Volume 16 Number 24 1988 Nucleic Acids Research

Evidence for genetic relationship between RNA and DNA viruses from the sequence homology of a putative polymerase gene of Bluetongue virus with that of vaccinia virus: conservation of RNA polymerase genes from diverse species

P.Roy[1,2]*, A Fukusho[1], G.D.Ritter[1] and D.Lyon[1]

[1]University of Alabama at Birmingham, School of Public Health, Department of Environmental Health Sciences, Birmingham, AL 35294, USA and [2]NERC Institute of Virology, Mansfield Road, Oxford, OX1 3SR, UK

Received August 31, 1988; Revised and Accepted November 21, 1988 Accession no. X12819

Abstract

The nucleotide sequence of segment 1 of the double stranded RNA genome of bluetongue virus serotype 10 (BTV-10), encoding the largest viral core protein, VP1, has been determined. Linear sequence analysis of the predicted amino acid sequence of the 149-K Da protein, a putative component of the viral RNA-directed RNA polymerase, revealed extensive homology with the vaccinia virus 147K Da DNA-directed RNA polymerase subunit. Similar homologies were detected between the VP1 polypeptide and the β chain subunit of *Escherichia coli* and common tobacco chloroplast RNA polymerases, yeast RNA polymerase II and III and fruit fly polymerase II.

INTRODUCTION

BTV is a member of the *Orbivirus* genus in the family Reoviridae. The genome of BTV, like reovirus, consists of ten double-stranded RNA species (dsRNA), each encoding at least one polypeptide. The dsRNA genome is encapsidated by an inner protein shell consisting of two major proteins, VP3 and VP7, as well as the minor proteins VP1, VP4 and VP6. The VP1 protein is coded by viral RNA L1 (1). The core particle is surrounded by a diffuse outer capsid, consisting of two proteins, VP2 and VP5. The purified BTV particle contains an RNA-directed RNA polymerase (transcriptase). Although, as in reoviruses, it has been shown that the transcriptase activity resides in the inner viral core, the proteins that constitute the viral transcriptase have not been identified. For reovirus it has been shown that the protein λ3, encoded by the reovirus L1 gene, determines the pH optimum of the viral transcriptase (2).

For BTV, VP1 protein is a candidate component of the virion transcriptase on the basis of its location, size and molar ratio in the virus particle. In these aspects it is comparable to transcriptase components of reoviruses. Determination of the nature of RNA polymerase is central both to an understanding of the viral infection process and the biology and evolution of these viruses. We report here the nucleotide sequence of the L1 gene of BTV-10 and its predicted gene product, VP1. Similarities between the amino acid sequences of the VP1 protein and that of the large RNA

polymerase subunit of vaccinia virus (a DNA virus) as well as those of *Escherichia coli*, common tobacco chloroplast, yeast and fruit fly RNA polymerase are presented.

MATERIALS AND METHODS

Growth of virus and isolation of dsRNA

United States serotype BTV-10 came from the collection of the Arthropod-borne Animal Disease Research Laboratory (Denver, Colorado) after being passaged six times in BHK-21 cells. The virus was plaque-cloned using monolayers of BHK-21 cells. The viral dsRNA was purified as described by Yamaguchi *et al.* (3) and the 10 individual segments were separated by agarose gel electrophoresis and the L1 gene was electroeluted as described previously by Purdy *et al.* (4).

DNA cloning of ds L1 RNA.

Polyadenylation of dsRNA and synthesis of cDNA copies of polyadenylated L1 RNA using oligo(dT)$_{12-18}$ primer were undertaken as described previously (4). The RNA templates were removed by 0.5 *M* KOH treatment and double-stranded cDNA was generated by self-annealing. The products were repaired by the Klenow fragment of DNA polymerase as described previously (4). The double-stranded cDNA hybrids were then tailed with dC at their 3' termini and annealed to a *Pst*I-cut, dG-tailed, pBR322 vector (5). After transformation, clones containing viral sequences were recovered and screened by colony hybridization (6) and the *Hin*fI restriction patterns of the derived hybridization positive inserts were compared as described previously (4).

RNA gel electrophoresis, blotting and hybridization.

