Nucleotide Sequence Analysis of RNA-3 and RNA-4 of Beet Necrotic Yellow Vein Virus, Isolates F2 and G1

By S. BOUZOUBA, H. GUILLEY, G. JONARD, K. RICHARDS* and C. PUTZ

Laboratoire de Virologie, Institut de Biologie Moléculaire et Cellulaire, 15, rue Descartes, 67084 Strasbourg Cédex, France and Station de Pathologie Végétale, I.N.R.A., 28, rue de Herrlisheim, 68021 Colmar Cédex, France

(Accepted 28 February 1985)

SUMMARY

The nucleotide sequences of cDNA clones corresponding to RNA-3 and RNA-4 of beet necrotic yellow vein virus isolates F2 and G1 have been determined. The cDNA of RNA-3 of isolate F2 is 1775 residues in length and contains a coding region of 219 codons. In isolate G1 this coding region has undergone an internal deletion of 354 nucleotides in such a way as to conserve a shortened reading frame. Otherwise, the RNA-3 sequences of the two isolates were closely similar. RNA-4 of isolate F2 has an extrapolated length of 1431 residues and contains an open reading frame of 282 codons. This open reading frame has undergone an internal deletion of 324 nucleotides in one cDNA clone of RNA-4(G1) with conservation of a shortened reading frame. Sequence analysis of other RNA-4(G1) cDNA clones revealed, however, that the exact boundaries of the deletion are not always the same. RNA-3 and RNA-4 of each isolate are more than 90% homologous for the 3'-terminal 200 nucleotides. Short homologous sequences are also present in RNA-3 and RNA-4 of isolate F2 flanking the regions deleted in each of these RNAs in the G1 isolate. These homologous sequences probably play a role in the deletion process.

INTRODUCTION

Beet necrotic yellow vein virus (BNYVV) causes a severe disease of sugar beet (Beta vulgaris) called rhizomania (Tamada & Baba, 1973; Tamada, 1975). Preparations of virions typically contain four plus-strand RNA species, with lengths of 7100, 4800, 1800 and 1500 nucleotides for isolate F2 (Richards et al., 1985). The viral coat protein is encoded by RNA-2 while hybridization studies have shown that the two shorter RNAs share no significant sequence homology with RNA-1 and RNA-2, thus ruling out the possibility that RNA-3 and RNA-4 are subgenomic RNAs derived from the longer species (Richards et al., 1985). It is not yet known, however, if RNA-3 and RNA-4 are essential for viral infectivity or whether one or both species are satellite RNA.

While RNA-1 and RNA-2 of different BNYVV isolates are the same size, RNA-3 and RNA-4 may vary considerably in length between isolates. We have already described isolate BNYVV(G1) in which RNA-3 and RNA-4 are 1400 and 1150 nucleotides long rather than 1800 and 1500 nucleotides as in F2 (Richards et al., 1985). We have also recently observed a BNYVV isolate with an RNA-3 of 1200 nucleotides and have now obtained a subisolate of F2 by single local lesion inoculation in which RNA-3 is less than 1000 nucleotides long (unpublished observations). In order to learn more about the structure of BNYVV RNA-3 and RNA-4 and to investigate their variations in length, we have cloned and sequenced cDNA copies of both RNA-3 and RNA-4 from the F2 and G1 isolates. The results have allowed us to characterize the coding regions of each RNA, to identify the sequences responsible for the length variation and to pinpoint conserved nucleotide sequences which may be important in the process by which such length variations occur.
METHODS

Viral RNA. Isolates BNYVV(F2) and BNYVV(G1) have been described previously (Richards et al., 1985). Virus was propagated in Chenopodium quinoa. Nucleic acid was extracted from purified virus with phenol and concentrated by ethanol precipitation (Richards et al., 1985). The RNA was then usually further purified by chromatography on a 1 cm column of oligo(dT)-cellulose (P-L Biochemicals) poured in a Pasteur pipette and equilibrated with 10 mm-Tris–HCl pH 7.5, 1 mm-EDTA, 0.5 M-LiCl, 0.1% SDS. After loading, the column was washed with several volumes of loading buffer and the bound BNYVV RNA was eluted with 2 ml of the same buffer but lacking LiCl. The RNA was concentrated by ethanol precipitation, dissolved in sterile water at a concentration of 1 mg/ml and stored frozen.

