

De novo generation of cymbidium ringspot virus defective interfering RNA

József Burgyan,† Luisa Rubino and Marcello Russo*

Dipartimento di Protezione delle Piante dalle Malattie, Università degli Studi and Centro di Studio del CNR sui Virus e le Virosi delle Colture Mediterranee, Via Amendola 165/A, 70126 Bari, Italy

Nicotiana clevelandii plants were inoculated with cymbidium ringspot tomosvirus RNA synthesized *in vitro*, after which further passages were made by sap inoculation. During the third passage, low M_r RNA species appeared which had the characteristics of deletion mutants of genomic RNA. Sequence analysis of several of these defective interfering RNAs sug-

gested a possible evolution of smaller from larger molecules. Computer-generated secondary structures of sequences surrounding recombination sites were extensive and stable and these sites occurred in interior or hairpin loops, thus providing a possible explanation for discontinuous RNA transcription and the formation of deletions in genomic RNA.

Introduction

Repeated serial high multiplicity passages of several animal viruses can lead to the generation of defective interfering (DI) particles which interfere with the replication of the helper genome and thereby diminish cytopathic effects, leading to the establishment of persistent infections (Perrault, 1981; Schlesinger, 1988).

Among plant viruses, two tomosviruses [petunia asteroid mosaic virus, also referred to as the cherry strain of tomato bushy stunt virus (TBSV) (Hillman *et al.*, 1987), and cymbidium ringspot virus (CyRSV) (Burgyan *et al.*, 1989)] and one carmovirus [turnip crinkle virus (TCV) (Li *et al.*, 1989)] have been shown to support the replication of DI RNA. Attempts to determine the origin of DI RNAs have been made by making high multiplicity passages of a DI RNA-free isolate of TBSV cherry (Morris *et al.*, 1988) or by inoculating plants with purified genomic RNA from CyRSV (Burgyan *et al.*, 1989). The results of these experiments provide circumstantial evidence for *de novo* generation of DI RNA from the parental virus genome. With TCV, the use of cloned inocula provided experimental evidence for *de novo* generation (Li *et al.*, 1989).

The genome of CyRSV is a positive-sense ssRNA which is 4733 nucleotides (nt) long and encapsidated in icosahedral particles (Grieco *et al.*, 1989a; Martelli *et al.*, 1989). In addition to genomic RNA, CyRSV particles encapsidate two subgenomic RNAs of 2118 nt and 936 nt, both of which are 3'-coterminal with genomic RNA.

The larger of these RNAs encodes the capsid protein (M_r 41 000; 41K); the smaller has two nested genes which encode two polypeptides with predicted M_r s of 22K and 19K (Grieco *et al.*, 1989b). CyRSV also supports the replication of other linear RNAs, either completely derived from genomic RNA (DI RNA; Burgyan *et al.*, 1989) or essentially unrelated to it (satellite RNA; Gallitelli & Hull, 1985). Molecules of both types have been cloned and sequenced (Rubino *et al.*, 1990).

The availability of full-length cDNA clones of CyRSV genomic RNA, from which infectious RNA can be synthesized (Burgyan *et al.*, 1990), prompted us to reinvestigate the origin of CyRSV DI RNA. The results of these studies, reported in the present paper, show that the generation of DI RNA does not require the presence of DI RNA-length template but is probably generated from genome-length RNA by a series of progressive deletions.

Methods

Virus RNA and plant inoculation. Infectious CyRSV RNA was synthesized *in vitro* from full-length cDNA clones and inoculated into *Nicotiana clevelandii* plants as previously described (Burgyan *et al.*, 1990). Infected tissues were ground in approximately the same volume of high pH inoculation buffer containing bentonite (Heaton *et al.*, 1989) between 8 and 10 days later and inoculated into *N. clevelandii* (second passage). Two further subinoculations were done in the same way.

