

Complete sequence of the NS1 gene (S6 RNA) of US bluetongue virus serotype 17

Lin-fa Wang, Roy H.Do, Bennie I.Osburn¹ and Ronald Y.Chuang²

Department of Biochemistry & Biophysics, ¹Department of Veterinary Pathology and ²Department of Pharmacology, University of California, Davis, CA 95616, USA
 Submitted August 8, 1989

EMBL accession no. X15891

The complete nucleotide sequence of the double-stranded RNA segment 6 (S6) of bluetongue virus serotype 17 (BTV-17), which encodes the nonstructural NS1 protein (1), has been determined from two overlapping cDNA clones. The S6 RNA was deduced to be 1769 base pairs in length, coding for a 552 amino acid protein (64,541 daltons). The sequence was 99% identical at both the RNA and protein levels to that of BTV-10 S6 (or M6) segment (2), confirming the highly conserved nature of this genome segment, as revealed previously by northern analysis (3). The deduced amino acid sequence and the coding frame were also consistent with what had been observed from our gene fusion and expression studies using this genome segment (1).

```

GTTAAAAAGTTCCTAGTGGCAACCCAAACATGGAGCGCTTTTGGAAAAATACAATATCAGTGGGGATTACCGAAATGCCAGGAACTTT      96
M E R F L R K Y N I S G D V A N A T R T F      21
TTTGGCTATTTACCACAGTGGACTTGCAGTCACTAAAAAGGAATTGCTATTAATGGAATGTGTGTTAAGCAGCATTITGAGAGACCGATGAT      192
L A I S P Q W T C S H L K R N C L F N G M C V K Q H F E R A M I      53
CCGGCAACTGATGCGGAGGAGCCGGGAAACGATACAAATGGTTGATGGCAAAGGAAGCAATGTATGATCGGGAACAGTCTGGCTCAATG      288
A A T D A E E P A K A Y K L V E L A K E A M Y D R E T V W L Q C      85
TTTCAAAGCTTTTCCCAACCGTACGAGGAGGATGTCGAAGGAAGATGAAGCGATGCGGAGCCAAATGTCTGAGGATTACCGCAAAAGTGGAT      384
F K S F S Q P Y E E D V E G K M K R C G A Q L L E D Y R K S G M      117
GATGAATGAAGCCGTGAAACAATCGCGTGGTTAATTCAGAAAGAAATTAGATTGGATGATTCACITTCGCGCAATGCCTACATCTACGTGCAAT      480
M N E A V K Q S A L V N S E R I R L D D S L S A M P Y I Y V P I      149
CAATGATGGTCAAATTTGTTAATCCGACATTTATATCAAGATATCGCAAAATTCATATTATTTTACAACCCAGATGCAGCTGATGGATTGA      576
N D G Q I V N P T F I S R Y R Q I A Y Y F Y N P D A A D D W I D      181
TCCAAATCTCTTTGGCATTGCGTGGACAGCAAAATCAGATTAACCGTGAAGTTGAGAGACAAATTAACACATGCCCCTTACACTGGATACAGAGGTAG      672
P N L F G I R G Q H N Q I K R E V E R Q I N T C P Y T G Y R G R      213
AGTGTTCAGAAATGTTCTTGGCGATTGAGCTGATCAATTTTTGAGAAATGGATGATTTTGGCAAGCATTITTAACAGGTACCGCTCGATGGCGAT      768
V F Q V M F L P I Q L I N F L R M D D F A K H F N R Y A S M A I      245
ACAACAATATCTGAGAGTGGTTATGCTGAAGAGATCAGATATGTACAACAGCTCTTCGGAAGGGTCCCAACAGGATGATTTCCATTACACAGAT      864
Q Q Y L R V G Y A E F I R Y V Q Q L F G R V F T G E F P L H Q M      277
GATGCTGATGAGACCGATCTCCGAACACCGGATCCGAGTATTTGGAGGCCCGGGTGGAGATCAGTGATGAGAATGGCAAAGCTGGCTATT      960
M L M R R D L L P T R D R S I V E A R V R S G D E N W Q S W L L      309
ACCTATGATCATCTCGTGGGGTGGATCATCAGGATCGTGGGAATGGTTTATGATTACATAGSATAAGCAATACATGTCACCTTTGCTA      1056
P M I I I R E G L D H Q D R W E W F I D Y M D R K H T C Q L C Y      341
CTTGAACACTCAAACAGATCCCGACTGTAGTGTGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT      1152
L K H S K Q I P A C S V I D V R A S E L T G C S P F K M V K I E      373
AGAGCATGTAGGAAATGATTCAGTGTTTAAACAGAAATAGTTCGCGATGAGCAAAATCGGACAGATGGAGATGATTAATATACAACAAATTTGTA      1248
E H V G N D S V F K T K L V R D E Q I G R I G D E G W Q E G I      405
CACTGGAGCGGAGGCATTGATTAACAACCGGATTCACATTCATCGCTGGATTAGGGGGTCTGGCACTGGAACGATGAAGGATGGCAGGAGGTAT      1344
T G A E A L I T T A I H I H R W I R G S G I W N D E G W Q E G I      437
F M L G R V L L R W E L T K A Q R S A L L R L F C F V C Y G Y A      1440
ACCACGCGCAGCGGACGATACCGGACTGGATAATCTTGGAACTTTTTGGATATCATTITGAAGGGCCGAGAAGTCTAGTGAAGATGAGGATGA      1536
P R A D G T I P D W N N L G N F L D I I L K G P E L S E D E D E      501
AAGAGCTTATGCTACAATTTTGAATGGTTGATGATTAATAACTCTATGCTATGCAAGAAAGGTTCACTTCGCTGGTTCGCTGCGCTCGCGTGTG      1632
R A Y A T M F E M V R C I I T L C Y A E K V H F A G F A A P A C      533
TGAAGCGGGGAAAGTAAATTAATCTTGGCGCATGCTCAGATGGATGGAGTATTAGTTACTGATTTTTGGTTTTTATTTCTTTCTTTTCATT      1728
E G G E V I N L A A R M S Q M W M E Y *      552
TCTATTTCTCTTACACTCTACTAGAAGCTTTTCACTTAC      1769
    
```

Acknowledgements

This work was supported by HHS research grants GMS 19673 (RHD) and DA 05901 (RYC).

References

(1) Wang, L-f., et al. (1989) *Biochem. Biophys. Res. Commun.* 162, 892-899.
 (2) Lee, J. and Roy, P. (1987) *Nucl. Acids Res.* 15, 7207.
 (3) Unger, R. E., et al. (1988) *Virology* 167, 296-298.