

Novel Structures of Two Virus-like RNA Elements from a Diseased Isolate of the Dutch Elm Disease Fungus, *Ophiostoma novo-ulmi*

Yiguo Hong,* Thomas E. Cole,* Clive M. Brasier,† and Kenneth W. Buck*¹

*Department of Biology, Imperial College of Science, Technology, and Medicine, London SW7 2BB, United Kingdom; and
†Forest Research Station, Alice Holt Lodge, Farnham, Surrey GU10 4LH, United Kingdom

Received October 17, 1997; returned to author for revision November 21, 1997; accepted December 10, 1997

The nucleotide sequences of 2 of the 10 mitochondrial double-stranded (ds) RNA segments in a diseased isolate, Log1/3-8d² (Ld), of *Ophiostoma novo-ulmi*, RNA-7 (1057 nucleotides) and RNA-10 (317–330 nucleotides), have been determined. Both RNAs are A-U-rich, but in Southern and Northern blots, no hybridization with mitochondrial DNA or RNA could be detected. Only very short open reading frames were found in both RNAs. As most of its sequence is unrelated to any of the other Ld dsRNAs, RNA-7 may be regarded as a satellite RNA. Northern blotting detected a full-length single-stranded (ss) form of RNA-7 in nucleic acid extracts from Ld. The 5′- and 3′-terminal 39 nucleotides of ssRNA-7 are imperfect inverted complementary repeats of each other, which could cause ssRNA-7 to form a panhandle structure. In addition, the 5′-terminal nucleotides 1–28 and 3′-terminal nucleotides 1032–1057 of ssRNA-7 each contained inverted complementary sequences, allowing the possibility for each terminus to form separate stem-loop structures. The combination of these two structural features has not been found previously in any dsRNA or ssRNA virus. RNA-10 was shown to have an unusual structure, consisting of a mosaic of sequences derived from regions of the 5′- and 3′-termini, or just the 5′-terminus, of RNA-7. RNA-10 has a high degree of inverted complementarity, with the potential to be folded into a very stable hairpin structure. A model for the formation of RNA-10 is presented, involving replicase-driven strand switching between (–)-strand and (+)-strand templates during RNA synthesis, followed by utilization of the nascent strand as a primer and template to form a snap-back RNA. © 1998 Academic Press

Key Words: mitochondrial RNA replicon; inverted complementary repeats; satellite RNA; mosaic structure.

INTRODUCTION

An isolate of the Dutch elm disease fungus, *Ophiostoma novo-ulmi*, Log1/3-8d² (hereafter designated Ld), is characterized by unusually slow growth, the formation of unstable amoeboid-like colonies, and reduced spore germination. Transmission experiments with isolates carrying nuclear markers showed that this phenotype is caused by an extranuclear genetic determinant, designated the d²-factor (Brasier, 1986), which is under investigation as a potential biological control agent (Webber, 1993). Isolate Ld contains 10 segments of unencapsidated, virus-like double-stranded RNA (dsRNA), which copurify with the mitochondria and which have estimated sizes (kbp) of 3.49 (RNA-1), 3.03 (RNA-2), 2.69 (RNA-3), 2.43 (RNA-4), 2.33 (RNA-5), 2.21 (RNA-6), 0.95 (RNA-7), 0.92 (RNA-8), 0.48 (RNA-9), and 0.33 (RNA-10) (Rogers *et al.*, 1986, 1987). Transmission of the dsRNA from Ld to healthy isolates of *O. novo-ulmi* results in conversion of the phenotype of the healthy isolates to that of the diseased isolates, including large reductions in the levels of cytochrome oxidase in the mitochondria (Rogers *et al.*, 1987), and

the *de novo* generation of DNA plasmids derived from the mitochondrial DNA (Charter *et al.*, 1993; Abu-Amero *et al.*, 1995). Loss of dsRNA segments 4, 7, and 10 during conidiogenesis or multiplication of the fungus in the yeast phase leads to reversion to the healthy phenotype, suggesting that one or more of these segments constitutes the d²-factor (Rogers *et al.*, 1986, 1987).

Analysis of single-conidial and yeast-phase isolates which have lost various combinations of dsRNA segments, in addition to segments 4, 7, and 10, suggests that the 10 dsRNA segments of *O. novo-ulmi* isolate Ld may be composed of several independent groups of RNA replicons (Rogers *et al.*, 1988). In addition to mixed infections with different viruses, satellite and defective RNAs can contribute to the multiplicity of dsRNA segments in fungi (reviewed by Buck, 1997). In order to start to unravel the complexity of the dsRNA segments in isolate Ld, we have obtained the complete nucleotide sequences of dsRNA segments 7 and 10, and show that both RNAs have unusual sequence arrangements. We further show that RNA-7 has the properties of a satellite RNA and exists in both double-stranded and single-stranded (ss) forms, and that RNA-10 consists of a mosaic of sequences derived from RNA-7.

¹To whom correspondence and reprint requests should be addressed. Fax: +44 171 584 2056. E-mail: k.buck@ic.ac.uk.

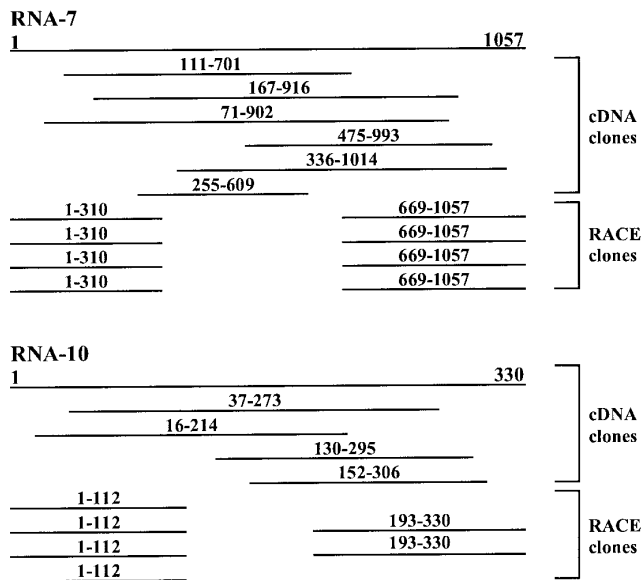


FIG. 1. Maps of cDNA and RACE clones used to construct the nucleotide sequences of RNA-7 and RNA-10. Clones are positioned and numbered according to the sequences in Fig. 2. All the clones shown were obtained from independent cDNA reactions and were sequenced in both directions.

