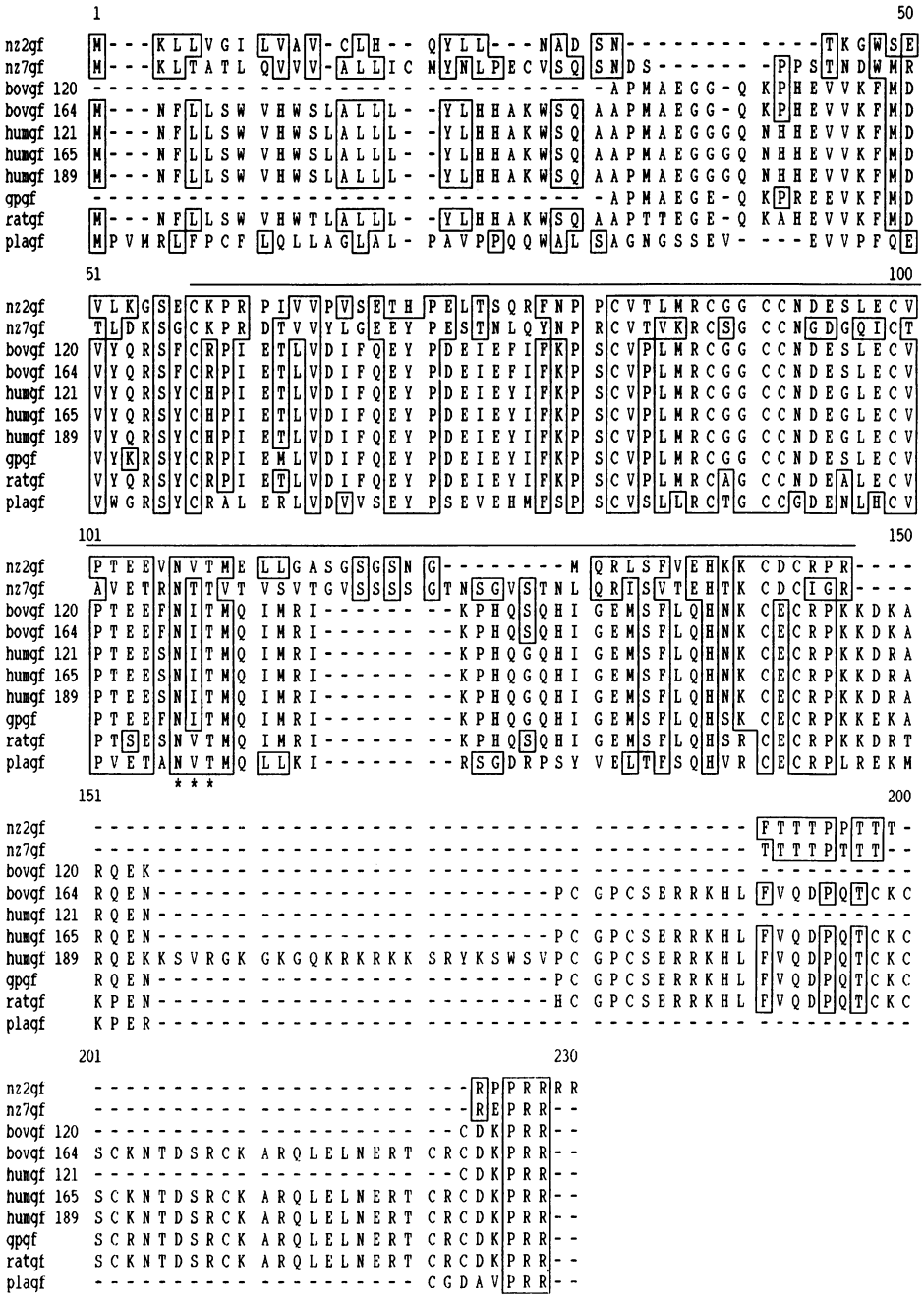


FIG. 5. Alignment of the inferred amino acid sequences of the NZ2 and NZ7 VEGF-like genes with sequences of the mammalian VEGFs, using CLUSTAL (19). The mammalian VEGF polypeptides compared are bovine $VEGF_{120}$ (bovgf 120 [46]), bovine $VEGF_{164}$ (bovgf 164 [22, 46]), human $VEGF_{121}$ (humgf 121 [22]), human $VEGF_{165}$ (humgf 165 [22]), human $VEGF_{189}$ (humgf 189 [20]), guinea pig VEGF (gpgf [2a]), rat VEGF (ratgf [4]), and human placental VEGF (plagf [23]). The amino acid residues conserved between the two viral proteins and between one or both of the viral proteins and the mammalian VEGFs are boxed. The bar delineates the region equivalent to the minimal transforming region of v-*sis* (40); the putative glycosylation recognition site is indicated by asterisks.

strains, we identified regions whose predicted polypeptide sequences share homology with the carboxy terminus of the VAC gene F9L. From the alignment of the OV genome with that of VAC, a homolog of F9L would be expected to be at the left rather than the right end, suggesting that some OV genes may be rearranged compared with those in VAC.