Purified BTV-10 RNA was resolved on an agarose gel, blotted onto a Genescreen hybridization transfer membrane (New England Nuclear, Boston, MA) and hybridized to nick-translated cloned DNA by procedures described previously (4).

Sequencing of BTV DNA clones and protein analysis.

Sequencing was carried out on end-labeled, strand-separated restriction fragments of plasmid DNA containing viral DNA inserts using the Maxam and Gilbert method (7). Protein homology searches were made with the FASTP and Staden alignment programme (8, 9).

RESULTS

Cloning and nucleotide sequence of the BTV-10 L1 gene encoding the VP1 protein:

The cloning of the BTV-10 L1 gene was carried out as described under Materials and Methods. The *Hin*fI restriction patterns of putative recombinant plasmids showed that at least ten of 50 clones that were tested possessed similar DNA inserts that were between 1000 and 2500 base pairs long. In order to confirm that they

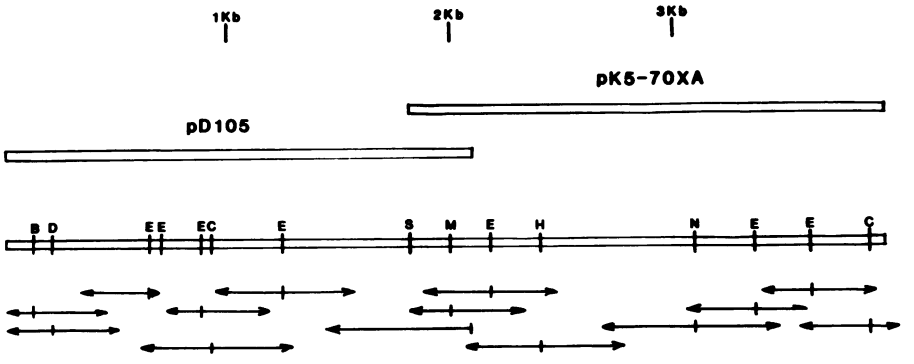

Fig.1. The strategy used to determine the sequence of the cloned BTV-10 segment 1 gene. The restriction enzymes used are shown. The distances and directions in which individual strands were sequenced by the Maxam and Gilbert method (7) are shown by the solid arrows and the dotted arrows. Restriction enzymes used are as follows: B, *BelI;* D, *DdeI*; E, *ECoRI*; M, *MluI*; H, *HindIII*; N, *NsiI*; S, *SphI*.

represented the L1 segment, , BTV-10 RNA was resolved by agarose gel electrophoresis, blotted on Genescreen paper and hybridized to nick translated DNA representing each of the ten positive clones. All the clones annealed specifically to the L1 RNA of BTV-10 (data not shown). Two clones, one containing 5' terminal sequence (#pD105) and another containing 3' terminal sequence (pK5 70XA), were eventually selected for determining the complete sequence of the L1 gene. The sequence was obtained on strand-separated, end labelled restriction fragments from these two overlapping clones (Fig. 1). The majority of the sequence (over 90%) was obtained from analysis of both DNA strands using the Maxam and Gilbert method (7).

The complete nucleotide sequence of the DNA clone of L1 gene is presented in Fig. 2 with the predicted amino acid sequence of the single extensive open reading frame. Excluding the homopolymeric tails, the entire sequence is 3954 nucleotides long. The 5' non-coding region of the L1 RNA of BTV-10 contained only 11 nucleotides, shorter than all the other RNA species of BTV-10 sequenced so far (10 - 16). The 3' non coding region, excluding the stop codon, is 34 nucleotides long, equivalent to those of the other RNA species of BTV-10. The conserved 5' (GTTAAAA....) and 3' (CATTCACA...) terminal sequences of the BTV RNA segments (17) were identified in the DNA clones. Based on the data, the double-stranded L1 RNA is estimated to have a mass of 2.7 x 10^6 Daltons in agreement with previous reports (18).