RESULTS

Production of cDNA clones and sequencing dsRNA-7 and dsRNA-10

The ten dsRNA segments from *O. novo-ulmi* isolate Ld were separated by agarose gel electrophoresis and bands corresponding to segments 7 and 10 were extracted from the gel. The dsRNAs were denatured and first-strand cDNA was synthesized using reverse transcriptase and random hexanucleotide primers. After second-strand synthesis, the double-stranded DNA was cloned into the plasmid vector LITMUS 28. The identity of clones was confirmed by using them as probes in Northern blots of dsRNA from *O. novo-ulmi* isolates Ld, W2tol1, and sci21. W2tol1 contains dsRNA segments 1 and 8 (Rogers *et al.*, 1986), whereas sci21, a single conidial isolate derived from Ld, was shown by polyacrylamide gel electrophoresis with ethidium bromide staining to contain 7 of the original 10 dsRNAs and to lack segments 4, 7, and 10 (not shown), similar to several other single conidial isolates previously described (Rogers *et al.*, 1986). Bands in the position of dsRNA-7 and dsRNA-10 were detected in blots of dsRNA from Ld, but not from W2tol1 or sci21 (not shown). The terminal sequences were determined by 5' RACE. Because the RNA templates were double-stranded, 5' RACE was used to determine the 5'-termini of both strands and the 3'-termini were taken as the complements of the 5'-termini.

The cDNA clones used to obtain the complete sequences of dsRNA-7 and dsRNA-10 are illustrated in Fig. 1. Sequences were confirmed from two to eight indepen-

dent clones. The nucleotide sequences of one strand of dsRNA-7 and dsRNA-10 are shown in Fig. 2.

Analysis of the nucleotide sequences of dsRNA-7 and dsRNA-10

The nucleotide sequences of dsRNA-7 and dsRNA-10 were examined for the presence of open reading frames (ORFs) on both strands. Because these RNAs are associated with the mitochondria (Rogers *et al.*, 1987), the sequence UGA was taken to encode tryptophan as in other fungal mitochondria (Osawa *et al.*, 1992; Paquin *et al.*, 1997). The longest ORF on either strand of RNA 7 could encode a protein of only 30 amino acids. If putative introns containing 5' GU and 3' AG sequences were removed, an ORF encoding a protein of 135 amino acids could be created. However, no putative spliced RNAs of the calculated size were detected using RT-PCR or Northern blotting (data not shown). Similar analysis indicated that RNA-10 could not encode a protein of more than 41 amino acids.

Nucleotide polymorphisms were detected at 16 positions out of the 1057 nucleotides of RNA-7, most of which were A-G or C-U transitions (Fig. 2A). DsRNA-10 has more sequence heterogeneity than dsRNA-7 with polymorphisms at 34 positions, including nucleotide transitions, transversions, and deletions/insertions, with an overall size range of one strand from 317–330 nucleotides (Fig. 2B). The frequencies of polymorphisms are much greater than would be expected from errors in cDNA synthesis by reverse transcriptase or in PCR reactions with *Taq* DNA polymerase (Goodman, 1995) and probably largely represent sequence heterogeneity in the populations of RNA-7 and RNA-10. RNA polymerases that replicate RNA generally lack proofreading activities (Domingo and Holland, 1994) and viral RNA molecules may consist of populations of closely related molecules or "quasi-species" (Eigen and Biebricher, 1988). Heterogeneity in populations of dsRNA molecules has also been described for the W dsRNA of *Saccharomyces cerevisiae* (Rodriguez-Cousino *et al.*, 1991), the dsRNA of *Leishmania* RNA virus 1 (Stuart *et al.*, 1992) and a dsRNA replicon in rice (Moriyama *et al.*, 1995). The opportunity for the accumulation of mutations in populations of *O. novo-ulmi* RNA-7 and RNA-10 may be enhanced if these RNAs do not encode protein and require only the retention of cis-acting signals for RNA replication.

The 5'-terminal nucleotides 1–28 of RNA-7 contain complementary sequences which, using the program FOLDRNA, can be folded to form a stem-loop structure (ΔG –9.3 kcal/mol; Fig. 3A). A similar stem-loop structure could be formed by the 3'-terminal nucleotides 1032–1057 (ΔG –8.5 kcal/mol; Fig. 3B). The 5'-terminal nucleotides 1–13 (CUUUAUCCUUU) and the 3'-terminal nucleotides 1045–1057 (AAAGGAUAAUAG) are also completely complementary to each other when read in



FIG. 2. Nucleotide sequences of the (+)-strand of dsRNA-7 (A) and dsRNA-10 (B). Nucleotides are numbered from the 5'-end. Conserved nucleotides at positions where polymorphisms occurred are indicated in italics, and less conserved nucleotides are indicated underneath. Asterisks (*) indicate positions where a nucleotide was absent in some clones. The positions of the specific nested forward-strand primers used to obtain the 5'-RACE clones of the (-) strand are underlined, and the complementary sequences to these primers were used for the 5'-RACE clones of the (+)-strand. The polarity of strands was determined from the Northern blot in Fig. 5. The RNA-10 sequence is shown as the strand with the highest sequence identity to the (+)-strand of RNA-7.

opposite directions. These sequences, together with further adjacent complementary sequences, could result in the formation of a panhandle structure (Fig. 3C). The free energy of this structure, calculated by adding to the stem a loop with the same sequence as that in Fig. 3A or 3B, is -34.8 or -32.2 kcal/mol, respectively. The panhandle structure was also formed when the whole of RNA-7 was folded (not shown). It is also noteworthy that the 3'-terminal nucleotides 1032-1057 constitute a partial direct repeat (88.5% sequence identity) of nucleotides 1-28.

The 5'-terminal nucleotides 1-24 of RNA-10 are repeated three times with two or three mismatches from nt 98-118, 217-240, and 306-329 (Fig. 4). Each repeat contains complementary sequences which could be folded to form stem-loop structures similar to those shown for RNA-7 in Figs. 3 A and 3B. In addition the 5'-terminal nucleotides 1-24 and the 3'-terminal nucleotides 306-329 are fully complementary to each other when read in opposite directions. A high degree of complementarity (in opposite directions) extends over most of the sequence, enabling a very stable loop-back structure to be formed (ΔG -162.3 kcal/mol; Fig. 5). Furthermore, the full-length sequences of the two complementary strands of segment 10 are 80.5% identical when both are compared in a 5'-3' direction.

Molecular relatedness of RNA-7 and RNA-10

In Northern blots of dsRNA from isolate Ld, a cDNA probe from RNA-7 hybridized, not only with dsRNA-7, but also with dsRNA-10 (Fig. 6, lane Ld). The hybridization with dsRNA-10 was less intense than that with dsRNA-7, but it should be noted that the RNA-7 probe covered only nucleotides 71-902. Similarly, RNA-10 cDNA probes could detect both dsRNA-10 and dsRNA-7, with the intensity of hybridization with RNA-7 being less than that with RNA-10 (not shown). These results suggested that RNA-10 may be related to the 5'- and /or 3'-ends of RNA-7.