Full-length cloning and sequencing of DI RNAs. Total RNA was extracted from one leaf of infected plants and electrophoresed in 1.2% low melting point agarose (LMP; BRL) as described (Burgyan & Russo, 1988; Russo *et al.*, 1988). RNA species migrating in positions expected for molecules approximately 400 to 700 nt long were eluted

† Permanent address: Agricultural Biotechnology Center, P.O. Box 170, H-2101 Gödöllő, Hungary.

(Maniatis *et al.*, 1982) and resuspended in 10 µl sterile deionized water. These RNAs were denatured with 1 µl 100 mM-methylmercuric hydroxide, incubated for 10 min at room temperature and then mixed with 2 µl 700 mM-2-mercaptoethanol and 40 units RNase inhibitor (HPRI; Amersham); cloning was essentially as described by Ahlquist & Janda (1984). First strand synthesis was primed with the oligonucleotide 5' CCCCTGCATTGCTGCAA 3', which is complementary to the last 17 nucleotides of the CyRSV genomic RNA sequence; second strand synthesis was primed with the oligonucleotide 5' GGAAATCCTCCAGGACA 3', which is homologous to the first 17 nucleotides of CyRSV RNA. (The first base is unknown but it is assumed to be G; Burgyan *et al.*, 1990.) Another strategy, which was used to ensure that scarce DI RNA species did not escape detection, was to use 10 µg total RNA extract as the template for the first strand synthesis and then to amplify cDNA by 35 cycles of the polymerase chain reaction (PCR) with the two primers described before and AmpliTaq polymerase (GeneAmp DNA Amplification Reagent Kit; Perkin Elmer Cetus). Melting, annealing and polymerization times and temperatures were 1 min at 94 °C, 1 min at 37 °C and 3 min at 72 °C, respectively. The resulting dsDNA was cloned in pUC18 digested with *Sma*I.

Recombinant plasmids were extracted by the boiling method (Maniatis *et al.*, 1982) and the sizes of the inserts were estimated after restriction with *Hind*III and *Eco*RI. For sequencing, plasmids were extracted according to the method of Hattori & Sakaki (1986) and sequenced with T7 DNA polymerase (Sequenase; U.S. Biochemicals). Sequence data were analysed with the DNA Strider program (Marck, 1988); secondary structures were constructed according to the program of Zucker & Stiegler (1981).

Northern blots. RNA samples were denatured with formamide and formaldehyde and electrophoresed in 1.2% agarose gels in MOPS buffer (20 mM-MOPS, 5 mM-sodium acetate, 1 mM-EDTA) containing 2.2 M-formaldehyde (Maniatis *et al.*, 1982) and blotted onto Hybond membranes (Amersham). Blots were hybridized with a nick-translated DI RNA clone designated pDI4 (Rubino *et al.*, 1990).

Results

De novo generation of small RNA species

Northern blots of RNA preparations from two *N. clevelandii* plants inoculated with *in vitro* transcripts of CyRSV RNA showed the presence of genomic and subgenomic RNAs but not of low *M_r* RNA species (Fig. 1, lane 3), as found previously by Burgyan *et al.* (1990). The same pattern was obtained with RNA extracts from sap inoculated plants at the second passage (not shown). However, RNA preparations extracted from plants infected at the third passage contained, in addition to genomic and subgenomic RNAs, a number of RNA species migrating in the range of approximately 400 to 700 nt (Fig. 1, lanes 1 and 2). In samples collected a day later from younger leaves of the same plant, two different electrophoretic patterns were found (Fig. 1, lanes 4 and 5). The numbers and relative quantities of RNA species smaller than subgenomic RNA 2 found in different leaf samples varied. Nevertheless these small RNAs were present in all samples from the third passage onwards