Predicted VP1 protein analysis:

A single open reading frame in the L1 gene codes for a primary gene product

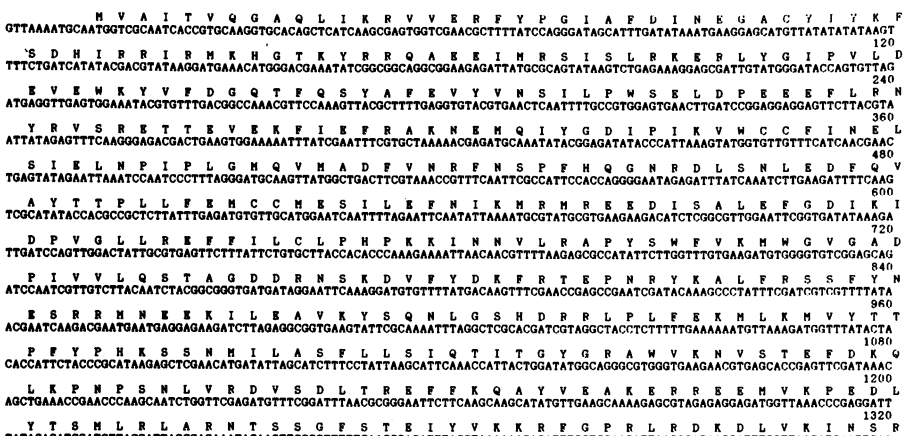

Fig. 2. The nucleotide sequence of the cloned L1 gene of BTV-10. The open reading frame begins at nucleotide residues 12-14 and terminates by aTGA codon at residuea 3918-3920. The single-letter amino acid codes are shown above their respective DNA codon (see Table 1).

Table 1 Amino acid compositions of BTV-10 VP1 protein

Alanine (A)	76
Arginine (R)	91
Aspartic (D)	77
Asparagine (N)	53
Cysteine (C)	12
Glutamic (E)	84
Glutamine (Q)	42
Glycine (G)	69
Histidine (H)	21
Isoleucine (I)	95
Leucine (L)	111
Lysine (K)	87
Methionine (M)	50
Phenylalanine (F)	65
Proline (P)	51
Serine (S)	83
Threonine (T)	78
Tryptophan (W)	11
Tyrosine (Y)	56
Valine (V)	90
Total	1,302
Net charge	+27.5
Size	149,588

composed of 1302 amino acids with a calculated size of 149,588 Daltons, slightly larger than that of previous estimates (140,000 Daltons, 19).

The protein is high in leucine and isoleucine and low in alanine compared to the "average protein" (Table 1, 20). Assuming that glutamic and aspartic acids each have charges of -1, arginine and lysine + 1, and histidine + 0.5, the calculated net charge of the predicted molecule was + 27.5 at physiological pH. In the middle region of the molecule there are at least four tracts of 18 to 25 hydrophobic amino acids without any intervening charged residues (e.g., aa residues 364-383, 521-541, 725-743 and 912-937). The distribution of the charged amino acids throughout the molecule is even except near the carboxyl terminus which contains many arginine and lysine residues (Fig. 2). It remains to be determined whether these are involved in interactions between VP1 protein and genomic RNA or has other functions. A number of hydrophobic regions of the molecule were identified in hydropathic analysis (Fig. 3), dispersed throughout the molecule.

Evidence for homology between bluetongue virus VP1 protein and vaccinia virus RNA polymerase:

The derived amino acid sequence of VP1 protein of BTV-10 was compared with PIR protein library using the FASTP computer program of Lipman and Pearson (9). The search revealed homology with the vaccinia virus 147-K Da polymerase subunit (21). The similarity was considered significant since vaccinia virus is a DNA virus and the