cDNA synthesis. cDNA of BNYVV RNA was cloned by the method of Heidecker & Messing (1983) with minor modifications. The reaction mixture for cDNA synthesis contained 10 mm-MgCl2, 70 mm-KCl, 50 mm-Tris–HCl pH 8.2, 2 mm-dithiothreitol, 800 μM each of dATP, dGTP, dCTP, dTTP, 5 μCi [α-32P]CTP (3000 Ci/mmol), 25 μg/ml actinomycin D, 1 μg oligo(dT)-tailed pUC9 (P-L Biochemicals) and 2 μg BNYVV RNA, all in a final volume of 15 μl. Thirty units of reverse transcriptase (Life Sciences, St Petersburg, Fla., U.S.A.) were added and the mixture was incubated at 37 °C for 90 min. The reaction mix was then diluted to 50 μl with water, extracted with phenol and the DNA was ethanol-precipitated three times in such a way as to minimize co-precipitation of unincorporated deoxynucleotide triphosphates (Heidecker & Messing, 1983). In order to add dC tails the pellet of cDNA-plasmid conjugates was taken up in 10 μl water and 4 μl 1 M-potassium cacodylate pH 7.0, 1 μl 1 mm-dCTP, 1 μl 50 mm-dithiothreitol, 2 μl 20 mm-MnCl2 and 20 units deoxynucleotidyl terminal transferase (Bethesda Research Laboratories) were added. The reaction was carried out at 37 °C for 20 min followed by phenol extraction and ethanol precipitation and the pellet of dC-tailed cDNA-plasmid conjugates was taken up in 10 μl water.

The dC-tailed cDNA-plasmid conjugates were diluted to 50 μl in 5% sucrose in 0.2 M-NaOH, 0.8 M-NaCl, 1 mm-EDTA and layered on 5 to 20% linear gradients of sucrose poured over 0.5 ml 60% sucrose cushions made up in the same alkaline solution. Gradients were centrifuged at 36000 r.p.m. for 17 h at 4 °C in a Beckman SW50.1 rotor. Fractions of 0.3 ml were collected and their radioactivity was determined by Cerenkov counting. The two or three fractions corresponding to the peak of radioactivity at the interface between the gradient and the cushion were added and mixed with 2.5 μg oligo(dG)-tailed pUC9 (P-L Biochemicals). The material was then dialysed overnight at 4 °C against 10 mm-Tris–HCl pH 7.5, 10 mm-NaCl, 1 mm-EDTA and concentrated by ethanol precipitation with 20 μg carrier Escherichia coli tRNA. The precipitated DNA was dissolved in 100 μl water, and concentrated solutions of NaCl, Tris–HCl pH 8 and formamide were added to give final concentrations of 50 mm-NaCl, 10 mm-Tris–HCl and 32% formamide in a volume of 1 ml. The mixture was incubated for 24 h at 37 °C, dialysed overnight at 4 °C against 10 mm-Tris–HCl pH 8, 100 mm-NaCl, 1 mm-EDTA and the annealed DNA concentrated by ethanol precipitation.

The precipitate of annealed DNA was taken up in 50 μl of a solution containing 20 mm-Tris–HCl pH 7.5, 10 mm-MgCl2, 50 mm-NaCl, 1 mm-dithiothreitol and 100 μM each of dCTP, dGTP, dATP and dTTP. Five units DNA polymerase Klenow fragment (Boehringer) were added and the mixture was incubated for 1 h at 15 °C and 1 h at room temperature. After phenol extraction the DNA was precipitated with ethanol and resuspended in 50 μl H2O.