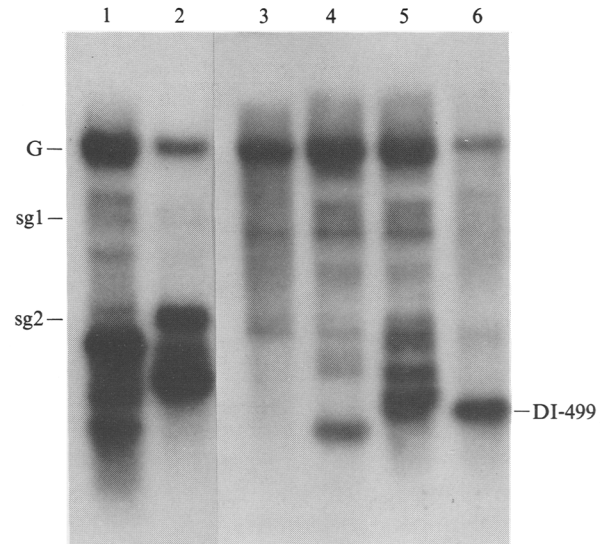


Fig. 1. Northern blots of RNA preparations from a plant infected with CyRSV RNA *in vitro* transcripts (lane 3) and plants inoculated with sap at the third passage, 7 (lanes 1 and 2) and 8 (lanes 4 and 5) days after inoculation. Lane 6 shows an RNA preparation containing a DI RNA (DI-499) of 499 nt previously described (Rubino *et al.*, 1990). G (4733 nt), sg1 (2118 nt) and sg2 (936 nt) are positions of CyRSV genomic and subgenomic RNAs, respectively, in lane 3.

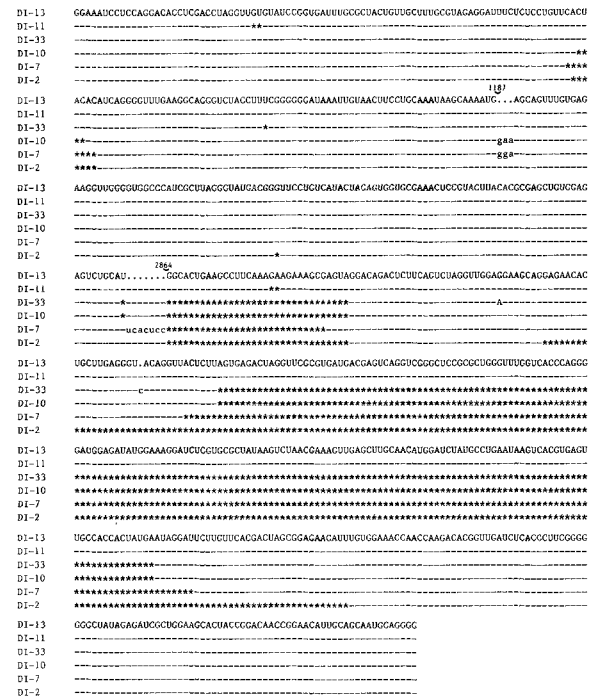


Fig. 2. Nucleotide sequence of DI RNAs generated *de novo*. The complete sequence is reported for DI-13. In the other DI RNAs, dashed lines indicate conserved nucleotides; stars are deletions; lower-case letters are insertions in positions indicated by dots in DI-13; upper case letters are substitutions. Numbers marked indicate the large deletions in genomic RNA.

