Characterization of a Subgenomic DNA Isolated from *Triticum aestivum* Plants Infected with Wheat Dwarf Virus

By H. MACDONALD,† R. H. A. COUTTS AND K. W. BUCK*

Department of Pure and Applied Biology, Imperial College of Science and Technology, London SW7 2BB, U.K.

(Accepted 26 February 1988)

**SUMMARY**

A subgenomic DNA has been isolated from wheat tissue infected with a Swedish isolate of wheat dwarf virus and cloned. Restriction endonuclease analysis and nucleotide sequence determination indicated that the subgenomic DNA (1472 nucleotides) was derived from the genomic DNA (2749 nucleotides) by two separate deletions. The subgenomic DNA had lost open reading frames (ORFs) encoding the virus coat protein and a putative protein of Mr 10146, but retained an ORF and an open reading region encoding putative proteins of Mr 30156 and 17292, respectively, and two structural features thought to be important for virus DNA replication i.e. a potentially stable stem–loop structure containing the conserved TAATATTAC sequence and a putative primer initiation site.

**INTRODUCTION**

Geminiviruses have double-icosahedral or 'geminate' particles, a genome of circular ssDNA molecules of 2.5 to 3.0 kb and a single capsid polypeptide species of Mr 28000 to 30000 (for reviews, see Harrison, 1985; Stanley, 1985; Lazarowitz, 1987). There are two subgroups of geminiviruses: (i) viruses with unipartite genomes, leafhopper vectors and monocotyledonous hosts, e.g. maize streak virus (MSV) (Mullineaux *et al.*, 1984; Grimsley *et al.*, 1987), digitaria streak virus (DSV) (Donson *et al.*, 1987) and (ii) viruses with bipartite genomes, dicotyledonous hosts and whitefly vectors, e.g. African cassava mosaic virus (ACMV) (synonym cassava latent virus) (Stanley, 1983), bean golden mosaic virus (Morinaga *et al.*, 1983) and tomato golden mosaic virus (TGMV) (Hamilton *et al.*, 1983).

Genomic ssDNA molecules, their dsDNA counterparts and dimeric DNA molecules have been detected in extracts of plants infected with geminiviruses of each subgroup, i.e. ACMV (Stanley & Townsend, 1985), TGMV (M. J. Slomka, K. W. Buck & R. H. A. Coutts, unpublished data), and DSV (Donson *et al.*, 1987). Also, subgenomic ss and dsDNA species of approximately half the length of the genomic DNA molecules, and derived from the smaller of the two genomic DNA molecules, have been isolated from plants infected with isolates of ACMV (Stanley & Townsend, 1985; Coutts & Buck, 1987), TGMV (MacDowell *et al.*, 1986) and potato yellow mosaic virus (Roberts *et al.*, 1988). We now report the isolation and characterization of a subgenomic DNA isolated from plants infected with a Swedish isolate of wheat dwarf virus (WDV), a leafhopper-transmitted geminivirus which we have recently shown by agroinfection (Grimsley *et al.*, 1986, 1987) with dimers of cloned virus DNA (MacDowell *et al.*, 1985) to have a unipartite genome (Hayes *et al.*, 1988b).

**METHODS**

*Enzymes and isotoates.* Restriction endonucleases, DNA polymerase I and T4 DNA ligase were obtained from NBL Enzymes Ltd. (Northumbria Biochemicals, Cramlington, U.K.) and used according to the manufacturer's recommendations. DNA polymerase I Klenow fragment, [α-35S]dATP and [α-32P]dCTP were from Amersham.

† Present address: Friedrich Miescher-Institut, P.O. Box 2543, CH-4002 Basel, Switzerland.
Isolation of viral DNA. A Swedish isolate of WDV in wheat tissue (Triticum aestivum L. cv. Diamant) was kindly supplied by Dr K. Lindsten. Nucleic acid extracts from infected tissue were produced as described previously (MacDowell et al., 1985). Fractions enriched for supercoiled dsDNA were prepared by treatment of these extracts with alkali, followed by centrifugation to equilibrium in a caesium chloride–ethidium bromide density gradient as described by Sunter et al. (1985).