Direct sequence comparisons confirmed that RNA-10 could be derived almost completely from nucleotides 2-86 of RNA-7. Nucleotides 1-85 of RNA-10 constitute a partial direct repeat of nucleotides 3-86 of RNA-7 (96.3% sequence identity). The remainder of RNA-10 is made up of a mosaic of sequences derived from regions of RNA-7 nucleotides 2-86. Figure 7 depicts a scheme for the origins of RNA-10, which gives the highest sequence identity over the largest number of nucleotides. However, because of the high degree of self-complementarity of RNA-10, many regions of RNA-10 have high degrees of sequence identity with both strands of RNA-7. Further-

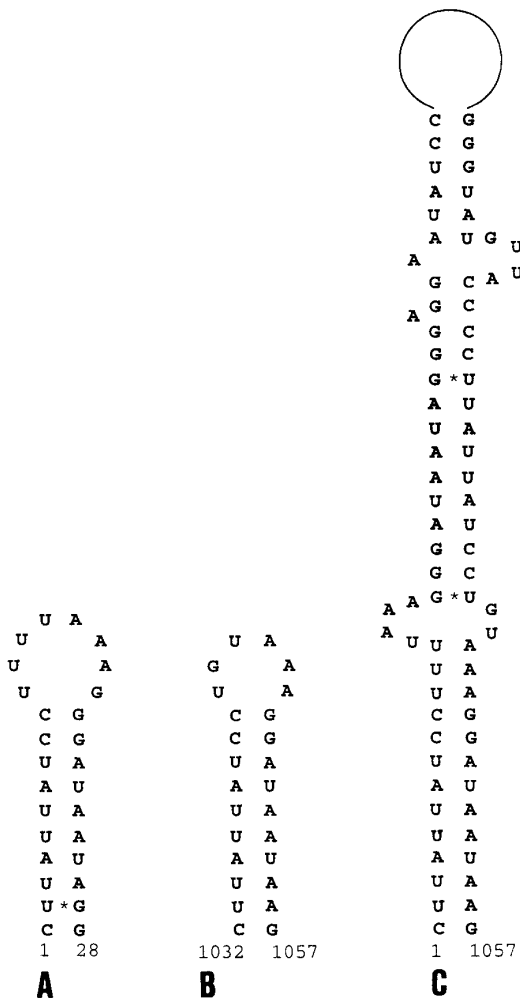


FIG. 3. Potential secondary structures involving the 5'- and 3'-terminal sequences of the (+)-strand of RNA-7. (A) 5' stem-loop structure. (B) 3' stem-loop structure. (C) Panhandle structure involving complementary base-pairing of the 5'- and 3'-terminal sequences.

more, because of the four repeats in RNA-10 (Fig. 4) and the fact that the 3' sequence of RNA-7 can be regarded as either a direct repeat of, or complementary to, the 5' sequence of RNA-7, there are several other possibilities for the derivation of RNA 10 from RNA-7. In particular, it is noteworthy that the 3'-terminal nucleotides 305–329 of

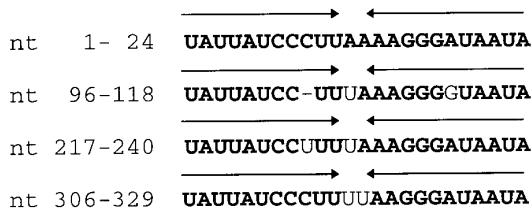


FIG. 4. Repeated sequences in RNA-10. Each repeat has inverted complementary sequences indicated by the arrows. Nucleotides identical or nonidentical to those in the corresponding position of the first repeat are shown in upper- or lowercase respectively. (—) indicates a gap used to align the sequences.

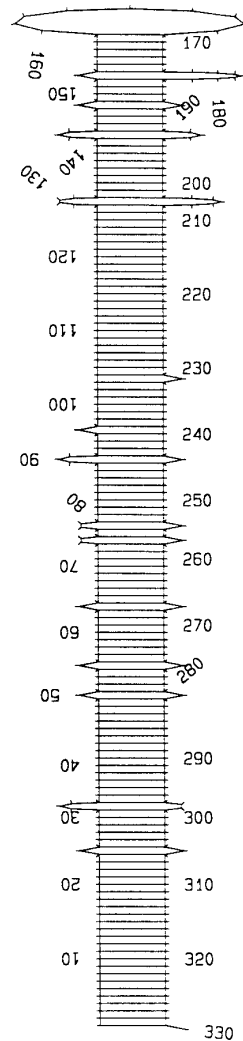


FIG. 5. Predicted secondary structure of one strand of RNA-10. The RNA was folded with the program FOLDRNA and the result displayed graphically with the program SQUIGGLES. Nucleotides are numbered as in Fig. 2B.

RNA-10 and the reverse (complementary) sequence of nucleotides 306–330 of RNA-10 both have 92% sequence identity to the 5'-terminal nucleotides 2–27 of RNA-7. Also the forward and reverse (complementary) sequences of the 3'-terminal nucleotides 305–330 of RNA-10 have 88.5% sequence identity to the 3'-terminal nucleotides 1032–1056 of RNA-7.

Detection of viral single-stranded RNA

Total nucleic acids were extracted from *O. novo-ulmi* isolate Ld and ssRNA was precipitated with 2 M lithium chloride. DsrRNA was then precipitated from the supernatant by adjusting the lithium chloride concentration to 5 M. After agarose gel electrophoresis, Northern blot analysis of the ssRNA and dsRNA fractions was carried out using strand-specific RNA probes synthesized from a cDNA clone of RNA-7.

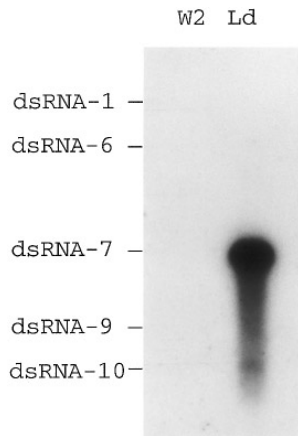


FIG. 6. Northern blot analysis of dsRNA purified from *O. novo-ulmi* healthy isolate W2*tol1* (lane W2) and diseased isolate Ld (lane Ld). The blot was probed with a ^{32}P -labelled cDNA to RNA-7 nucleotides 71–902. The positions of dsRNA segments 1, 6, 7, 9, and 10 of Ld are shown on the side of the blot.

An RNA species, detected in the ssRNA fraction, was confirmed to be single-stranded by its resistance to DNase I (Fig. 8A, lane 6) and its sensitivity to S1 nuclease (Fig. 8A, lane 7). The size of this RNA was estimated by comparison of its electrophoretic mobility with ssRNA markers, produced by *in vitro* transcription of cDNA clones of RNA-7 in LITMUS 28, to be 1.1 kb, which corresponds to the size (1057 nt) of the full-length single-stranded form of dsRNA-7. This RNA was readily detected with a probe corresponding to one strand of RNA-7 (Fig. 8A, lanes 5 and 6), but not with the complementary-strand probe (Fig. 8B, lanes 5 and 6). It is des-

igned the (+) strand of RNA-7 and has the sequence shown in Fig. 2.