Cloning. Cells of E. coli strain JM103 were rendered competent by the CaCl2 procedure (Mandel & Higa, 1970) and 106 cells in 2 ml were mixed with recombinant DNA. The mixture was kept in ice for 20 min, at 42 °C for 2 min and at room temperature for 30 min before plating 200 μl aliquots on B broth plates (Messing, 1983) containing 50 μg/ml ampicillin and upon which 40 μl 100 mm-isopropyl β-D-thiogalactoside (Boehringer) and 40 μl 2% 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (Boehringer) dissolved in dimethylformamide had been spread. Plates were incubated at 37 °C overnight and the white colonies which contain cDNA inserts were selected for further analysis. About 500 clones were generally obtained per μg of input RNA. All clones used in this study were prepared from a single preparation each of BNYVV(F2) and BNYVV(G1) RNA.

Screening the clones. Small scale isolations of plasmid DNA were carried out on 1·5 ml of bacterial culture by the alkaline lysis method (Birnboim & Doly, 1979). The size of the cDNA inserts was estimated by agarose gel electrophoresis after double digestion with HindIII and SmaI restriction endonucleases, enzymes which cut in the polylinker at each extremity of the insert in pUC9 (Vieira & Messing, 1982). Plasmids with an insert of 1000 or more nucleotides were further analysed by using them as probes for Northern blots. A portion of the plasmid minipreparation corresponding to about 0.5 μg DNA was 32P-labelled by nick translation (Rigby et al., 1977) and hybridized to BNYVV RNA immobilized on diazobenzyloxymethyl paper (Alwine et al., 1977). Autoradiography of the blots permitted each clone to be assigned to the RNA species from which its was derived (Richards et al., 1985).

Sequence analysis. The cDNA insert from clones selected for sequence analysis was isolated by preparative agarose gel electrophoresis after double digestion with restriction enzymes chosen to cut in the polylinker on each side of the insert: HindIII and either BamHI, SmaI or EcoRI. DNA bands were visualized in the ethidium bromide-stained gel under long wavelength u.v. illumination and the DNA was extracted from the gel band by the
method of Vogelstein & Gillespie (1979). Inserts were digested with restriction enzymes and the fragments were 5'-\(^{32}\)P-end-labelled with polynucleotide kinase or 3',5'-\(^{32}\)P-end-labelled with reverse transcriptase and the appropriate (\(\alpha\)-\(^{32}\)P)deoxynucleoside triphosphate (Franck et al., 1980; Guilley et al., 1982). After strand separation on polyacrylamide gels the fragments were sequenced by the method of Maxam & Gilbert (1977). Sequence data were analysed using the University of Wisconsin Genetics Computer Group software package (Devereux et al., 1984) run on a VAX 11/750 computer.

RESULTS AND DISCUSSION

Sequence analysis

Table 1 summarizes the characteristics of the clones chosen for sequencing. The cDNA inserts were isolated as described in Methods and cut with one or more appropriate restriction enzyme. The resulting DNA fragments were 32P-end-labelled, purified and sequenced by the partial chemical degradation method of Maxam & Gilbert (1977). Overlaps among the various subsequences were used to define a non-ambiguous primary structure for each insert. The inserts of pBA4, pBF11, pBA3 and pBF4 were sequenced on both strands over 80% or more of their total lengths. Details concerning derivation of the sequence will be provided on request.

One reason we chose to clone the cDNA inserts in a pUC plasmid was because the restriction sites present in the polylinker flanking the insert in such constructs should be useful for sequencing the extremities of the insert. This proved to be the case for the 5' termini of the inserts but attempts to sequence the 3'-terminal regions from a polylinker site were unsuccessful. At the 3' terminus the insert is joined to the polylinker by a poly(dA) sequence of about 70 nucleotides and sequence gels of fragments encompassing this region were impossible to interpret beyond the homopolymer tract. The unreadable gels had a strong echo pattern in the portion of the sequence ladder above the poly(A), suggesting length heterogeneity in the poly(A) sequence. This is probably caused by the bacterial DNA polymerase ‘slipping’ on the long homopolymer tract during plasmid replication. Because of this difficulty it was necessary to sequence the 3'-terminal portion of the inserts from restriction sites preceding the poly(A) tract.

