

Nucleotide sequences of three Nodavirus RNA2's: the messengers for their coat protein precursors

Ranjit Dasgupta and Jean-Yves Sgro

Institute for Molecular Virology, University of Wisconsin, Madison, WI 53706, USA

Submitted August 4, 1989

EMBL accession nos X15959-X15961 (incl.)

Nodamura virus (NOV), Flock House virus (FHV) and Boolarra virus (BOV) are members of the family of insect virus called *Nodaviridae*. They are serologically distinct from each other. NOV was originally isolated from mosquitoes in Nodamura, Japan. The natural hosts for FHV and BOV were separate species of grass grubs from New Zealand and Australia respectively. All nodaviruses infect and grow well in wax moth larvae. FHV and BOV are routinely propagated and grown in cultured *Drosophila melanogaster* cells. NOV does not infect *Drosophila* cells. It is unique among the *Nodaviridae* in its ability to cause fatal infection in vertebrates (suckling mice, ref 1).

The genomes of Nodaviruses consist of two single stranded messenger sense RNAs contained in a single virion. RNA1 codes for the replicase function and RNA2 codes for the precursor for the capsid protein. Both RNAs have a 5'-terminal cap m7GpppGp and a blocking moiety, presumably a protein at their 3'-termini (2). We report here the complete nucleotide sequence of NOV, FHV and BOV RNA 2's derived from virion RNA's and their full length or near full length cDNA clones. Oligonucleotides used as primers at the 3'-end for the synthesis of single stranded cDNAs were designed from the known sequence of Black Beetle Virus (BBV) RNA2 (2) assuming a homology between these viruses. Transcripts derived from cloned FHV cDNA were infectious to cultured *Drosophila* cells (3). Sequences of the cloned cDNAs were determined from the fragments produced by restriction digestions and also after making nested deletions using exonuclease III, by means of the dideoxy chain termination method. 95 bases at the 3'end of NOV RNA2 and 60 bases at the 3'end of BOV RNA2 were also determined by direct enzymatic RNA sequencing of ³²P-labelled fragments produced by limited ribonuclease T1 digestion of virion RNA2. The initiator AUG codons for all these RNAs (underlined in Fig. 1) are very close to the 5'-termini. NOV, FHV and BOV code for 399, 407 and 403 amino acids respectively. NOV RNA2 has a direct repeat of 21 nucleotides (b 1235-1255 and b 1262-1282) near the 3'-end. Comparison of these sequences and also the published sequence of Black Beetle Virus RNA2 (2) shows that FHV and BBV are very closely related to each other (80% homology) whereas BOV and NOV are somewhat distantly related to BBV (50% homology).

```

N   1  CUAAACAACCAADUACAUC UGGUAUCCAAAGCAGCAGCC CGCCGAGAGCGGCCCCUCG ACAACAACACCGACAACAGU
O  81  CGAACCCGUGCAUCAACCCAG CCUGUCGGAGGAGAGUCUG CGUACUCGAGACAGCAGCA GAUUGGCAGCACAAACAC
V 161  AUCCUGAAAAUGAGCGCUC CGGACUCGADUUUCUGAAAU GUGCCUUUGCUCCGCCGGAU UUUUCCACCGAUCCCGGCA
241  AGGUAUCCAGACAADUUUC AGGGUCUCGUUUACCGAAG AAACAUUCUCUGACCCAGUC GAUAACGUUUACUCCGGGA
R 321  AACAGACGAUGCUUGUGGUU GCACCUAUUCUGGAAUUGC UUGCCUGAAGGCAGACGGA AUGUUGCGCCAUCCUUUCA
N 401  GGUGUCCUCUUGCCUCAGU UGAAUUUCCAGGGUUGGACC AGCUGUUUGGCACGUCAGCA ACTGACACCCGAGCUAUGU
A 481  CACUGCUUUCGGUANGGU CAUUGGCAGCCGGUUGUAC CCCACGACAAUCUAAUGCA AUUUGCUGGGUCAADACAG
2 561  UGUACAAGAUACUCUAAA CAGGUUCUUAUCUCUACUC UCAGACCGUAGCGACCGUAC CACCUACCAACUUGGUCAG
641  AACACAAUUGCAADAGAUGG ACUAGAAGCGUAGADGCAC UACCAAUAACAACUACUCC GGUUUUUUUUCGAGGGUUG