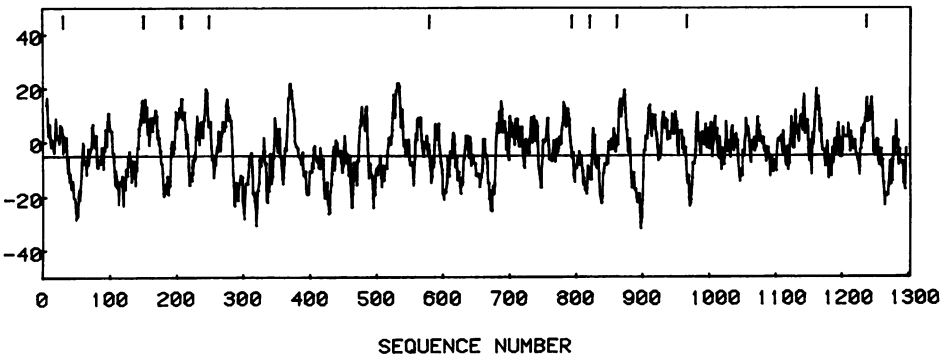

Fig. 3. Hydropathic plot and distribution of cysteine residues for the predicted L1 gene product of BTV-10. The regions of the predicted proteins with a net hydrophobicity (areas above the centre line) or hydrophilicity (areas below the centre line) as well as the distribution of cysteine residues (vertical bars) are displayed (23). The plot involves a span setting of 21 amino acids.

```
Vacc  MAVISKVTYSLYDQKEINATDIIISHVKN-DDDIG-TVKDGRLGAMDGALCKTCGKTELE     60
BTV   MVAIT-----------VQGAQLIKRVVERFYPGIAFDINEG-----------------A
      M**I*           ***  *I   V      I   * *G

Vacc  CFGHWGKVSIYK--THIVK------PEFISEIIRLLNHICIHCGLLRSREPYS----DDI    120
BTV   CY-------IYKFSDHIRRIRMKHGTKYRRQAEEIMRSIS-----LRKERLYGIPVLDEV
      C*       IYK   HI *         *   *  **  I      LR    Y    D**

Vacc  NLK-ELSGHALRRLK-----DKILSKKKSCWNS-----ECMQPY---------QK-ITFS    180
BTV   EWKYVFDGQTFQSYAFEVYVNSILP-----WSELDPEEEFLRNYRVSRETTEVEKFIEFR
        K  * G  *  *         IL      W       E**  Y           K I F

Vacc  KK----------KV--CFVNKLDDINV-PNSLIYQKLISIHEKF-WPLLEIHQYPANL-    240
BTV   AKNEMQIYGDIPIKVWCCFINELS-IELNPIPLGMQVMADFVNRFNSPFHQGNRDLSNLE
       K           KV  CF*N L  I * P  L *Q ** *  *F  P*         NL

Vacc  -FYTDYFPIPPLI-----IRPAISFWIDSIPKETN--ELTY---------LLGMIVKNCN    300
BTV   DFQVAY--TTPLLFEMCCMESILEFNIKMRMREEDISALEFGDIKIDPVGLLREFFILC-
       F   Y    PL*     *  ** F I    *E    L *          LL  **  C

Vacc  LNADEQVIQKAVIEY----------DDIKII----SNNTSSINLSY----ITSGK-NNM    360
BTV   LPHPKKINNVLRAPYSWFVKMWGVGADPIVVLQSTAGDDRNSKDVFYDKFRTEPNRYKAL
      L     * * * * Y          D I **          S  * Y          *   *

Vacc  IRS--YIVARRKDQTARSVIGP---STSITVNEVGMPAYIRNTLTEKIFVNAFTVDKVKQ    420
BTV   FRSSFYNESRRMNEE--KILEAVKYSQNLGSHDRRLPLFEK--MLKMVYTTPFYPHKSSN
      *RS  Y   RR       **     S  *   *  *p** *  *   **   F   K  *

Vacc  ------LLASNQVKFY----------FNKRLN----------QLTR--IRQGKFIKNK    480
BTV   MILASFLLSIQTITGYGRAWVKNVSTEFDKQLKPNPSNLVRDVSDLTREFFKQA-YVEAK
           LL  * *  Y           F K L           LTR  **Q  **  K

Vacc  IHLLPGDWVEVAVQE--YTSII-FGRQPS----LHRY--NVIASSIRATEGDTIKISPGI    540
BTV   ER-----REEMVKPEDLYTSMLRLARNTSSGFSTEIYVKKRFGPRLR--DKDLVKINSRI
              E**  E  YTS** * R* S       Y    **  *R  * D *KI   I

Vacc  AN----SQ-NADFDGDEEWMILEQNPKAVIEQSILMYPTTLLKH-DIHGAP-VYGSIQDE    600
BTV   KALVIFTKGHTVFTDEE----LHKK-----YNSVELYQTKGSRDVPIKATRTIY-SINLS
            *     F* *E    L          *S* *Y T * *   I *   *Y SI*

Vacc  IVAAYSLFRIQDLCLDEVLNILGKYGR-EFDPKGKCKFSGKDIYTYL--IGEKI-NYPGL    660
BTV   VLVP-QL--IVTLPLNEYFSRVGGITRPDYK-----KIGGKVIVGDLEATGSRVMDAADC
      ***   L  I  L L E**  *G * R **       K* GK I*  L   G **  *  *

Vacc  LKDG---EIIANDVD-SNFVVAM-RHLSLAGLLSDHKSNVE---GINFIIKSSYVFKRYL    720
BTV   FRNSADRDIFTIAIDYSEYDTHLTRHNFRTGMLQGIREAMAPYRALRY---EGYTLEQII