Sequence of RNA-3

The sequence of clone pBF11, corresponding to RNA-3 of isolate F2, is presented in Fig. 1. The insert is 1775 base pairs in length plus a poly(A) tract of about 70 base pairs, giving a total length of about 1850 nucleotides which compares well with the length of 1800 nucleotides estimated for RNA-3 of isolate F2 by agarose gel electrophoresis (Richards et al., 1985). Experiments in which decapped 5'-\(^{32}\)P-labelled RNA-3 was sequenced by direct methods show that the pBF11 insert may correspond to full-length RNA-3 of isolate F2, lacking at most only a few residues (unpublished observations).

Fig. 1 also presents the sequence of pBA4, a clone of RNA-3 from the G1 isolate. The pBA4 insert is one base pair shorter than pBF11 at the 5' terminus, perhaps because of premature termination of cDNA synthesis. It is evident from Fig. 1 that the RNA-3 sequences from the F2
Sequence of BNYVV RNA-3 and RNA-4

Fig. 1. Sequence of cDNAs corresponding to BNYVV RNA-3, isolates F2 and G1. The F2 sequence is from clone pBF11 and the G1 sequence from clone pBA4. Point mutation differences between the F2 and G1 sequences are indicated by asterisks beneath the sequences. Asterisk above the sequence indicates termination codon ending the long open reading frame discussed in the text.

and G1 isolates share a great deal of sequence homology. The two sequences are identical in all but 41 positions. With one exception the sequence variations are simple base substitutions, 30 transitions and 10 transversions. The most striking difference between the two sequences, however, is a segment of 354 nucleotides present in RNA-3 of isolate F2 between positions 730 and 1083 which is missing from RNA-3 of isolate G1 (Fig. 1). This extra sequence can account for essentially all of the observed length difference between RNA-3 of the F2 and G1 isolates. From the sequence data alone we cannot conclude whether RNA-3 of the G1 isolate has been derived from RNA-3 of F2 (or a common ancestor) by deletion of the 354 residue sequence or if F2 has arisen from G1 (or a common ancestor) by insertion of this sequence. Consequently, we shall refer to the supplementary sequence as an insertion/deletion (I/D) element. As will be shown below, however, the sequence variation in the population of RNA-4(G1) molecules near the boundaries of the insertion/deletion site favours the idea that deletion is the explanation.
The RNA-3(F2) sequence contains a long open reading frame beginning with an AUG at nucleotide 447 and extending to a UAG at nucleotide 1106 (Fig. 2) which can encode a polypeptide (P25) of 219 residues in length with a mol. wt. of about 25000 (25K). This polypeptide undoubtedly corresponds to the RNA-3(F2) in vitro translation product estimated previously to have a mol. wt. of about 27K by PAGE (Richards et al., 1985). Interestingly enough, the I/D element which distinguishes between RNA-3 of the F2 and G1 isolates falls within the open reading frame (Fig. 2). Since the length of the RNA-3 I/D element, 354 nucleotides, is a multiple of 3, RNA-3(G1) should encode a polypeptide of about 12K which has the same N and C termini as P25 but lacks an internal polypeptide sequence of 118 amino acids (Fig. 2).

Sequence of RNA-4

Fig. 3 presents the sequence of clones derived from RNA-4 of the F2 and G1 isolates. The cDNA insert contained within clone pBA3, corresponding to RNA-4(G1), has a length of 1107 base pairs plus a poly(A)tract of about 70 base pairs for a total length of 1180 base pairs. This compares well with the length of 1150 nucleotides estimated for RNA-4(G1) by agarose gel electrophoresis suggesting that pBA3 contains a nearly full-length insert. Direct RNA sequence analysis of the Y-terminal sequence of RNA-4 has been hampered by contamination of decapped 5'-32P-end-labelled RNA by RNA-3 degradation products. Putz et al. (1983), however, have reported that the 5'-terminal structure of RNA-4 from another isolate of BNYVV is probably m7GpppA, suggesting that pBA3 lacks a few nucleotides at its 5' terminus.