An RNA species in the dsRNA fraction was detected by both the (+)-strand and (–)-strand probes (Figs. 8A and 8B, lanes 2–4) and was confirmed to be double-stranded by its resistance to both DNase I (Figs. 8A and 8B, lane 3) and S1 nuclease (Figs. 8A and 8B, lane 4). This dsRNA was identified as dsRNA-7 by comparison of its electrophoretic mobility with those of the Ld dsRNAs, and distinguished from dsRNA-8, which migrates only slightly faster than dsRNA-7, by the absence of hybridization of the probes with dsRNA from isolate sci21 (Fig. 8, lane 1), which contains all the Ld dsRNAs except segments 4, 7, and 10, or with dsRNA from isolate W2*tol1* (Fig. 6, lane W2), which contains only dsRNA segments 1 and 8 (Rogers *et al.*, 1986). Under the conditions of electrophoresis employed here, the single-stranded and double-stranded forms of RNA-7 had similar electrophoretic mobilities.

The intensities of the ssRNA bands in the Northern blot (Fig. 8) were greater than those of the dsRNA bands (compare lanes 2 to 4 with lanes 5 and 6). Since the amounts of ssRNA and dsRNA analysed were derived from the same amounts of total nucleic acids, this infers either that there is more ssRNA-7 than dsRNA-7 in isolate Ld or that the efficiencies of transfer to the membrane, and/or hybridization with the probe, were different for ssRNA-7 and dsRNA-7. The latter explanation seems less likely, because the RNAs were denatured with alkali prior to transfer to the membrane.

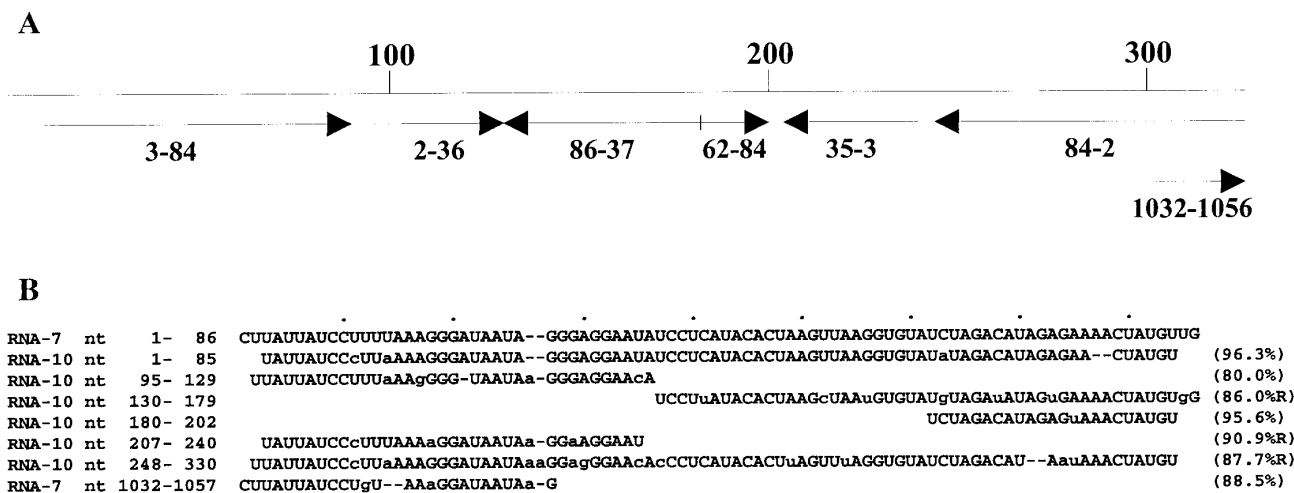


FIG. 7. (A) Scheme for the derivation of RNA-10 from RNA-7. The horizontal line depicts the 330 nucleotide RNA-10. The arrows with numbers underneath indicate the sequences of RNA-7 from which the various regions of RNA-10 are proposed to be derived. Arrows pointing to the right indicate the forward sequence (Fig. 2); arrows pointing to the left indicate the reverse (complementary) sequence. (B) Alignment of nucleotide sequences of regions of RNA-10 with nucleotides 1–86 and 1032–1057 of RNA-7. Numbers in parentheses on the right of the sequences are percentage sequence identities with the corresponding regions from nucleotides 1–86 of RNA-7. R indicates the reverse (complementary) sequence. (–) indicates a gap introduced in the sequence alignments. Each sequence was aligned individually with RNA-7 using the GAP program. Nucleotides in upper- or lowercase indicate identity or nonidentity with the corresponding nucleotide of RNA-7 nucleotides 1–86, respectively.

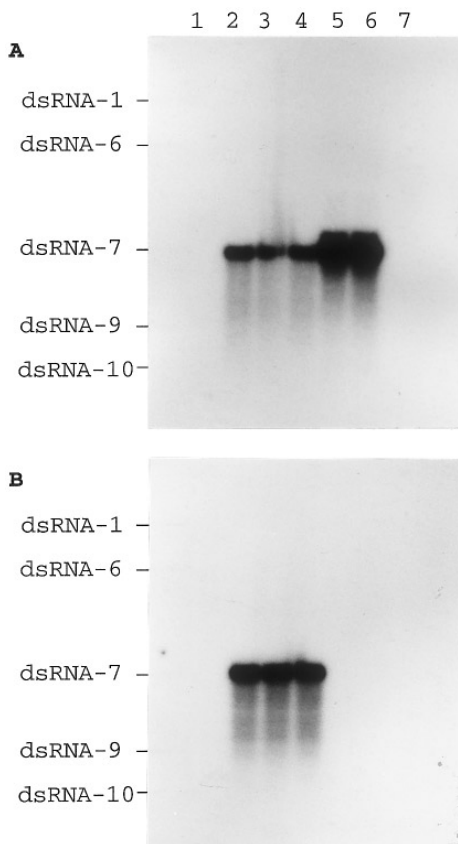


FIG. 8. Northern blot analysis of ssRNA (2 M LiCl-fraction) and dsRNA (5 M LiCl-fraction) from *O. novo-ulmi* isolate Ld. Amounts of ssRNA and dsRNA, equivalent to 10 μ g of total nucleic acids prior to LiCl precipitation, were electrophoresed in 1.4% nondenaturing agarose gels and blotted onto a membrane. Hybridization was carried out using RNA-7 (+)-strand-specific (A) and (-)-strand-specific (B) 32 P-labelled riboprobes, transcribed *in vitro* from a cDNA clone corresponding to RNA-7 nucleotides 167–916. Lane 1, dsRNA from *O. novo-ulmi* isolate sci21; lanes 2, 3, and 4, dsRNA from *O. novo-ulmi* isolate Ld; lanes 5, 6, and 7, ssRNA from *O. novo-ulmi* isolate Ld. Prior to electrophoresis, RNA samples were incubated with DNase I (lanes 3 and 6), DNase I plus S1 nuclease (lanes 4 and 7), or no enzyme (lanes 2 and 5). The positions of dsRNA segments 1, 6, 7, 9, and 10 of isolate Ld are shown on the side of the blot.