      **      *I*   *D S *   * RH    G*L   *  *      * *      Y *  **

Vacc  SIYGFG---VTFKDLRPNST--FTNKLEA-INVEKIELIKEAYAKYLNDVRDGKIVPLSK    780
BTV   D-FGYGEGRVA-NTLWNGKRRLFKTTFDAYIRLDESERDKGSF-----KVPKG-VLPVSS
        *G*G   V    L        F   **A I **  E  K  *         V  G **P*S

Vacc  -------ALEADYVESMLSNLTNLNIREIEEHMRQ---TLIDDPDNNL----LKMAKAKA    840
BTV   VDVANRIAVDKGF-DTLIAATDGSDLALIDTHLSGENSTLIANSMHNMAIGTL-IQRQRA
             A**  * ****         *  I* H*      TLI     N*    L * * *A
```

```
Vacc  GYKVNPTELMYILGTY-GQQRI---DGEPAETRVLGRVLPYYLPDSKDPEGRGYILNSLT   900
BTV   VGREQPGILTFLSEQYVGDDTLFYTKLHTTDITVFDKVAASIF-DTVAKCGHE---ASPS
      * * *P  L **   Y G   *         *  V* *V*  ** D*   G      S *

Vacc  KGLTGSQYYFSMLVARSQSTDIVC--ETSRTGTLARKIIKKMEDMVVDGYGQVVIGNTLI   960
BTV   KTMMTP---YSVEKTQTHAKQ-GCYVPQDRMMIISSERRKDIED--VQGYVRSQV-QTMI
      K *      *S*    *       C     R   *     K *ED  V GY    * *T*I

Vacc  KYAANYTKILGSVCKPV-DLIYPDESMTWYLEISALWNK---IKQGFVYSQKQKLAKKTL  1120
BTV   ------TKVSRGFCHDLAQLILMLKTT--F--IGA-WKMKRTIKENAMYRDR-KFDSNDE
            TK*   *C  *  LI*   *   *  I A WK    IK  **Y  * K*

Vacc  APFNFLVFVKPTTEDNAIKVKDLYDMIHNVIDDVR--EKYF--FTVSNID-FMEYI--FL  1180
BTV   DGFTLIQIRNPLALYVPIGWNG-YGAHPAALNIVMTEEMYVDSIMISKLDEIMAPIRRIV
        F ** *  P       I     Y *   **  V   E Y*  *  *S *D *M  I  **

Vacc  THLNPSR--IRITKETAITIFE-KFYEKL-----NYTLGGGTPIGIISAQVLSEKFTQQA  1240
BTV   HDIPPCWNETQGDKRGLISATKLSFFSKMARPAVQAALSDPQIMNLVEELPLGE-FSPGR
        * P      * K  *I**    F* K*      ** L     *  **    L E F*

Vacc  LSS--FHTTE-KSGAVKQKL--GFN-EFNN-----LTNLSKNKTEIITLVSDDISKLQSV  1300
BTV   ISRTMMHSALLKESSAKALLSSGYRLEYQKALNGWIAQVSMRLGEESGVISTSYAKL--F
      *S   *H*   K   *K  L  G*  E**       *  **S    E   **S  *  KL  *

Vacc  KINFEEFVCLGELNPDHSLFEKKQDRYVVDIIVNKLYIKRAEITELVVEYMIERFIS-FSV  1360
BTV   DVYFE----GELDGAPYMFPDQN-------LSPQFYIQKMMIGPRVSSRVRNSYVDRIDV
       * FE    GEL     *F   *        *   *YI ** I   V    *   **  * V

Vacc  IV-KEWGMETFIEDEDNIRFTVYLNFVEP--EELNLSKFMMVLPGAANKGKISKFKIPIS  1420
BTV   ILRKDVVMRGFITAN-----TI-LNVIEKLGTNHSVGDLVTVFTLMNIETRVAE---ELA
      I* K*  M  FI         T* LM**E        *  ** V*  *     **       *

Vacc  DYTGYDCFNQTKKLNKMTVELMNLKELGS--FDLENVNVYPGVWNTYDIFGIEAAREYLC  1480
BTV   EYMT----SEKIRFDALK--LLK-KGIAGDEFTM-SLNV--------------ATQDFI-
      *Y          **  *   L*  K *    F *  *NV                A  ***

Vacc  EAMLNTYGEGFDYLYQPCDLLASLLCASYEPESVNKFKFGAASTLKRATFGDNKALLNAA  1540
BTV   ----DTY---LAYPYQ---LTKT------EVDAISLY--CTQMVMLRAALGLPKKKMKIV
           TY   * Y YQ   L          E * *  *        * RA *G  K  * **

Vacc  LH---KKSEPINDNSSCHFFSKVPNIGTGYYKYFIDLGLLMRMERKLSDKISSQKIKEEM  1600
BTV   VTDDAKKRYKIRLQ---RFRTHVPKI---------------KVLKKLIDP-NRMTVRNNL
      *    KK   I  *    F * VP I                ** *KL D       **  *

Vacc  EETEDF-
BTV   EN--QFV                                                       1607
      E    F
```