Electrophoresis and blotting. DNA was electrophoresed in agarose gels and transferred to GeneScreen Plus membranes (New England Nuclear) as described by Hayes et al. (1988a). The blots were pre-hybridized and hybridized to 32P-labelled, nick-translated DNA (Rigby et al., 1977) according to the manufacturer’s (New England Nuclear) recommendations. Restriction fragments were eluted from agarose gels according to Dretzen et al. (1981).

Cloning and sequencing. DNA manipulations were as described by Maniatis et al. (1982) unless otherwise stated. DNA was transformed into Escherichia coli DH5a (Jessee, 1986) using the methods of Hanahan (1985) and Murray (1986) with cloning vectors pEMBL9 (Dente et al., 1985) and M13mp18 and 19 (Messing, 1983). Nucleotide sequencing of WDV subgenomic DNA clones was by the dideoxynucleotide chain termination method (Sanger et al., 1977) using 6% denaturing polyacrylamide gels (Sanger & Coulson, 1978) and a strategy similar to that used for sequencing the genomic clone (MacDowell et al., 1985).

RESULTS

Detection and cloning of WDV subgenomic DNA

Analysis of DNA from WDV-infected plants, by agarose gel electrophoresis and Southern blotting (Fig. 1), revealed a number of WDV-specific DNA species (lane 2). These included the virion ssDNA, genomic length dsDNA forms (supercoiled, open circular and linear) and a novel species, designated band X. DNA preparations enriched for supercoiled DNA forms (lane 3) contained increased proportions of genomic length supercoiled DNA and the faster migrating band X, suggesting that the latter is a supercoiled subgenomic DNA. This suggestion was confirmed by digestion of the supercoiled DNA with restriction endonucleases HindIII (lane 4) and BclI (lane 5), each of which cut WDV genomic DNA once (MacDowell et al., 1985). In each case, two restriction fragments were formed, one of 2.7 kbp, which corresponded to linearized DNA of genomic size, and another of 1.4 kbp arising by linearization of the supercoiled subgenomic DNA. It is clear, therefore, that the subgenomic DNA retains the recognition sites for both HindIII and BclI (Fig. 2). The small amount of virion ssDNA remaining in the preparation of supercoiled DNA (lane 3) was unaffected by digestion with HindIII (lane 4) or BclI (lane 5).

When the products of HindIII digestion of supercoiled DNA were cloned into the HindIII site of pEMBL9, two classes of WDV-specific clone were obtained. The most abundant clones contained inserts of 2.7 kbp, corresponding to genomic length WDV dsDNA. Other clones had inserts of 1.4 kbp corresponding to subgenomic length WDV dsDNA. Further clones of subgenomic DNA were obtained by separating the HindIII digestion products of supercoiled DNA by agarose gel electrophoresis, eluting the band corresponding to the subgenomic DNA and cloning it into HindIII-cut M13mp19.

Restriction endonuclease and sequence analysis of WDV subgenomic DNA

Restriction endonuclease mapping of 28 independent subgenomic clones indicated that they were all similar to the genomic DNA but carried one or more deletions totalling 1.3 kbp somewhere between the two EcoRI sites at nucleotides 2678 and 2141 in the genomic DNA (Fig. 2). The complete nucleotide sequence of one clone of subgenomic DNA (WD2) was determined and compared to the sequence of a full-length genomic clone (WD1) derived from the same batch of WDV-infected wheat tissue (MacDowell et al., 1985). There were three differences (Fig. 2): (i) a deletion of 1210 bp between nucleotides 181 and 1392 of the full-length genomic clone, (ii) a second deletion of 67 bp between two GGT sequences of the genomic clone at nucleotides 2727 and 45 respectively (one of the GGT sequences was retained in the subgenomic...
Wheat dwarf virus subgenomic DNA

Fig. 1. Southern blot of DNA from WDV-infected wheat plants. Lane 1, DNA from purified WDV virions; lane 2, total DNA from WDV-infected plants; lane 3, DNA from WDV-infected plants, enriched for supercoiled forms; lanes 4 and 5, DNA as in lane 3, digested with HindIII and BclI, respectively. The blot was probed with 32P-labelled, nick-translated pWD1. oc, lin, sc and ss refer to the genome-length open circular, linear and supercoiled ds forms and virion ssDNA respectively. Band X is described in the text. Sizes of linear dsDNA molecules (kbp) are shown on the right hand side of the gel.