It was unexpected to find that the two VEGF-like genes were barely homologous at the DNA level and encoded variant proteins. Poxviruses, and indeed other viruses, encode polypeptides with homologies to cellular polypeptides, and it is reasonable to assume that during evolution, these genes have been acquired from their hosts. The similarity of the OV NZ2 DNA sequence with sequences of the mammalian VEGF genes is consistent with the hypothesis that the viral gene has

been acquired by an event that occurred recently, in evolutionary terms, possibly from the natural host species, sheep. However, the gene in OV NZ7 would then represent a variant which has diverged at an extraordinarily high rate, the change being restricted to this gene, since the flanking DNA shows over 98% identity to that in OV NZ2. Furthermore, it might be expected that in a G+C-rich virus, divergence would have involved a drift toward a higher rather than a lower G+C content. An alternative explanation for the high degree of variation found between the two genes could be that they represent separate acquisitions of genes with different sequences, and their virtually identical locations indicate that the near-terminal regions of poxvirus genomes are favored sites for gene acquisition.

Another unusual feature of the regions sequenced is the G+C content. We have characterized 14 genes in OV and have found 5 with G+C contents lower than might be expected for a genome with a overall G+C content of 63%. These are ORF 1 (55%) (13), a dUTPase-like gene (49%) (26), a gene encoding homolog of the 14-kDa VAC fusion protein (42%) (29), and two ORFs downstream from the latter gene in a region equivalent to that occupied in cowpox virus by the type A inclusion gene (37 and 45%) (30). The remaining genes range in G+C content from 60 to 75%, with an average of 68% (12, 45). We can offer no simple explanation as to why there is such a large variation in the G+C content of genes in OV, taking into account their positions in the genome and their putative functions.

Preliminary studies have revealed that supernatants from cell cultures infected with either strain have mitogenic effects on vascular endothelial cells compared with control cells (10). The lesions induced after infection of sheep (3, 47) and humans (17, 21, 39, 49) with OV are remarkable for the extensive capillary proliferation, dilatation, and dermal swelling seen in histological sections, and it seems likely that this response is a direct effect of the activity of the VEGF-like gene. How the VEGF activity might be an advantage to the virus is unclear, but it could increase the flow of metabolites essential to the virus or indirectly induce epithelial proliferation. A similar pathology is seen in humans after infection with milkers' node virus (17), another member of the parapoxvirus genus, and we predict that the presence of a VEGF-like gene will be characteristic of the parapoxviruses.

ACKNOWLEDGMENT

We are grateful for the excellent technical assistance of Ellena Whelan.