DsRNA-7 is not derived from any of the other Ld dsRNA segments

Probes corresponding to RNA-7 nucleotides 167–916, which excludes the terminal sequences in common with RNA-10, did not hybridize with any of the other Ld dsRNAs (Fig. 8). The Northern hybridization assay was found to be at least 100 times more sensitive for detection of dsRNA segments than was ethidium bromide staining and could detect bands not visible by staining. Although the relative abundance of the different dsRNAs in Ld varies over about a 10-fold range (Rogers *et al.*, 1987), with RNA-7 and -10 being the most abundant and RNA-4 being the least abundant, all the 10 dsRNA bands in Ld are readily detected by ethidium bromide staining. Therefore, if any of the other dsRNAs in Ld were related to

nucleotides 167–916 of RNA-7, this relationship would have been detected by Northern hybridization.

DsRNA-7 and dsRNA-10 are not related to *O. novo-ulmi* mitochondrial DNA or RNA

DsRNAs 7 and 10 are associated with the mitochondria (Rogers *et al.*, 1987) and fungal mitochondrial DNAs are characteristically A-U-rich (Paquin *et al.*, 1997), raising the possibility that the dsRNAs might be transcribed from the mitochondrial DNA. However in Northern blots, no hybridization was detected when RNA from healthy *O. novo-ulmi* isolates W2tol1 or sci21 was probed with labelled RNA-7 riboprobes (Figs. 6 and 8). Furthermore, no hybridization was detected when Southern blots of W2tol1 mitochondrial DNA, either uncut or cleaved with *Bcl*I, *Bgl*II, *Eco*RI, *Eco*RV, or *Pvu*II (for a restriction map of W2tol1 mitochondrial DNA, see Charter *et al.*, 1993), were probed with oligo-labelled probes obtained from cDNA clones of RNA-7 and RNA-10 (not shown). Controls of cDNA clones of RNA-7 and RNA-10, included in the blots at one 20th of the molarity of the mitochondrial DNA, were readily detected by the probe. Also there was no significant sequence identity between RNA-7 or RNA-10 and a DNA plasmid derived from the mitochondrial DNA of Ld (Abu-Amro *et al.*, 1995) or any other sequences in the EMBL and GenBank databases. Isolate Ld was originally derived from isolate W2tol1 by natural transmission of dsRNA segments to it from wild *O. novo-ulmi* genotypes in an elm tree (Brasier, 1986) and sci21 is a single conidial isolate derived from Ld. W2tol1 and sci21 are therefore isogenic with Ld, apart from the presence of additional dsRNA segments and a mitochondrial DNA plasmid in the latter. It is therefore concluded that RNA-7 and RNA-10 are not derived from the *O. novo-ulmi* mitochondrial DNA, in agreement with the results of Rogers *et al.* (1987) who could not detect hybridization between *in vitro* labelled total dsRNAs of Ld and the mitochondrial DNA. This supports the hypothesis that the dsRNAs in *O. novo-ulmi* mitochondria are virus-like genetic elements.

DISCUSSION

The 5'- and 3'-termini of dsRNA-7 and dsRNA-10 each contain imperfect inverted repeats, allowing the possibility, in the single-stranded form, of forming stem-loop structures at each of the termini (Figs. 3 A and 3B). In addition, the 5'- and 3'-terminal sequences are complementary to each other with the potential to form a panhandle structure (Fig. 3C). Complementary 5'- and 3'-termini are characteristic of viruses with negative-stranded RNA genomes, such as members of the *Myxoviridae* (Desselberger *et al.*, 1980), and *Rhabdoviridae* (Keene *et al.*, 1979) families, or ambisense RNA genomes, such as members of the *Bunyaviridae* (Elliott, 1990; Kormelink *et al.*, 1992) and *Arenaviridae* (Lukashev-

ich *et al.*, 1997) families, and the *Tenuivirus* (Ramirez and Haenni, 1994) genus. For those viruses with multisegment genomes, the terminal sequences of all the segments of a particular virus are highly conserved. The formation of panhandle structures *in vivo* has been demonstrated for several of these viruses, and in the case of influenza virus, evidence that the structure is important in the transcription and replication of the virus was obtained (Hsu *et al.*, 1987; Hagen *et al.*, 1994). Inverted repeats have also been found close to the 5'- and 3'-termini of dsRNA viruses of the *Phytoreovirus* genus of the *Reoviridae* family, although the sequences of the repeats are different for different dsRNA segments of the same virus (Omura, 1995). Stem-loop structures close to the 3'- and/or 5'-termini are found in many positive-stranded RNA viruses and constitute essential cis-acting recognition signals for RNA replication (Buck, 1996), but for these viruses, the 5'- and 3'-stem-loops have unrelated sequences and the stem does not generally extend completely to the 3'- or 5'-terminus. The presence of inverted repeats at both the 5'- and 3'-termini, combined with the complementarity of the 5'- and 3'-termini, has not been described previously for any replicating RNA, including a mitochondrial dsRNA from *Cryphonectria parasitica* (Polashack and Hillman, 1994), mycovirus dsRNAs of the *Hypoviridae* (Hillman *et al.*, 1995), *Partitiviridae* (Ghabrial *et al.*, 1995a; Oh and Hillman, 1995; Nogawa *et al.*, 1996), and *Totiviridae* families (Ghabrial *et al.*, 1995b), and the positive-stranded RNA mushroom bacilliform virus (*Barnaviridae* family) (Revill *et al.*, 1994). These terminal structures could therefore be indicative of a novel mechanism for the initiation of RNA synthesis. It is noteworthy that the 3'-terminal hairpins would allow self-priming of RNA replication, as occurs with some DNA viruses, such as members of the *Parvoviridae* family (Berns, 1990).

The nucleotide sequences of RNA-7 and RNA-10 indicate that they either do not encode a protein or encode only very short proteins. The replication of these RNAs therefore probably depends on one or more of the other eight dsRNAs in *O. novo-ulmi* isolate Ld to encode an RNA-dependent RNA polymerase and possibly other proteins (Buck, 1996). As most of the sequence of dsRNA-7 is not derived from any of the other Ld dsRNAs, RNA-7 can be regarded as a satellite RNA, defined as an RNA which is dependent on another RNA (helper RNA) for its replication, but which contains substantially distinct nucleotide sequences from the helper RNA or host (Mayo *et al.*, 1995). Satellite RNAs have been described for several animal, fungal, and plant viruses. Some of these encode a protein, e.g., the M₁ dsRNA of *S. cerevisiae* (Wickner, 1996), a satellite dsRNA of *Gaeumannomyces graminis* (Romanos *et al.*, 1981), and ssRNA satellites of plant nepoviruses (Fritsch *et al.*, 1993), whereas others do not encode functional proteins, e.g., dsRNA-3 of *Atkinsonella hypoxylon* partitivirus (Oh and Hillman, 1995)

and satellites of several positive-stranded RNA plant viruses (Roossinck *et al.*, 1992). Although the hybridization data indicate that RNA-7 is not closely related to the *O. novo-ulmi* mitochondrial genome, its A-U-rich nature suggests that it might have originally arisen from a mitochondrial RNA by recombination with a virus RNA to acquire cis-acting sequences for replication and subsequently diverged in sequence to optimise its replication ability.