Fig. 2. Map of the WDV genome (2749 nucleotides) showing the regions deleted in the subgenomic DNA (hatched areas). Arrows on the outside of the circle represent ORFs (solid) or an open reading region (broken) with the Mr (x 10^-3) of the putative protein which is encoded indicated by a number. The thick arrow indicates the position at which small DNA molecules are tightly bound to the virion genomic ssDNA (Hayes et al., 1988c). O indicates the position of a potentially stable stem-loop structure containing the sequence TAATATTAC and x indicates the position of the C to G transversion in subgenomic clone WD2. Relevant restriction endonuclease recognition sites are as follows. B, BclI; E, EcoRI; H, HindIII; Hp, HpaI.

DNA) and (iii) a C to G transversion at nucleotide 2530 of the genomic clone. The two deletions together account for the 1.3 kbp deletion observed by restriction mapping.

Eleven of the subgenomic clones were cleaved with EcoRI and the 0.9 kbp fragment was subcloned into M13mp19 in the appropriate orientation so that dideoxynucleotide sequencing could be performed across both deletions. In all cases the deletions were found to be identical to those in clone WD2. The deletion boundaries are shown in Fig. 3.

When supercoil-enriched DNA from WDV-infected plants was digested with restriction endonuclease HpaII and the products were analysed by agarose gel electrophoresis and Southern blotting, bands of the sizes predicted from the sequences of the genomic and subgenomic DNA clones were detected (Fig. 4). The 780, 450, 210 and 170 bp fragments were specific to the genomic DNA, the 330 bp fragment was specific to the subgenomic DNA and the 750 and 390 bp fragments were common to both (Fig. 2). The presence of the 330 bp fragment and the absence of any additional bands shows that the deletions were present in all the subgenomic DNA in infected plants and not only in the subgenomic clones isolated.
Subgenomic DNA (large deletion)

CGTGGAAGATGTTGAAATACAACACCCTG

Genomic DNA

CGTGGAAGATGTTGAAATACAACACCCTG

170 180 190 1380 1390 1400

Subgenomic DNA (small deletion)

TTATATAGGGCAGGTGTGGCGGTCGGGGGGG

Genomic DNA

TTATATAGGGCAGGTTTTGGCGGGAGAACA

AATATATCCCGTCCAAAACCGCCCTC

2720 2730 2740 40 50 60

Fig. 3

Fig. 3. Nucleotide sequences at the junctions of the large and small deletions in cloned WDV subgenomic DNA and the corresponding regions of cloned WDV genomic DNA. For the subgenomic DNA, the sequence corresponding to the virion sense DNA is shown. For the genomic DNA, the sequence of both strands is shown; the solid line denotes the positions of the sequence GGT and the arrows indicate the positions of the sequences CTT and GTT.

Fig. 4

Fig. 4. Southern blot of WDV dsDNA. Lane 1, DNA from WDV-infected plants, enriched for supercoiled forms; lane 2, DNA as in lane 1, digested with HpaII; lane 3, enlargement and lower exposure of part of lane 2 to allow resolution of the two bands of 780 and 750 nucleotides. The blot was probed as in Fig. 1. Sizes of HpaII restriction fragments (bp) are shown at the side of the gel. ss, virion ssDNA of genomic length; g, supercoiled DNA of genomic length; sg, supercoiled subgenomic DNA.

DISCUSSION

Restriction mapping and sequencing have shown that the WDV subgenomic DNA could be derived from the genomic DNA by two deletions. The clones were obtained from DNA prepared from two different batches of wheat plants infected with the same WDV isolate and the deletions in all the clones which were sequenced were identical. Hence, there appears to be only one major subgenomic DNA species associated with this WDV isolate.

The large deletion completely removes the open reading frames (ORFs) which encode putative proteins of Mr 29400 (29-4K) and 10-1K. On the basis of amino acid sequence homology with the coat protein of MSV (MacDowell et al., 1985), the 29-4K protein is probably the WDV coat protein which, by analogy with ACMV (Stanley & Townsend, 1986) and TGMV (Brough et al., 1988), may not be required for systemic infection. The ORF for the 10-1K protein is conserved in MSV (Mullineaux et al., 1984) and DSV (Donson et al., 1987), suggesting an essential function. It is likely, therefore, that the subgenomic DNA is unable to replicate in the absence of the genomic DNA.