For RNA-4(F2) the sequence presented in Fig. 3 was obtained from analysis of two clones. Clone pBF4 was found to correspond to the 3'-terminal 1000 residues of the sequence and had a perfect inverted repeat of 59 base pairs at its 5' extremity. This inverted repeat sequence was not present in the corresponding region of the otherwise very similar RNA-4(G1) sequence, suggesting that it was a cloning artefact. Such sequence rearrangements are frequently found at the 5' end of inserts of less than full-length cDNA clones (Fields & Winter, 1981; Volckaert et

---

**Fig. 2.** Organization of BNYVV RNA-3 and RNA-4. The long open reading frames referred to in the text are indicated by boxes. The hatched areas represent the insertion/deletion (I/D) elements present in the RNA of F2 isolate but absent from RNA of G1. The size of the potential translation product for the RNA from each isolate is indicated above the corresponding open reading frame (P25 = 25K etc.). Nucleotide (NT) numbers referring to Fig. 1 and 3 are indicated beneath each diagram.
Fig. 3. Sequence of cDNA clones corresponding to BNYVV RNA-4, isolates F2 and G1. The F2 sequence is from clones pBF4 and pBFI5 and the G1 sequence from clone pBA3. Point mutations are indicated by asterisks beneath the sequences. Asterisk above the sequence indicates termination codon ending the long open reading frame discussed in the text.
Fig. 4. Sequence flanking RNA-4 deletion point in the F2 isolate (pBF4) and in three cDNA clones from the G1 isolate (pBA3, pBA7 and pBA5). The amino acid-coding capacity is given above each sequence in single letter code. A dash has been inserted in the pBA5 sequence for alignment purposes.

al., 1981; Van Wezenbeek et al., 1983). In order to extend the RNA-4(F2) sequence beyond this point a second cloning experiment was performed starting with the same preparation of BNYVV RNA. The insert of pBF15, the longest RNA-4 cDNA clone obtained, was sequenced from its 5' terminus to nucleotide 1149. Where they overlapped, pBF15 and pBF4 were identical except that the 5'-terminal inverted repeat of pBF4 was absent from the longer pBF15, consistent with the idea that the rearranged sequence is indeed a cloning artefact. The insert of pBF15 extends beyond the 5' terminus of pBF4 by 413 base pairs. It also has an inverted repeat, of 11 base pairs, at its 5' terminus. This inverted portion of pBF15 is likely to be a cloning artefact and is not included in the RNA-4(F2) sequence presented in Fig. 3. The RNA-4(F2) clone is 61 residues shorter than the RNA-4(G1) clone at its 5' terminus, perhaps because of premature termination of cDNA synthesis. In view of their overall sequence homology we assume that the 5'-terminal sequence of RNA-4(F2) will be closely similar to RNA-4(G1), but this remains to be shown by direct RNA sequence analysis.

As in the case of RNA-3, the sequences of the RNA-4 clones from the two isolates are, with one exception, co-linear with over 97% sequence homology (17 base changes with 14 transitions and three transversions). Again, as in the case of RNA-3, the most significant difference takes the form of an I/D element 324 nucleotides in length which is present in RNA-4(F2) but absent from RNA-4(G1) (Fig. 3). No obvious sequence homology exists between the I/D elements of RNA-3 and RNA-4.

The RNA-4(F2) sequence contains an open reading frame extending from an AUG at nucleotide 344 to a UAG at nucleotide 1192 which can encode a 31K polypeptide (Fig. 2), in good agreement with the size of the in vitro translation product of RNA-4(F2) as measured by PAGE (Richards et al., 1985). The I/D element falls within this open reading frame so that the translation product of RNA-4(G1) should be shorter by 108 amino acids (Fig. 2).