```

Nucleic Acids Research

721	UUUAUUCACAAUCCGGUGUGCA	AUGAACCCUAGUUGUAGUUC	CADCCCAUCAUGGAGGGUUA	CGCGUCGGUCCACC
801	ACGUAAACCAUUCGCUAAGCU	AGCAUUGUACACCAAUUCUAC	CUUCUCAGGGCCACGAUUA	CCGGUCUUGGGCAGUUGGAC
881	GCAAUUGCGAUCCUUGUGAC	AGCCGCCACCCGGUGCCGUCA	ACACGGCCGUGCUGAAGGUC	UGGGCCUUGCGUAGAGUACG
961	UCCGAAUUCCAAUUCCACCC	UCUAUGAGUUUCGUCGUGAG	UCACCAGCAAACGACGAUA	CGCGUCGCCCGCUAUAGGA
1041	AAAUUGCCAGAGAUAUUCG	AUUGCCGUCGCUUCCAAAGA	CAACGCCACAUUUUGGGAAC	GGCUCCGAUCCAUUCUGAAA
1121	UCUGGGUCUAAUUCUUGCUUC	GACCAUACCUUGGCCCGUAG	GAGUGGCAGCGACAGGGAUC	AAAGGCACUAUUGAAACCAU
1201	UGGUUCUUGUGGGUUGGAU	UCCACCACAGAAAGCGUUGA	CGACGCAAAACGUCUUA	GGGUUGACGACGCAAAACCU
1281	CCCCAAGUCGCGACACCGA	CCCUAUACCCAUUCUAGGG	UCUUAACCCUUCUGX	1335
F 1	GUAAACAAUUCCAAGUUC	AAAUUGGUAAUAAACACAGA	CCAAGACGUCACAGGUC	ACGGGUUGCGUCACAACAA
H 81	CCCAAACAGCCGCUUCCA	CAGCAAAAACGGUCACGUA	UGGUAGACGCCGACGUAUC	GCACGAGCGUAUUCGCCGA
V 161	CGUGUGCCGGGAUUAACAU	GGCGCGCUAACCAUAUUA	GUCACCCUGGUUUGCGGUU	CUCAAUUGGCAUUGCACC
241	ACCUAGAUUCACACCGGACC	CCGGUAAGGGAUACCUGAU	AGAUAUUAAGGCCAAAGUGU	CAGCCGAAGGAUUGCCUCA
R 321	AUCAAUUCUACAGUUUACU	GCCGGACAGGACAUUUUAU	ACUCAUCGCAUUCACCCCGC	GAGUCGCUACUGGAGUGCU
N 401	AGCGUUCUGUGGUAUCUU	UCCUAUCUAGUGCGCAUCU	UUAACCCCGUUAUUUACCG	GGUUUACAUUGAUUGUCCG
A 481	AAACAUCUACAACUUCAGG	CCGAUCAGGUGUCUUAUCU	AGGUACGCUUCCAUGAACGU	GGGUUAUUACCAACGUGCA
2 561	ACUUGAUGCAGUUUGCCCGA	AGCAUAACUUUUGGAAUUG	CCCUUCUAAAGCUUGAGUAC	UGCAAUUCCCGUUUGCAACA
641	GAUCAGCCACAGUUCGCU	AGUUCAUAUCUUGUUGGUU	UAGAUGGUUUCUAGCGGUG	GGCGUCGACAAUUCUCUGA
721	GUCAUUCUACAAAGGAGUGU	UUUCACAGUCGCGUUGAAC	GAGCCUGACUUUGAAUUCAA	UGACAUAUUGGAGGUAUCC
801	AGACAUUCGACUUCGUAUUG	GUUCUUUCUUGUUCUACGGG	UCAACUUUACCAUUGGACU	CAGGAGCAGAAGCCACCAGU
881	GGAGUAGUCGGAUUGGGCAA	UAUGGCACAGAUUGCAUCC	GUGUCUCCGCGCCUGAGCGC	GCAGUUUACUCUGCCAUUCU
961	CAAGGCAUUGGCUUCGAUUG	AGUAUCGACCAAAUCCAAAC	GCCAUUGUUAACCAAUUCGG	CCAUUAUUGCCUUCUUCG
1041	AUGAGGCGCGCUUCAGGAA	UACCGUACGGUUCGCGAGAU	UUUGCCGGUUCGCAUGADAG	CGGCCAAAUAUGCAUAAUG
1121	UGGGAGAGUGAAUUCUUA	CAUUAUUUCCUCCUGGCUU	CUCAAGCCAAUUCUCCCGCG	CCGAUCGUGUUGCCCGCAAG
1201	UGGUUAUAGUGGACUUCAG	CCUUUUUUGAAGGAUUUGGC	UUUUAGAAGCAUCGCGACGC	CAACCUAAACCGGCAUGAU