Fig. 4. Homology between bluetongue virus segment 1 gene product and the 147-KDa subunit vaccinia virus RNA polymerase. The deduced amino acid sequences of BTV-10 L1 gene product were compared with the published sequence data of vaccinia virus RNA polymerase large subunit (21) using an optimum alignment computer program of Staden (8). Homologous amino acids in the two sequences are indicated by single letter amino acid abbreviations under each matching pair. Conservative amino acid changes are shown by stars.

Table 2 Alignment of BTV-10 VP1 protein with polymerases

Polymerases	No. of aa matches			No. of breaks	
	Possible matches	Real score	% identity	Real break	% break
Vaccinia 147K	1054	266	25	119	11
Yeast 215K	1118	262	23	141	12
Yeast 160K	1117	264	24	135	13
E. coli β chain	1130	259	23	127	11
Common tobacco Chloroplast β chain	922	267	30	120	13
Fruit fly 215K	463	151	33	70	15

viral polymerase is a DNA-dependent RNA polymerase. As presented in Fig. 4, with a minimal number of gaps, 266 amino acid residues out of 1054 possible matches are homologous between the two polypeptides (a score of 25% homology). However, when possible (conservative) mutational changes (K to R, D to E, S to T, N to Q or hydrophobic to hydrophobic residues) were included in the comparison (see asterisks in Fig. 4), the homologies between the two proteins revealed an even greater score. The total homology, including conservative changes was around 50%. Although the homologous amino acid changes were distributed evenly throughout the molecules, certain blocks of amino acids were more conserved than others. For example, amino acid residues 194-239, 572-598 and 1510 to 1541 scored between 54 to 60% homology.

DISCUSSION

The amino acid sequence of the VP1 protein of bluetongue virus was inferred from the complementary DNA sequence of BTV RNA segment 1. Computer-assisted analysis of the coding capacity of each of the six possible reading frames indicated the presence of a single large reading frame extending from nucleotide residues 12 to 3920. The deduced amino acid sequence indicates that the gene product has a predicted size of 149,588 Da. The molecule is highly basic protein with a positively charged carboxyl terminus (the last 50 residues have an overall charge of +13.5). Proteins that bind to nucleic acids often exhibit a net basic charge with clustered K and/or R residues possibly involved in interactions with the negative charged phosphoryl backbone (22).

The bluetongue virus VP1 bears a remarkable resemblance to vaccinia virus 147-K Da polymerase subunit. Both proteins are similar in size and exhibit a net basic charge at neutral pH. Broyles and Moss have recently reported (21) that the vaccinia virus 147-K Da subunit shares extensive homology with the largest subunits of the prokaryotic RNA polymerase and eukaryotic RNA polymerase 11 and 111. A computer alignment program was used to compare the predicted VP1 amino acid sequence with each of these polymerase subunits. The alignment comparison indicated significant homology (30%) between the BTV VP1 protein and the β-chain of the common tobacco chloroplast RNA polymerase (see Table 2). Comparison of the VP1 sequence with the sequences of the large subunits of the *Saccharomyces cerevisiae* RNA polymerase 11 (215-K Da) and 111 (160-K Da), the β-chain of the *E. coli* RNA polymerases (215-K Da) and the sequences of the fruit fly large subunit (215K Da) RNA polymerase revealed similar homologies (23-33%) (Table 2). It is likely that the large subunits of all six RNA polymerases have similar functional domains.