Sequence analysis of other RNA-4(G1) clones

In order to obtain information about variability within the BNYVV RNA population, three additional clones of RNA-4(G1) were partially sequenced (Table 1). The sequence of these clones was strictly identical to that of pBA3 except in the vicinity of the deletion point for the I/D element. For clones pBA7 and pBA401, the deletion is 19 base pairs shorter, being 305 base pairs rather than 324 base pairs long (Fig. 4). Consequently, the open reading frame in RNA-4 is prematurely closed at the deletion point in pBA7 and pBA401 and we predict that translation of an RNA molecule corresponding to this class of cDNA insert should yield a polypeptide of 14K rather than the 19K polypeptide predicted for pBA3. In clone pBA5 the boundaries of the deletion are the same as in pBA7 and pBA401 but there is an additional single nucleotide deletion at position 724 which compensates for the shift in reading frame introduced by the longer deletion (Fig. 4). It is also noteworthy that the residues immediately flanking the deletion in pBA5, pBA7 and pBA401 are mutated with respect to the pBF4 sequence. This is not the case, however, for the residues bounding the deletion in pBA3 (Fig. 4).
This sequence variation of the cDNA inserts undoubtedly reflects sequence heterogeneity in the RNA-4(G1) population. The G1 inoculum used in this study was derived from a single local lesion but it has been shown that such a practice does not ensure homogeneity of a virus population (Garcia-Arenal et al., 1984). The vicinity of the I/D site in RNA-4(G1) appears to be a mutational hot spot as all sequence variation so far detected is confined to this region. Furthermore, the fact that there is variation within the RNA-4(G1) population as to the exact boundaries of the sequence eliminated with respect to RNA-4(F2) argues strongly that the RNA length differences observed between the two isolates result from deletions within the longer RNA rather than insertions into a shorter ancestral sequence. BNYVV is transmitted by the soil-borne fungus *Polymyxa betae* (Tamada, 1975) and it is therefore of interest that soil-borne wheat mosaic virus, which is also transmitted by a fungus of the genus *Polymyxa*, can apparently undergo spontaneous deletions of as much as 1000 nucleotides of its RNA II component (Shirako & Brakke, 1984).

**Sequence homology between RNA-3 and RNA-4**

Hybridization experiments detected cross-reaction between the cDNA clones of RNA-3 and RNA-4, indicating that these two sequences share elements of sequence homology (Richards et al., 1985). Fig. 5 shows a computer-generated matrix (Maizel & Lenk, 1981) comparing sequences of RNA-3 and RNA-4 from isolate F2. The matrix exhibits a high degree of sequence homology for the 3'-terminal 200 residues of the two RNAs. The homologous region consists of three blocks of almost perfect homology separated by short gaps or dissimilar sequences (Fig. 6). Presumably, these sequences are responsible for functions common to the two RNA molecules such as initiation of replication or recognition of viral coat protein during initiation of virus assembly.

No other extended regions of sequence homology were detected between RNA-3 and RNA-4 from isolate F2 at the level of resolution used for matrix analysis but visual inspection of the sequences has revealed short highly homologous regions near the boundaries of the I/D elements (Fig. 7). The deletions which give rise to the shortened RNA-3 and RNA-4 of isolate G1 occur near or within these sequences. It therefore seems reasonable to conclude that the mechanism engendering the deletions is similar for RNA-3 and RNA-4 and that the conserved sequences play a role in the process. One possibility is that the deleted regions are eliminated by splicing but there are no sequences resembling the consensus splice junction donor and acceptor sites (Mount, 1982) flanking the deletions. In the case of RNA-4(F2) there is a sequence GTGAAAGCTAA (nucleotides 723 to 732) near the left-hand boundary of the I/D element which is repeated with one mismatch at the right-hand boundary (nucleotides 1048 to 1055). In clone pBA3 the deletion has taken place in such a manner that one of the two repeats is partially eliminated (Fig. 3). Deletions of this sort involving homologous regions could arise during replication if the growing RNA chain switches from one such repeat to the other with the intervening RNA looping out of the replicative form to be subsequently eliminated by an unknown process. Such a mechanism, however, cannot readily account for the deletion in RNA-3, which has no such sequence repeats flanking the I/D element, nor can it easily be reconciled with the existence of the variant RNA-4(G1) sequences (i.e. pBA5 and pBA7) in which the deletion does not neatly span the sequence repeats.