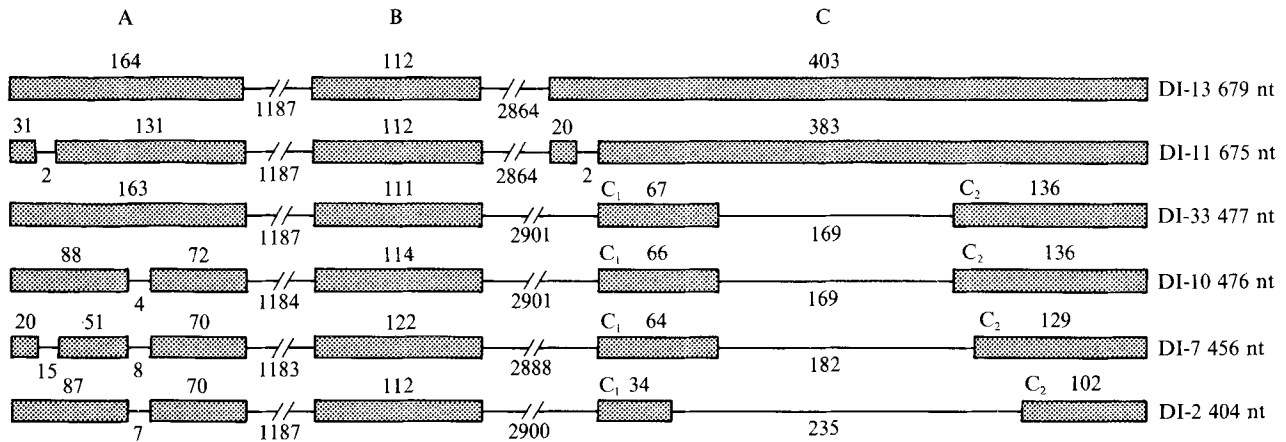


Fig. 3. Conserved sequence blocks in DI RNAs generated *de novo*. Numbers above shaded areas are nucleotides present in conserved blocks; numbers below lines are deletions in genomic RNA.

(patterns of the fourth passage are not shown), supporting the idea that these RNA species were generated *de novo*.

Cloning and sequencing of DI RNAs

The RNA preparations shown in Fig. 1 (lanes 1, 2, 4 and 5) were used for cloning. Clones obtained by two procedures (cloning from RNA eluted from gels and amplification by PCR from cDNA made on unfractionated RNA templates) had inserts with sizes ranging from about 400 bp to 700 bp. Sixteen full-length clones, containing both primer sequences used for cloning, were sequenced (four or five for each size class of 400, 500 and 700 bp). As shown in Fig. 2 and 3, there were six types of molecule of identical sequence. The longest DI RNA (DI-13, Fig. 3) had 679 nucleotides and was made up of three blocks (A, B and C) entirely derived from CyRSV genomic RNA (Grieco *et al.*, 1989a). Block A is composed of the first 164 nucleotides of genomic RNA including the 5' leader sequence and the initiation codon of the 92K protein [open reading frame (ORF) 1]; block B is 112 nt long and represents the central part of the putative polymerase gene (ORF 2); block C is 403 nt long and corresponds to 49 nucleotides of the carboxyl terminus of the 22K protein gene (ORF 4) and the entire non-coding region of 354 nucleotides. The constitution of DI-13 RNA was the simplest of the DI RNAs generated *de novo* as it represents a linear deletion mutant of genomic RNA with two long deletions.

Sequence analysis of the smaller RNAs showed that they are formed by essentially the same genomic RNA sequence blocks with a progressive reduction in size caused by the introduction of further deletions inside

Block A	<u>G G A A A U C C U C C A G G A</u>	
Block B	<u>G A A A G C A G U U U G U G A</u>	DI-10
Block B	<u>G G A A G C A G U U U G U G A</u>	DI-7
Block B	<u>A G C A G U U U G U G A G A A</u>	DI-13, -11, -33, -2
Block C	<u>G G C A C U G A A G C C U U C</u>	DI-13, -11
Block C ₁	<u>G G A C A G A C U C U U C A G</u>	DI-33, -10, -2
Block C ₁	<u>A G U A G G A C A G A C U C U</u>	DI-7
Block C ₂	<u>A U A G G A U U C U U G U U C</u>	DI-33, -10
Block C ₂	<u>U C U U G U U C A C G A C U A</u>	DI-7
Block C ₂	<u>U G U G G A A A C C A A C C A</u>	DI-2

Fig. 4. Alignment of the first 15 nucleotides of conserved blocks of DI RNAs formed *de novo*.

blocks A and C of DI-13 RNA. As compared with DI-13, block A was almost identical in DI-33, except for a single base deletion (not shown in Fig. 3), and had one or two short deletions in DI-11, -10, -7 and -2. The termini of block A were conserved in all six types of molecule. Block B was the least affected; it had no deletions and only minor variations at its extremities (Fig. 2). Block C showed the greatest changes; there was progressive reduction in its size due to deletions in the central part and at the 5' end (Fig. 2 and 3).