Comparison of the nucleotide sequences of RNA-7 and RNA-10 indicated that RNA-10 could be derived from RNA-7 (Fig. 7). RNA-10 probably retains sequences from RNA-7 which are required for recognition by its cognate RdRp and other sequences which optimize replication. These would be expected to include 5'- and 3'-terminal sequences, and RNA-10 could be derived from the 5' nucleotides 2–86 and 3' nucleotides 1032–1056 of RNA-7. However, because the 3'-terminal sequence of RNA-7 is both an imperfect direct and an inverted repeat of the 5'-terminal sequence, derivation of RNA-10 from nucleotides 2–86 alone of RNA-7 is also possible. RNA-10 could be formed by RNA recombination mechanisms, analogous to those proposed for the formation of animal and plant defective RNAs (Lai, 1992; Simon and Bujarski, 1994; Nagy and Simon, 1997). In the most favored replicase-driven template switching or copy choice mechanism, the replicase and nascent RNA strand move from a position on one RNA template and reinitiate synthesis either at a different position on the same template or on a different template.

Based on the most extensive sequence identities, RNA-10 appears to be a mosaic formed by copying different regions of RNA-7 (–)- and (+)-strand templates (Fig. 7). A possible model for the formation of RNA-10 involves copying from near the 3'-end of an RNA-7 (–)-strand template to synthesize nucleotides 3–84 of the RNA-7 (+)-strand followed by four template switches: (1) to repeat nucleotides 2–36 of the RNA-7 (+)-strand by recopying from the 3'-end of the (–)-strand template; (2) to synthesize nucleotides 86–37 of the RNA-7 (–)-strand by copying the (+)-strand template; (3) to synthesize nucleotides 62–84 of the RNA-7 (+)-strand by copying the (–)-strand template. A final template switch would involve the nascent strand folding back and acting as its own template to produce a "copy-back" RNA, mediated by complementarity between sequences copied from the (–)- and (+)-strand templates (Fig. 7A). This would explain the high degree of self-complementarity of RNA-10 (Fig. 5) and its "mirror image" structure (Fig. 7A). The model suggests that the regions around nucleotides 35–36 and 84–86 of RNA-7 may be recombination hot-spots. Models of RNA-7 (–)- and (+)-strands constructed with the FOLDRNA program predicted high degrees of secondary structure, suggesting that replicase pausing during RNA synthesis may be common, but did not reveal any particular features around these regions. Replicase

pausing could also be common because of the A-U-rich nature of RNA-7. Pausing at such regions can result in the synthesis of nontemplated nucleotides (Nagy and Simon, 1997) and may explain the apparently nontemplated nucleotides 86–94 and 203–206 in RNA-10. Again the regions around nucleotides 35–36 and 84–86 were not more A-U-rich than other parts of RNA-7. It is possible therefore that the recombinant constituting RNA-10 may have been selected from a range of recombinants on the basis of its ability to replicate. The relatively small divergence of RNA-10 from the RNA-7 sequence may also have resulted from the accumulation of mutations selected on a similar basis.

Defective RNAs have been described for several fungal virus RNAs, such as those derived from the M and L-A dsRNAs of *S. cerevisiae* (Lee *et al.*, 1986; Esteban and Wickner, 1988) and from the hypovirus L-dsRNA of *C. parasitica* (Shapira *et al.*, 1991). These have simpler structures than the *O. novo-ulmi* Ld RNA-10, consisting of a single internal deletion, although some of the hypovirus defective RNAs had nontemplated nucleotides inserted at the breakpoints. More complex defective RNAs, with several deletions and duplications, have been described for some plant RNA viruses, e.g., tomato bushy stunt virus (Hillman *et al.*, 1987; White and Morris, 1994), and a defective RNA of influenza virus consisted of a mosaic of sequences from RNA segments 1 and 3 (Fields and Winter, 1982). Snap-back RNAs have been described for several animal RNA viruses, e.g., defective RNAs of vesicular stomatitis virus (Perrault and Leavitt, 1977) and poliovirus full-length (+)- and (–)-strands covalently linked at one end (Young *et al.*, 1985). However, a combination of a mosaic of sequences from (+)- and (–)-strands and a snap-back structure, such as that found in *O. novo-ulmi* Ld RNA-10, has not been described previously.

MATERIALS AND METHODS

Fungal isolates

O. novo-ulmi isolates Ld and W2tol1 have been described previously (Brasier 1986). Single-conidial isolate sci21 was derived from isolate Ld as described by Rogers *et al.* (1986). All fungal isolates were maintained on malt extract agar (MEA, Oxoid) as described (Rogers *et al.*, 1986).

RNA isolation

Total nucleic acids were isolated as described (Covey and Hull, 1981) from mycelia after 14 days of growth on cellophane membranes overlaying MEA at 22°C. ssRNA was precipitated by addition of LiCl to a final concentration of 2 M (Baltimore, 1966) and, after resuspension, further purified by an additional round of 2 M LiCl precipitation. DsRNA was precipitated from the 2 M LiCl

supernatant by addition of LiCl to a final concentration of 5 M (Diaz-Ruiz and Kaper, 1978) and further purified by an additional round of 2 M LiCl and 5 M LiCl precipitation. DsRNAs of isolate Ld were separated by electrophoresis through a 1.4% agarose gel and extracted from the gel using an RNaid kit (Bio 101). DsRNAs were also analysed by polyacrylamide gel electrophoresis as described by Rogers *et al.* (1986).

cDNA synthesis, cloning, and sequencing

cDNA libraries were constructed using gel-purified dsRNA-7 and dsRNA-10, respectively. DsRNA, mixed with random hexamer primers, was denatured at 94°C for 3 min. First-strand cDNA was then synthesized using Superscript II reverse transcriptase (GIBCO-BRL). Second-strand cDNA synthesis was as described (Gubler and Hoffman, 1983) using a Superscript choice system for cDNA synthesis (GIBCO-BRL). Double-stranded cDNA was ligated either directly into the *EcoRV* site of the vector LITMUS 28 (New England Biolabs), or, after addition of an *EcoRI/Not I* adapter, into the *EcoRI* site of LITMUS 28, and transformed into *Escherichia coli* DH5 α . cDNA clones were primarily screened by PCR using forward (5'-GGGCAGATCTTGAATGCATC-3') and backward (5'-GACGTCACCATGGGAAGCTTC-3') primers corresponding to nucleotides 2466–2486 and 2555–2535 of the LITMUS 28 sequence, respectively, and were further confirmed by Northern blotting. cDNA clones of the ends of dsRNA-7 and dsRNA-10 were determined by the rapid amplification of cDNA ends (RACE) procedure (Frohman *et al.*, 1988) using a GIBCO-BRL 5' RACE system. The 5' RACE PCR products were cloned into the vector pGEM-T Easy (Promega). Sequences of cDNA clones and sub-clones were obtained by the Sanger chain termination method (Sanger *et al.*, 1977) using both [α -³⁵S]dATP (Amersham)/T7 DNA polymerase (Pharmacia) and Dye-Terminator cycle sequencing with AmpliTaq DNA polymerase FS (ABI PRISM 377). Sequence analysis was performed using the University of Wisconsin Genetics Computer Group programs (version 8.1) (Devereux *et al.*, 1984), made available by the Daresbury Laboratory SEQUENT service. The RNA-7 and RNA-10 sequences are available from the Genbank and EMBL databases with the Accession Nos. AJ003120 and AJ003121, respectively.