1281	CCGAACAUCGGAUUCUUGG	CCACAUAAGCCCAUUUUGG	UUGAAGAUUAAGUAGUGAG	CCCCUUAGCGGAAACCGG
1361	AAUUUAUUCCAAACCGAGU	UUAGGCAACAGACUAAGGU	1400	
B 1	GAUUUCAACAUGGACCGCCAC	GACGACAACAACGUCCCAAA	GGCCAAUUGGCCAAGCUAA	ACAAGCUAAACAACCUUCAG
O 81	CCCGUUCUAGACGCCACCGU	AGCGCGACAGCAGCAGCAU	CACCGAGAACAUCUACDGA	UGCUAUCUGAACCCGGACUC
V 161	AGAUUUUGAAUUGUGCGUU	CGCUUCGCGGACUCCAACA	CGGAUUCUGGCCAAGGGUAU	CCUUAUAUUUUUGAAGGUA
241	UCUACUCUUCAGAAAGAUUG	UCUAUACAGAAACUGGUGUA	AAUUUUAGUGGCGCAACAAC	GCAAAUAUGUUAUCUUAAC
R 321	UCADAGUUCUACCAACCCA	GGUGUGGCUUCUUGCGCGU	UAUUAAAACAGCUACCGCGC	CAGCGCAACCAGCCGCAUUG
N 401	ACAACGACGGAUUGUUUUAC	AGCCGUUCUUUCUUGAUU	UCACUUCACUUUCGUGUAC	ACCGCUACAUAUCGUGCCGA
A 481	UCAAUUGCCGCUUUAAGU	ACGCCAGCAUAAUUUUUGU	UUGUACCCUACUUGCAACUC	CACCGAGUAUUUUGGUGCA
2 561	UCAGUUGUUGAAGGGGGCU	GUCCAAAUGUCAUCACACA	GUUUCGUGUUGAUACCCAC	CUAGUCGUCUCCAAUCUAGU
641	CACGCUAUACUGGUCUCGA	AAGCGCAUUAAGGUUGGUG	ACGAAAUAUACAGCGGACUG	UUUAUCGACGGUUGUUUAC
721	CCAAUCCAUAAACGGCAAGC	CAGAAUUUCGUCUAUCUUC	AUCUUGGAAGGUGUACAGAC	UCUGGCCAGGGCAAAUUGUA
801	CCGUUGCCCAAGCCGCUAUG	CCUUUUUCAUUAAGUUCUGG	AGCGGCAACUUGUCCCGGAU	UUACUGGAUUUGGAGGCAUG
881	GAUGCUUAUUAUAAUAGU	GACAGCAGCAGCUGGUCACG	UCAUACGCGCCACCAUUA	ACAUGGGCUUGAUUUGAGUA
961	CCGCCAAACACUAACACUG	CACUUUAUAAUAUGCACAU	GAUUCACCCGCUAAGACAU	UAUAGCAUUACAACAUAACA
1041	GGAAAGUUACAAUUCUUA	CCUGUUGCCGUACGAGCGAA	GUCUACCGCCAAUUGGUGG	AGAGAGUUAAACGGUUAUUC
1121	AAAGCCGGUUUGGUCGCGCG	UUUCUUAUUGUCCCGGUCUG	UCCGUGGAAUUGCCACCGGU	GUUCAACAUAUUGGAGAUUC
1201	GAUUGCAGAGCUUCUUUUU	GAGAGGACCAACGUGGCGCA	CGUACAAACGCCCCUAGCG	CAGUGAACUUAUGUUUCACAC
1281	CAACUUAUUGUACACAGAC	UAAGX	1305	

Figure. 1 Nucleotide sequences of NOV (top), FHV (middle) and BOV RNA2's (bottom). Start and stop codons are underlined. The sequence of one or two bases at the 3'-end of NOV and BOV RNA2 (indicated by an X) could not be accurately determined due to the presence of a blocking group.

References: (1) Bailey, L. & Scott, H. A. (1973) Nature, 241, 545.
 (2) Dasgupta et al. (1984) Nucl. Acids. Res. 12, 7215-7223.
 (3) Dasmahapatra et al. (1986) Proc. Natl. Acad. Sci. USA 83, 63-66.