Alignment of the first 15 nucleotides of each block of DI RNAs did not show a clear consensus sequence (Fig. 4) but the initiation sequences showed marked similarities. In particular, seven out of 10 had one or two Gs in the first two positions, followed by between one and three As in the next three positions; nine out of 10 had one or two pyrimidines in positions 10 and 11.

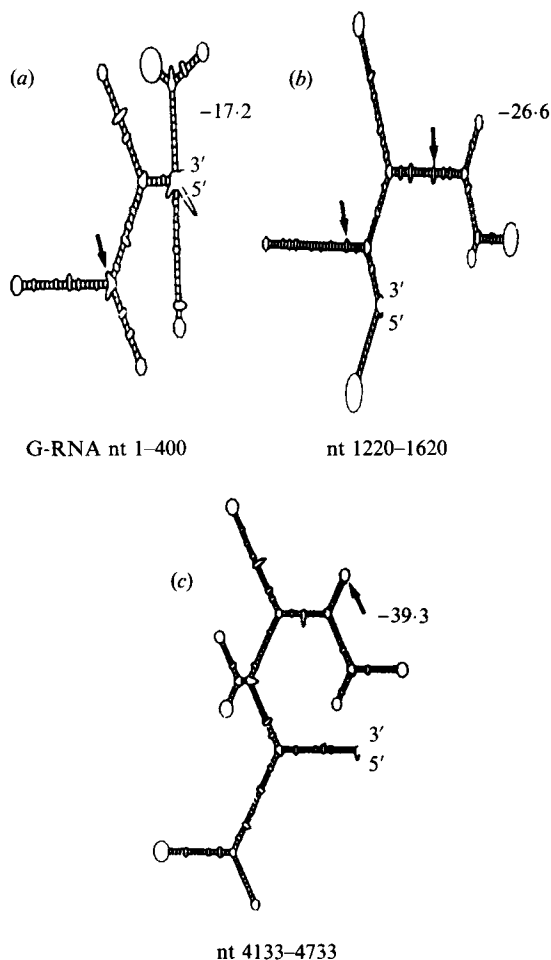


Fig. 5. Computer-generated secondary structure of sequences surrounding recombination sites in CyRSV genomic RNA (a to c). Free energy is indicated in kJ/mol. Arrows point to recombination sites.

Sites of DI RNA recombination

The most stable predicted secondary structures of the sequences surrounding the deleted regions of genomic RNA in DI-13 RNA were obtained using the algorithm of Zucker & Stiegler (1981). As shown in Fig. 5(a to c), these regions can be folded to form extensive secondary structures with the four rearrangement sites located within single-stranded sequences in interior or hairpin loops. Folding the entire DI-13 RNA molecule also generated a stable (ΔG , -41.9 kJ/mol) secondary structure in which recombination sites possibly involved in the formation of DI-2 RNA were found in unpaired sequences (not shown).

Discussion

DI RNAs of CyRSV generated *de novo* have essentially the same characteristics as DI RNA present with the

virus in the initial inoculum (Rubino *et al.*, 1990), i.e. the sequence is entirely derived from genomic RNA, there are no extra or repeated sequences and both 5' and 3' termini are conserved. In previous work we have found only one species of DI RNA, either because only the species contained in the inoculum was replicated (Rubino *et al.*, 1990) or because, when purified genomic RNA was used as inoculum followed by successive subinoculations using sap, other species escaped attention (Burgyan *et al.*, 1989). In this context, time of sampling and/or leaf position may be very important; for instance, the pattern of the sample of Fig. 1, lane 4 would indicate the presence of essentially one DI RNA, whereas a sample collected from different leaves of the same plant the day before contained several DI RNA species (Fig. 1, lane 1). The size distribution of novel DI RNAs thus seems to be a time-dependent event; DI RNA molecules that form first are longer than those formed later. The smaller molecules seem to be a stable form in the size range of 400 to 500 nt which is maintained in further passages.