Computer-assisted comparison of the RNA-3 and RNA-4 sequences with those of RNA from other plant viruses has revealed several points of homology. The most notable example is a 23 nucleotide block of BNYVV RNA-3 (nucleotides 1235 to 1257) which is also present in RNA-3 of cucumber mosaic virus (nucleotides 1889 to 1911) (Gould & Symons, 1982). In the case of BNYVV RNA-4, there is a block near the 3' terminus (nucleotides 1360 to 1376) which matches at 15 of 17 positions with a 3'-terminal portion of alfalfa mosaic virus RNA-1 (nucleotides 3522 to 3538) (Cornelissen et al., 1983) and a 10 residue match (nucleotides 710 to 719) preceded by shorter blocks of homology with solanum nodiflorum mottle virus and velvet tobacco mottle virus RNA-2 (nucleotides 66 to 75) (Haseloff & Symons, 1982). The evolutionary and/or functional implications of these similarities are not yet known.
Fig. 5. Matrix displaying sequence homology between BNYVV RNA-3 (vertical axis) and RNA-4 (horizontal axis) of isolate F2. For purposes of comparison we have assumed that the first 61 nucleotides of RNA-4(F2) are identical in sequence to the corresponding portion of RNA-4(G1). All possible blocks of 21 residues are compared pairwise and a dot is put on the matrix at the coordinates corresponding to the midpoint of each block if 14 or more residues of the two sequences are identical.

If RNA-3 and RNA-4 share a common ancestor they have diverged to a remarkable extent over much of their sequence while still conserving elements of homology at the 3' terminus and at the I/D element junctions. A better understanding of the relationship between RNA-3 and RNA-4, however, must await more knowledge about the function of these RNA species in the BNYVV infection process, in particular knowing whether or not RNA-3 and RNA-4 are satellite RNAs. In view of the location of the I/D elements it appears unlikely that the long open reading frames of RNA-3 and RNA-4 encode essential functions unless these functions lie exclusively within the N-terminal portion of the corresponding polypeptide. Furthermore, examination of the various point mutations occurring within the open reading frame of RNA-3 of the F2 and G1 isolates does not reveal a pattern characteristic of a functional coding region:
**Fig. 6.** Alignment of the 3'-terminal region of cDNA sequences corresponding to RNA-3 and RNA-4 of isolate F2. Regions of extensive homology are underlined. The alignment was generated using the program GAP of Deveraux et al. (1984) with a gap weight of 5 and a gap length weight of 0.3.

**Fig. 7.** Conserved sequences flanking the I/D elements of BNYVV RNA-3 and RNA-4 of isolate F2. The 5' and 3' boundaries of the deletion in RNA-3(G1) are indicated by the inverted triangles above the sequences. The 5' and 3' boundaries of the deletion in RNA-4(G1) cDNA clones pBA7 and pBA3 are indicated by the filled and hollow triangles, respectively, beneath the sequences. The homologous sequences referred to in the text are underlined.

In RNA-3 only four of 11 point mutations are conservative with respect to amino acid substitution while in RNA-4, six of 10 are conservative. On the other hand, it is remarkable that, with the exception of clone pBA7, elimination of the I/D elements in both RNA-3 and RNA-4 has occurred in such a way as to conserve the open reading frame.

The skilful technical assistance of Danièle Beck is gratefully acknowledged and we thank Karl Gordon for helpful discussions. S.B. is the recipient of a grant from the Institut Technique de la Betterave Industrielle. Additional financial support was provided by the CNRS.

**REFERENCES**


(Received 7 January 1985)