It is noteworthy that the subgenomic DNA retained two features thought to be important for virus DNA replication (Fig. 2): (i) a potentially stable stem–loop structure containing the sequence TAATATTAC which is conserved in all geminivirus DNAs which have been sequenced (MacDowell et al., 1985; Howarth & Goodman, 1986; Donson et al., 1987) and (ii) the 3' part (in the virion DNA sense) of a sequence to which, in the genomic virion ssDNA, small DNA molecules are tightly bound (Hayes et al., 1988c) and which may be important for the priming of DNA synthesis in vivo.
The subgenomic DNA also retained the ORF and open reading region encoding putative proteins of M, 30-2K and 17-2K (Fig. 2) which correspond to the N-terminal and C-terminal region of a conserved ORF in the bipartite geminiviruses, encoding a protein of M, 40K believed to be important for virus DNA replication (MacDowell et al., 1985). Expression of these sequences in the subgenomic DNA may make it less dependent on the genomic DNA. In the clone (WD2) for which the complete sequence was obtained, only one base change (a C to G transversion) was found in this region and this would lead to an amino acid change (glutamate to glutamine) at residue 22 of the 30-2K protein. This glutamate is not conserved in the corresponding proteins of MSV, TGMV or ACMV (MacDowell et al., 1985). It is likely, therefore, that the ORF encoding the protein of 30-2K in the WDV subgenomic DNA is functional. However, it is noteworthy that the downstream 'A-T' box from nucleotides 1277 to 1299 in the genomic sequence, which contains a putative transcriptional polyadenylation signal, was absent in the subgenomic DNA.

Although this is the first subgenomic DNA of a unipartite geminivirus to be described, subgenomic DNAs derived from the smaller of the two DNA components of ACMV (Stanley & Townsend, 1985) and TGMV (MacDowell et al., 1986) have been characterized. These subgenomic DNAs did not retain any functional ORFs, although the 200 bp 'common regions', containing the conserved TAATATTAC sequence, were retained. The ACMV and TGMV subgenomic DNA species were approximately half the size of the genomic DNAs, suggesting that their ssDNA forms might be encapsidated in the half-size isometric particles usually found in geminivirus preparations. The WDV subgenomic DNA (1472 nucleotides) is just over half the size of the genomic DNA (2749 nucleotides). It is possible that the ss form of the subgenomic DNA, found in some but not all preparations of WDV virion DNA, may also be encapsidated in half-size particles.

The nucleotides adjacent to the deletion points in the ACMV genomic DNA showed repeat sequences of two to seven nucleotides, one copy of which was retained in the subgenomic DNA (Stanley & Townsend, 1985). It was suggested that these repeats may contribute to the exact location of the recombination event postulated to be responsible for generation of the subgenomic DNA. A similar repeat (GGT) was found adjacent to the small deletion in WDV DNA and, again, one copy was retained in the subgenomic DNA. However, repeats adjacent to the large WDV deletion, or to the TGMV deletion points (MacDowell et al., 1986), were not found.

MacDowell et al. (1986) suggested that subgenomic DNAs might be generated by topoisomerase I-mediated non-homologous recombination. CTT and GTT sequences, for which topoisomerase I shows some preference, occur close to the deletion boundaries of both ACMV and TGMV deletions and also occur near the junctions of both WDV deletions (Fig. 3). This model, first proposed to explain the positions of cross-over points for excision of non-tandem chromosomal inserts of simian virus 40 DNA in mammalian cells (Bullock et al., 1985), may also be applicable to non-homologous recombination in plant cells.

We thank K. Lindsten for WDV-infected wheat tissue, V. E. Prince for assistance and the Science and Engineering Research Council for the award of a research grant to K.W.B. and R.H.A.C. Work with WDV was carried out with a licence (no. PHF 29A/143) issued by the Ministry of Agriculture, Fisheries and Food.

REFERENCES


(Received 22 December 1987)