The hypothesis that smaller DI RNAs originate from larger precursors is supported by the observation that smaller molecules are nearly always contained entirely within the larger molecules with no extra sequences. The only two exceptions are 10 and three nucleotides, at one extremity of DI-7 and DI-10, respectively, that are not present in DI-13. However, this could be due to the variability of the population of RNA precursors, which are similar but not identical to DI-13. In this connection it is worth noting that the presence of all the oligonucleotides of smaller DI RNAs of Semliki Forest virus in larger DI RNAs was taken as evidence for the evolution of small from large DI RNAs in high multiplicity passages (Stark & Kennedy, 1978). Another possibility is that DI RNAs of different size are generated simultaneously from the helper genome and that some molecules become predominant.

As pointed out by Cascone *et al.* (1990), the same mechanism may be operating in RNA recombination and generation of DI RNA. In the 'copy choice' mechanism (Lazzarini *et al.*, 1981) the RNA polymerase may stop before reaching the 5' end of the template RNA and jump, together with nascent complementary RNA, to another site on a different or the same template. In the first case, RNA recombination would occur; in the second, deletions would be introduced in template RNA to form DI RNA. Romanova *et al.* (1986) developed a model explaining the formation of poliovirus RNA recombinants. They observed that crossovers do not take place at random but in single-stranded sequences or stretches with weak secondary structures. Such preferential sites of recombination are surrounded by secondary structures with a negative free energy, which constrains

the action of RNA polymerase. Occasionally the enzyme is unable to overcome such regions and it separates from the original template RNA and may resume synthesis at another template site. The conditions for this model to operate (i.e. the presence of stable secondary structures surrounding single-stranded loops) seem to exist in the genomic RNA structure of CyRSV. Resumption of RNA synthesis by RNA polymerase also requires the presence of initiation signals with similar nucleotide sequences in different parts of the RNA. As shown in Fig. 4, many recombination sites start with G followed by A, which is part of the consensus recognized in all replicating molecules of the CyRSV system, i.e. genomic, subgenomic and satellite RNAs (Rubino *et al.*, 1990). A model based on discontinuous action of the RNA polymerase has been proposed to explain the generation of DI RNA from the mouse hepatitis coronavirus genome (Makino *et al.*, 1988). Likewise, dissociation of RNA polymerase from the template due to the presence of a hairpin structure is at the base of the model presented by Cascone *et al.* (1990) for the generation of TCV DI RNA and RNA recombinants.

We thank Dr P. D. Nagy for help given in the initial part of this work, T. Dalmay for the generation of secondary structures and Professor G. P. Martelli for helpful discussions and revision of the manuscript.