Northern blotting

DsRNA and ssRNA fractions, pretreated with RNase-free DNase I (Promega) with or without S1 nuclease (GIBCO-BRL) according to the manufacturer's instructions, were separated by electrophoresis through a 1.4% agarose gel, denatured with 50 mM NaOH/150 mM NaCl, neutralized with 1 M Tris-HCl (pH 7.5)/1.5 M NaCl, and blotted to Hybond-N membrane (Amersham). Blots were hybridized with either cDNA probes labelled with [α -³²P]dCTP (Amersham) using a random-primer proto-

col (Feinberg and Vogelstein, 1984) or strand-specific riboprobes prepared by *in vitro* transcription using T7 RNA polymerase (NBL) and [α - 32 P]UTP (Amersham) (Sambrook *et al.*, 1989).

ACKNOWLEDGMENTS

We thank the Gatsby Charitable Foundation for support of this work. T.E.C. is the recipient of a CASE award from the Biotechnology and Biological Sciences Research Council.

REFERENCES

- Abu-Amero, S. N., Charter, N. W., Buck, K. W., and Brasier, C. M. (1995). Nucleotide-sequence analysis indicates that a DNA plasmid in a diseased isolate of *Ophiostoma novo-ulmi* is derived by recombination between two long repeat sequences in the mitochondrial large subunit ribosomal RNA gene. *Curr. Genet.* **28**, 54–59.
- Baltimore, D. (1966). Purification and properties of poliovirus double-stranded RNA. *J. Mol. Biol.* **18**, 421–428.
- Berns, K. I. (1990). Parvovirus replication. *Microbiol. Rev.* **54**, 316–329.
- Brasier, C. M. (1986). The d-factor in *Ceratocystis ulmi*: Its biological characteristics and implications for Dutch elm disease. In "Fungal Virology" (K. W. Buck, Ed.), pp. 177–208. CRC Press, Boca Raton, FL.
- Buck, K. W. (1996). Comparison of the replication of positive-stranded RNA viruses of plants and animals. *Adv. Virus Res.* **47**, 159–251.
- Buck, K. W. (1997). Molecular variability of viruses of fungi. In "Molecular Variability of Fungal Pathogens" (Y. Couteaudier, P. D. Bridge, and J. M. Clarkson, Eds.), in press. CAB International, Wallingford.
- Charter, N. W., Buck, K. W., and Brasier, C. M. (1993). De-novo generation of mitochondrial DNA plasmids following cytoplasmic transmission of a degenerative disease in *Ophiostoma novo-ulmi*. *Curr. Genet.* **24**, 505–514.
- Covey, S. N., and Hull, R. (1981). Transcription of cauliflower mosaic virus DNA. Detection of transcripts, properties and location of the gene encoding the virus inclusion body protein. *Virology* **111**, 463–474.
- Desselberger, U., Racaniello, V. R., Zazra, J. J., and Palese, P. (1980). The 3' and 5' end terminal sequences of influenza A, B, and C virus RNA segments are highly conserved and show partial inverted complementarity. *Gene* **8**, 315–328.
- Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 539–549.
- Diaz-Ruiz, J. R., and Kaper, J. M. (1978). Isolation of viral double-stranded RNA using a LiCl fractionation procedure. *Prep. Biochem.* **8**, 1–13.
- Domingo, E., and Holland, J. J. (1994). Mutation rates and rapid evolution of RNA viruses. In "The Evolutionary Biology of Viruses" (S. S. Morse, Ed.), pp. 161–184. Raven Press, New York.
- Eigen, M., and Biebricher, C. (1988). Sequence space and quasispecies distribution. In "RNA Genetics" (E. Domingo, J. J. Holland, and P. Ahlquist, Eds.), Vol. 3, pp. 211–245. CRC Press, Boca Raton, FL.
- Elliott, R. M. (1990). Molecular biology of the Bunyaviridae. *J. Gen. Virol.* **71**, 501–522.
- Esteban, R., and Wickner, R. B. (1988). A deletion mutant of L-A double-stranded RNA replicates like M₁ dsRNA. *J. Virol.* **62**, 1278–1285.
- Feinberg, A. P., and Vogelstein, B. (1984). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**, 266–267.
- Fields, S., and Winter, G. (1982). Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structure of a small RNA segment. *Cell* **28**, 303–313.
- Fritsch, C., Mayo, M. A., and Hemme, O. (1993). Properties of satellite RNA of nepoviruses. *Biochimie* **75**, 561–567.
- Frohman, M. A., Dush, M. K., and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
- Ghabrial, S. A., Bozarth, R. F., Buck, K. W., Yamashita, S., Martelli, G. P., and Milne, R. G. (1995a). Family *Partitiviridae*. In "Virus Taxonomy, Sixth Report of the International Committee on Taxonomy of Viruses" (F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers, Eds.), pp. 253–260. Springer-Verlag, Wien/New York.
- Ghabrial, S. A., Bruenn, J. A., Buck, K. W., Wickner, R. B., Patterson, J. L., Stuart, K. D., Wang, A. L., and Wang, C. C. (1995b). Family *Totiviridae*. In "Virus Taxonomy, Sixth Report of the International Committee on Taxonomy of Viruses" (F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers, Eds.), pp. 245–252. Springer-Verlag, Wien/New York.
- Goodman, M. F. (1995). DNA polymerase fidelity: Misinsertions and mismatched extensions. In "PCR Strategies" (M. A. Innis, D. H. Geland, and J. J. Sninsky, Eds.), pp. 17–31. Academic Press, London.
- Gubler, U., and Hoffman, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263–269.
- Hagen, M., Chung, T. D. Y., Butcher, J. A., and Krystal, M. (1994). Recombinant influenza virus polymerase: Requirement of both 5' and 3' viral ends for endonuclease activity. *J. Virol.* **68**, 1509–1515.
- Hillman, B. I., Carrington, J. C., and Morris, T. J. (1987). A defective interfering RNA that contains a mosaic of a plant virus genome. *Cell* **51**, 427–433.
- Hillman, B. I., Fulbright, D. W., Nuss, D. L., and Van Alfen, N. K. (1995). Family *Hypoviridae*. In "Virus Taxonomy, Sixth Report of the International Committee on Taxonomy of Viruses" (F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers, Eds.), pp. 261–264. Springer-Verlag, Wien/New York.
- Hsu, M.-T., Parvin, J. D., Gupta, S., Krystal, M., and Palese, P. (1987). Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. *Proc. Natl. Acad. Sci. USA* **84**, 8140–8144.
- Keene, J. D., Schubert, M., and Lazzarini, R. A. (1979). Terminal sequences of vesicular stomatitis virus RNA are both complementary and conserved. *J. Virol.* **32**, 167–174.
- Kormelink, R., de Haan, P., Meurs, C., Peters, D., and Golbach, R. (1992). The nucleotide sequence of the M segment of tomato spotted wilt virus, a bunyavirus with two ambisense RNA segments. *J. Gen. Virol.* **73**, 2795–2804.
- Lai, M. (1992). RNA recombination in animal and plant viruses. *Microbiol. Rev.* **65**, 61–79.
- Lee, M., Pietras, D. F., Nemeroff, M. E., Corstanje, B. J., Field, L. J., and Bruenn, J. A. (1986). Conserved regions in defective interfering viral double-stranded RNA from a yeast virus. *J. Virol.* **58**, 402–407.
- Lukashevich, I. S., Djavani, M., Shapiro, K., Sanchez, A., Ravkov, E., Nichol, S. T., and Salvato, M. S. (1997). The Lassa fever virus: Nucleotide sequence, comparison, and precipitation of a predicted 250 kDa protein with monospecific antiserum. *J. Gen. Virol.* **78**, 547–551.
- Mayo, M. A., Berns, K. I., Fritsch, C., Kaper, J. M., Jackson, A. O., Leibowitz, M. J., and Taylor, J. M. (1995). Subviral agents: Satellites. In "Virus Taxonomy, Sixth Report of the International Committee on Taxonomy of Viruses" (F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers, Eds.), pp. 487–492. Springer-Verlag, Wien/New York.
- Moriyama, H., Nitta, T., and Fukuhara T. (1995). Double-stranded RNA in rice: A novel RNA replicon in plants. *Mol. Gen. Genet.* **248**, 364–369.
- Nagy, P. D., and Simon, A. E. (1997). New insights into the mechanisms of RNA recombination. *Virology* **235**, 1–9.
- Nogawa, M., Kaweyama, T., Nakatani, A., Taguchi, G., Shimosaka, M., and Okazaki, M. (1996). Cloning and characterization of mycovirus double-stranded RNA from the plant pathogenic fungus, *Fusarium solani* f. sp. *robiniae*. *Biosci. Biotechnol. Biochem.* **60**, 784–788.