The degree of relatedness among the five types of DNA-directed RNA

polymerases and the RNA-directed RNA polymerase of BTV has implications for the evolution of the bluetongue virus. Bluetongue virus is an RNA virus with an RNA-directed enzyme, vaccinia virus is a DNA virus with a DNA-directed RNA polymerase. The indicated similarities of their putative polymerases suggests that they may have evolved from a common ancestor, an ancestral gene that has been derived from prokaryotic/eukaryotic cellular origins.

ACKNOWLEDGEMENTS

We thank the staff of the NERC Institute of Virology, especially Dr. J.J. Marshall for computer analysis and Mrs. J. Oswell for typing. This work was supported by EEC Contract BAP-0120 U.K. and by Alabama Research Institute grant No. 86-409.

*To whom correspondence should be addressed

REFERENCES

1. Mertons, P.P.C., Brown, F. and Sangar, D.V. (1984). *Virology* **135**, 207-217.
2. Drayna, D. and Fields, B.N. (1982). *J. Virol.* **41**, 110
3. Yamaguchi, S., Fukusho, A. and Roy, P. (1988b) (In press) *Virus Res* **11**
4. Purdy, M.A., Petre, J. and Roy, P. (1984). *J. Virol.* **51**, 754-759
5. Maniatis, F., Fritsch, E.F. and Sambrook, J. (1982). Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
6. Grunstein, J., and Hogness, D.S. (1975). *Proc. Nat. Acad. Sci.* **USA, 72**, 3961-3965
7. Maxam, A. and Gilbert, W. (1980). In Grossman, L and Moldave, K, (eds.) *Methods in Enzymology Academic Press*, **New York 65**, 499-560.
8. Staden, R. (1982). *Nucleic Acids Res.* **10**, 2951-2961
9. Lipman, D.J. and Pearson, W.R. (1985). *Science* **227**, 1435-1441
10. Ghiasi, H., Purdy, M.A. and Roy, P. (1985). *Virus Res.* **3**, 181-190
11. Purdy, M.A., Ghiasi, H., Rao, C.D. and Roy, P. (1985). *J. Virol.* **55**, 826-830
12. Purdy, M.A., Ritter, G.D. and Roy, P. (1986) *J. Gen. Virol.* **67**, 957-962
13. Lee, J. and Roy, P. (1986). *J. Gen. Virol.*, **67**, 2833-2837.
14. Lee, J. and Roy, P. (1987). *Nucleic Acids Res.*, **15**, 7207
15. Yu, Y., Fukusho, A. and Roy, P. (1987). *Nucleic Acids Res.* **15**, 7206.
16. Yu, Y., Fukusho, A., Ritter, D.G. and Roy, P. (1988) *Nucleic Acid Res.*, **16**, 1620
17. Rao, C.D., Kiuchi, A. and Roy, P. (1983). *J. Virol.*, **46**, 378-383.
18. Verwoerd, D.W., Louw, H. and Oellermann, R.A. (1970). *J. Virol.* **5**, 1-7
19. Huismans, H. (1979). *Virology*, **92**, 385-396.
20. Dayhoff, M.O., Hunt, L.T. and Hurst-Calderon, S. (1978). In Dayhoff, M.O. (Ed.) *Atlas of Protein Sequence and Structure. Natl. Biomed. Res. Found., Washington D.C.* **5(3)**, 363-373.
21. Broyles, S.S and Moss, B. (1986). *Proc. Natl. Acad. Sci. USA.* **83**, 3141-3145
22. Larder, B.A., Kemp, S.D. and Darby, G. (1987). *EMBO Journal*, **6**, 169-175
23. Kyte, T. and Doolittle, R.F. (1982). *J. Mol. Biol.*, **157**, 105-132.