References

- AHLQUIST, P. & JANDA, M. (1984). cDNA cloning and *in vitro* transcription of the complete brome mosaic virus genome. *Molecular and Cellular Biology* **4**, 2876–2882.
- BURGYAN, J. & RUSSO, M. (1988). Studies on the replication of a satellite RNA associated with cymbidium ringspot virus. *Journal of General Virology* **69**, 3089–3092.
- BURGYAN, J., GRIECO, F. & RUSSO, M. (1989). A defective interfering RNA molecule in cymbidium ringspot virus infections. *Journal of General Virology* **70**, 235–239.
- BURGYAN, J., NAGY, P. D. & RUSSO, M. (1990). Synthesis of infectious RNA from full-length cDNA to RNA of cymbidium ringspot tomosvirus. *Journal of General Virology* **71**, 1857–1860.
- CASCONI, P. J., CARPENTER, C. D., LI, X. H. & SIMON, A. E. (1990). Recombination between satellite RNAs of turnip crinkle virus. *EMBO Journal* **9**, 1709–1715.
- GALLITELLI, D. & HULL, R. (1985). Characterization of satellite RNAs associated with tomato bushy stunt virus and five other definitive tomosviruses. *Journal of General Virology* **66**, 1533–1543.
- GRIECO, F., BURGYAN, J. & RUSSO, M. (1989*a*). The nucleotide sequence of cymbidium ringspot virus RNA. *Nucleic Acids Research* **17**, 6383.
- GRIECO, F., BURGYAN, J. & RUSSO, M. (1989*b*). Nucleotide sequence of the 3'-terminal region of cymbidium ringspot virus RNA. *Journal of General Virology* **70**, 2533–2538.
- HATTORI, M. & SAKAKI, Y. (1986). Dideoxy sequencing method using denatured plasmid templates. *Analytical Biochemistry* **152**, 232–238.
- HEATON, L. A., CARRINGTON, J. C. & MORRIS, T. J. (1989). Turnip crinkle virus infection from RNA synthesized *in vitro*. *Virology* **170**, 214–218.
- HILLMAN, B. I., CARRINGTON, J. C. & MORRIS, T. J. (1987). A defective interfering RNA that contains a mosaic of plant virus genome. *Cell* **51**, 427–433.
- LAZZARINI, R. A., KEENE, J. D. & SCHUBERT, M. (1981). The origins of defective interfering particles of the negative-strand RNA viruses. *Cell* **26**, 145–154.
- LI, X. H., HEATON, L. A., MORRIS, T. J. & SIMON, A. E. (1989). Turnip crinkle virus defective interfering RNAs intensify viral symptoms and are generated *de novo*. *Proceedings of the National Academy of Sciences, U.S.A.* **86**, 9173–9177.
- MAKINO, S., SHIEN, C. K., SOE, L. H., BAKER, S. C. & LAI, M. M. C. (1988). Primary structure and translation of a defective interfering RNA of murine coronavirus. *Virology* **166**, 550–560.
- MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory.
- MARCK, C. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Research* **16**, 1829–1836.
- MARTELLI, G. P., RUSSO, M. & GALLITELLI, D. (1989). Tombusvirus group. *AAB Descriptions of Plant Viruses*, no. 352.
- MORRIS, T. J., MULLIN, R., HEARNE, P. & HILLMAN, B. I. (1988). *De novo* generation of tomato bushy stunt defective interfering RNAs. *Phytopathology* **78**, 1601 (Abstract).
- PERRAULT, J. (1981). Origin and replication of defective interfering particles. *Current Topics in Microbiology and Immunology* **93**, 151–207.
- ROMANOVA, L. I., BLINOV, V. M., TOLSKAYA, E. A., VIKTOROVA, E. G., KOLESNIKOVA, M. S., GUSEVA, E. A. & AGOL, V. I. (1986). The primary structure of crossover regions of intertypic poliovirus recombinants: a model of recombination between RNA genomes. *Virology* **155**, 202–213.
- RUBINO, L., BURGYAN, J., GRIECO, F. & RUSSO, M. (1990). Sequence analysis of cymbidium ringspot virus satellite and defective interfering RNAs. *Journal of General Virology* **71**, 1655–1660.
- RUSSO, M., BURGYAN, J., CARRINGTON, J. C., HILLMAN, B. I. & MORRIS, T. J. (1988). Complementary DNA cloning and characterization of cymbidium ringspot virus RNA. *Journal of General Virology* **69**, 401–406.
- SCHLESINGER, S. (1988). The generation and amplification of defective interfering RNAs. In *RNA Genetics*, vol. II, pp. 167–185. Edited by E. Domingo, J. J. Holland & P. Ahlquist. Boca Raton: CRC Press.
- STARK, C. & KENNEDY, S. I. T. (1978). The generation and propagation of defective-interfering particles of Semliki Forest virus in different cell types. *Virology* **89**, 285–299.
- ZUKER, M. & STIEGLER, P. (1981). Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Research* **9**, 133–148.

(Received 7 August 1990; Accepted 13 November 1990)