REFERENCES

1. **Ahn, B. Y., and B. Moss.** 1992. RNA polymerase-associated transcription specificity factor encoded by vaccinia virus. Proc. Natl. Acad. Sci. USA **89:**3536–3540.
2. **Balassu, T. C., and A. J. Robinson.** 1987. Orf virus replication in bovine testis cells: kinetics of viral DNA, polypeptide, and infectious virus production and analysis of virion polypeptides. Arch. Virol. **97:**267–281.
2a. **Berse, B.** GenBank accession no. M84230.
3. **Cheville, N. F.** 1975. Cytopathology in viral diseases, p. 9–12. Karger, Basel.
4. **Conn, G., M. L. Bayne, D. D. Soderman, P. W. Kwok, K. A. Sullivan, T. M. Palisi, D. A. Hope, and K. A. Thomas.** 1990. Amino acid and cDNA sequences of a vascular endothelial cell mitogen that is homologous to platelet-derived growth factor. Proc. Natl. Acad. Sci. USA **87:**2628–2632.
5. **Conn, G., D. D. Soderman, M.-T. Schaeffer, M. Wile, V. B. Hatcher, and K. A. Thomas.** 1990. Purification of a glycoprotein vascular endothelial cell mitogen from a rat glioma-derived cell line. Proc. Natl. Acad. Sci. USA **87:**1323–1327.
6. **Davison, A. J., and B. Moss.** 1989. Structure of vaccinia virus early promoters. J. Mol. Biol. **210:**749–769.
7. **Deuel, T. F.** 1987. Polypeptide growth factors: roles in normal and abnormal cell growth. Annu. Rev. Cell Biol. **3:**443–492.
8. **Ferrara, N., and W. J. Henzel.** 1989. Pituitary follicular cells secrete a novel heparin-binding growth factor specific for vascular endothelial cells. Biochem. Biophys. Res. Commun. **161:**851–858.
9. **Ferrara, N., K. A. Houck, L. B. Jakeman, J. Winer, and D. W. Leung.** 1991. The vascular endothelial growth factor family of polypeptides. J. Cell. Biochem. **47:**211–218.
10. **Ferrara, N., D. J. Lyttle, and A. J. Robinson.** Unpublished data.
11. **Fleming, S. B., J. Blok, K. M. Fraser, A. A. Mercer, and A. J. Robinson.** 1993. Conservation of gene structure and arrangement between vaccinia virus and orf virus. Virology **195:**175–184.
12. **Fleming, S. B., K. M. Fraser, A. A. Mercer, and A. J. Robinson.** 1991. Vaccinia virus-like early transcriptional control sequences flank an early gene in orf virus. Gene **97:**207–212.
13. **Fraser, K. M., D. F. Hill, A. A. Mercer, and A. J. Robinson.** 1990. Sequence analysis of the inverted terminal repetition in the genome of the parapoxvirus, orf virus. Virology **176:**379–389.
14. **Goebel, S. J., G. P. Johnson, M. E. Perkus, S. W. Davis, J. P. Winslow, and E. Paoletti.** 1990. The complete DNA sequence of vaccinia virus. Virology **179:**247–266.
15. **Gordon, J., A. Mohandas, S. Wilton, and S. Dales.** 1991. A prominent antigenic surface polypeptide involved in the biogenesis and function of the vaccinia virus envelope. Virology **181:**671–686.
16. **Gospodarowicz, D., J. A. Abraham, and J. Schilling.** 1989. Isolation and characterization of a vascular endothelial cell mitogen produced by pituitary-derived folliculo stellate cells. Proc. Natl. Acad. Sci. USA **86:**7311–7315.
17. **Groves, R. W., E. Wilson-Jones, and D. M. MacDonald.** 1991. Human orf and milkers' nodule: a clinicopathological study. J. Am. Acad. Dermatol. **25:**706–711.
18. **Hannink, M., M. K. Sauer, and D. J. Donoghue.** 1986. Deletions in the C-terminal coding region of the v-*sis* gene: dimerization is required for transformation. Mol. Cell. Biol. **6:**1304–1314.
19. **Higgins, D. G., and P. M. Sharp.** 1988. CLUSTAL: a package for performing multiple sequence alignments on a microcomputer. Gene **73:**237–244.
20. **Keck, P. J., S. D. Hauser, G. Krivi, K. Sanzo, T. Warren, J. Feder, and D. T. Connolly.** 1989. Vascular permeability factor, an endothelial cell mitogen related to PDGF. Science **246:**1309–1312.
21. **Leavell, U. W., M. J. McNamara, R. J. Muelling, and F. Landrum.** 1965. Ecthyma contagiosum (orf). South. Med. J. **58:**239–243.
22. **Leung, D. W., G. Cachianes, W.-J. Kuang, D. V. Goeddell, and N. Ferrara.** 1989. Vascular endothelial growth factor is a secreted angiogenic mitogen. Science **246:**1306–1309.
23. **Maglione, D., V. Guerriero, G. Viglietto, and P. Dellibovi.** 1991. Isolation of a human placental cDNA coding for a protein related to the vascular permeability factor. Proc. Natl. Acad. Sci. USA **88:**9267–9271.
24. **McGeoch, D. J.** 1990. Protein sequence comparisons show that the 'pseudoproteases' encoded by poxviruses and certain retroviruses belong to the deoxyuridine triphosphatase family. Nucleic Acids Res. **18:**4105–4110.
25. **Mercer, A. A., K. Fraser, G. Barns, and A. J. Robinson.** 1987. The structure and cloning of orf virus DNA. Virology **157:**1–12.
26. **Mercer, A. A., K. M. Fraser, P. A. Stockwell, and A. J. Robinson.** 1989. A homologue of retroviral pseudoproteases in the parapoxvirus, orf virus. Virology **172:**665–668.
27. **Mercer, A. A., G. Green, D. Spehener, R. Drillien, and A. J. Robinson.** Unpublished data.
28. **Moss, B.** 1990. Poxviridae and their replication, p. 2079–2111. *In* B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), Virology. Raven Press, New York.
29. **Naase, M., B. H. Nicholson, K. M. Fraser, A. A. Mercer, and A. J. Robinson.** 1991. An orf virus sequence showing homology to the fusion protein gene of vaccinia virus. J. Gen. Virol. **72:**1177–1181.
30. **Naase, M., B. H. Nicholson, D. J. Lyttle, and A. J. Robinson.** Unpublished data.

31. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
32. **Plouet, J., J. Schilling, and D. Gospodarowicz.** 1989. Isolation and characterization of a newly identified endothelial cell mitogen produced by AtT20. EMBO J. **8:**3801–3806.
33. **Robinson, A. J., and T. C. Balassu.** 1981. Contagious pustular dermatitis (orf). Vet. Bull. **51:**771–781.
34. **Robinson, A. J., G. Barns, K. Fraser, E. Carpenter, and A. A. Mercer.** 1987. Conservation and variation in orf virus genomes. Virology **157:**13–23.
35. **Robinson, A. J., G. Ellis, and T. Balassu.** 1982. The genome of orf virus: restriction endonuclease analysis of viral DNA isolated from lesions of orf in sheep. Arch. Virol. **71:**43–55.
36. **Robinson, A. J., and D. J. Lyttle.** 1992. Parapoxviruses: their biology and potential as recombinant vaccines, p. 285–327. *In* M. M. Binns and G. L. Smith (ed.), Recombinant poxviruses. CRC Press, Boca Raton, Fla.
37. **Rodriguez, J. F., and M. Esteban.** 1987. Mapping and nucleotide sequence of the vaccinia virus gene that encodes a 14-kilodalton fusion protein. J. Virol. **61:**3550–3554.
38. **Ross, R., E. W. Raines, and D. F. Bowen-Pope.** 1986. The biology of platelet-derived growth factor. Cell **46:**155–169.
39. **Sanchez, R. L., A. Herbert, H. Lucia, and J. Swedo.** 1985. Orf. A case report with histologic, electron microscopic and immunoperoxidase studies. Arch. Pathol. Lab. Med. **109:**166–170.
40. **Sauer, M. K., M. Hannink, and D. J. Donoghue.** 1986. Deletions in the N-terminal coding region of the v-*sis* gene: determination of the minimal transforming region. J. Virol. **59:**292–300.
41. **Shuman, S., and B. Moss.** 1987. Identification of a vaccinia virus gene encoding a type I DNA topoisomerase. Proc. Natl. Acad. Sci. USA **84:**7478–7482.
42. **Slabaugh, M. B., and N. A. Roseman.** 1989. Retroviral protease-like gene in the vaccinia virus genome. Proc. Natl. Acad. Sci. USA **86:**4152–4155.
43. **Stockwell, P. A. (University of Otago).** Unpublished data.
44. **Stockwell, P. A., and G. B. Petersen.** 1987. HOMED: a homologous sequence editor. Comput. Appl. Biosci. **3:**37–43.
45. **Sullivan, J., A. A. Mercer, S. B. Fleming, and A. J. Robinson.** Unpublished data.
46. **Tischer, E., D. Gospodarowicz, R. Mitchell, M. Silva, J. Schilling, K. Lau, T. Crisp, J. C. Fiddes, and J. A. Abraham.** 1989. Vascular endothelial growth factor: a new member of the platelet-derived growth factor gene family. Biochem. Biophys. Res. Commun. **165:**1198–1206.
47. **Wheeler, C. E., and E. P. Cawley.** 1956. The microscopic appearance of ecthyma contagiosum (orf) in sheep, rabbits, and man. Am. J. Pathol. **32:**535–545.
48. **Wittek, R., C. C. Kuenzle, and R. Wyler.** 1979. High C+G content in parapoxvirus DNA. J. Gen. Virol. **43:**231–234.
49. **Yeh, H.-P., and K. Soltani.** 1974. Ultrastructural studies in human orf. Arch. Dermatol. **109:**390–392.
50. **Yuen, L., and B. Moss.** 1986. Multiple 3′ ends of mRNA encoding vaccinia virus growth factor occur within a series of repeated sequences downstream of T clusters. J. Virol. **60:**320–323.