- Oh, C.-S., and Hillman, B. I. (1995). Genome organization of a partitivirus from the filamentous ascomycete *Atkinsonella hypoxylon*. *J. Gen. Virol.* **76**, 1461–1470.
- Omura, T. (1995). Genomes and primary protein structures of phyto-reoviruses. *Semin. Virol.* **6**, 97–102.
- Osawa, S., Jukes, T. H., Watanabe, K., and Muto, A. (1992). Recent evidence for the evolution of the genetic code. *Microbiol. Rev.* **56**, 229–264.
- Paquin, B., Laforest, M. J., Forget, L., Roewer, I., Wang, Z., Longcore, J., and Lang, B. F. (1997). The fungal mitochondrial genome project: evolution of fungal mitochondrial genomes and their gene expression. *Curr. Genet.* **31**, 380–395.
- Perrault, J., and Leavitt, R. W. (1977). Inverted complementary terminal sequences in single-stranded RNAs and snap-back RNAs from vesicular stomatitis defective-interfering particles. *J. Gen. Virol.* **38**, 35–50.
- Polashack, J. J., and Hillman, B. I. (1994). A small mitochondrial double-stranded RNA element associated with a hypovirulent strain of the chestnut blight fungus and ancestrally related to yeast cytoplasmic T and W RNAs. *Proc. Natl. Acad. Sci. USA* **91**, 8680–8684.
- Ramirez, B. C., and Haenni, A. L. (1994). Molecular biology of tenuiviruses, a remarkable group of plant viruses. *J. Gen. Virol.* **75**, 467–475.
- Revill, P. A., Davidson, A. D., and Wright, P. J. (1994). The nucleotide sequence and genome organization of mushroom bacilliform virus: A single-stranded RNA virus of *Agaricus bisporus* (Lange) Imbach. *Virology* **202**, 904–911.
- Rodríguez-Cousino, N., Esteban, L. M., and Esteban, R. (1991). Molecular cloning and characterization of W double-stranded RNA, a linear molecule present in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **266**, 12772–12778.
- Rogers, H. J., Buck, K. W., and Brasier, C. M. (1986). Transmission of double-stranded RNA and a disease factor in *Ophiostoma ulmi*. *Plant Pathol.* **35**, 277–287.
- Rogers, H. J., Buck, K. W., and Brasier, C. M. (1987). A mitochondrial target for double-stranded RNA in diseased isolates of the fungus that causes Dutch elm disease. *Nature* **329**, 558–560.
- Rogers, H. J., Buck, K. W., and Brasier, C. M. (1988). DsRNA and disease factors of the aggressive subgroup of *Ophiostoma ulmi*. In "Viruses of Fungi and Simple Eukaryotes" (Y. Koltin, and M. J. Leibowitz, Eds.), pp. 327–351. Dekker, New York/Basel.
- Romanos, M. A., Buck, K. W., and Rawlinson, C. J. (1981). A satellite double-stranded RNA from a virus from *Gaemannomyces graminis*. *J. Gen. Virol.* **57**, 375–385.
- Roossinck, M. J., Sleat, D., and Palukaitis, P. (1992). Satellite RNAs of plant viruses: Structures and biological effects. *Microbiol. Rev.* **56**, 265–279.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- Shapira, R., Choi, G. H., Hillman, B. I., and Nuss, D. L. (1991). The contribution of defective RNAs to the complexity of viral-encoded double-stranded RNA populations present in hypovirulent strains of the chestnut blight fungus *Cryphonectria parasitica*. *EMBO J.* **10**, 741–746.
- Simon, A. E., and Bujarski, J. J. (1994). RNA-RNA recombination and evolution in virus-infected plants. *Annu. Rev. Phytopathol.* **32**, 337–362.
- Stuart, K. D., Weeks, R., Guilbride, L., and Myler, P. J. (1992). Molecular organization of *Leishmania* RNA virus 1. *Proc. Natl. Acad. Sci. USA* **89**, 8596–8600.
- Webber, J. F. (1993). D-factors and their potential for controlling Dutch elm disease. In "Dutch Elm Disease Research: Cellular and Molecular Approaches" (M. B. Sticklen and J. L. Sberald, Eds.), pp. 322–332. Springer-Verlag, New York.
- White, K. A., and Morris, T. J. (1994). Nonhomologous RNA recombination in tombusviruses: Generation and evolution of defective interfering RNAs by stepwise deletions. *J. Virol.* **68**, 14–24.
- Wickner, R. B. (1996). Double-stranded RNA viruses of *Saccharomyces cerevisiae*. *Microbiol. Rev.* **60**, 250–265.
- Young, D. C., Tuschall, D. M., and Flanagan, J. B. (1985). Poliovirus RNA-dependent RNA polymerase and host cell protein synthesize product twice the size of poliovirus RNA in vitro. *J. Virol.* **54**, 256–264.