Sequences of ten circular ssDNA components associated with the milk vetch dwarf virus genome

Yoshitaka Sano,1 Masafumi Wada,2 Yoshifumi Hashimoto,1 Tsuguo Matsumoto1 and Makoto Kojima2

1 Department of Applied Biology, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan
2 Lab. Plant Pathology, Faculty of Agriculture, Niigata University, Ikarashi 2-8050, Niigata 950-2102, Japan

Milk vetch dwarf virus (MDV) is a member of the proposed genus Nanovirus, and its genome is composed of multiple, circular ssDNA components of about 1 kb. We have cloned and sequenced ten ssDNA components and designated them MDV-C1 to C10. Each DNA component contains one potential major open reading frame, and contains a putative stem–loop structure in the non-coding region. Notably, four components (C1, C2, C3 and C10) encode distinct replication-associated (Rep) proteins of 33 kDa, which show only limited (42–57%) amino acid identity. The six other components encode proteins with calculated molecular masses ranging from 12.7 to 19.7 kDa. Comparison of the sequences with those of other nanoviruses reveals that MDV is closely related to faba bean necrotic yellows virus (FBNYV) and subterranean clover stunt virus (SCSV). Six putative MDV genome products, including one Rep and five non-Rep proteins, show high (70.4–90.9%) amino acid identity to the corresponding six FBNYV proteins, whereas two other Rep proteins encoded by MDV-C2 and C3 are 82.3% and 73.0% identical to those encoded by SCSV-C2 and C6, respectively. These results indicate that MDV, FBNYV and SCSV have diverged from a common origin, which had multiple Rep components. In addition, the putative proteins encoded by MDV-C4 and its homologues contain a consensus retinoblastoma-binding motif, suggesting that they may be involved in controlling the host cell cycle.

Introduction

Milk vetch dwarf virus (MDV) is an aphid-borne virus that infects several legume crops in Japan. The name is derived from Chinese milk vetch (Astragalus sinicus L.), a common green-manure crop, from which the disease was first reported (Matsuura, 1953; Inouye et al., 1968). The virus is transmitted by Aphis craccivora in a persistent manner, and causes yellowing and dwarfing of Chinese milk vetch, broad bean (Vicia faba L.), pea (Pisum sativum L.) and soybean [Glycine max (L.) Merr.] plants. MDV was originally thought to be a luteovirus because of its symptoms, vector transmission and the observation of 26 nm isometric particles in MDV-infected broad bean leaves (Ohki et al., 1975). However, isometric particles measuring 18 nm in diameter and containing ssDNA of about 1 kb were later isolated from infected pea leaves (Isogai et al., 1992; Sano et al., 1993). Furthermore, MDV was shown to be serologically related to faba bean necrotic yellows virus (FBNYV), an isometric ssDNA virus occurring in North Africa and western Asia (Katul et al., 1993; Franz et al., 1996). Thus, MDV is now considered to be a member of the genus Nanovirus, which includes FBNYV, subterranean clover stunt virus (SCSV; Chu & Helms, 1988), banana bunchy top virus (BBTV; Harding et al., 1991) and coconut foliar decay virus (CFDV; Randles & Hanold, 1989).

To obtain basic information on the ssDNA genome of MDV, and to define its taxonomic relationship to other nanoviruses, we cloned and sequenced MDV DNA. Here, we report the sequences of ten ssDNA components of MDV, and compare them with those available for related viruses, together with the sequences of four newly identified DNA components of FBNYV (C7–C10) reported in the preceding paper (Katul et al., 1998).

Methods

Virus and DNA. The N isolate of MDV, obtained from naturally infected broad bean, was maintained in broad bean plants, and propagated in pea plants using A. craccivora as a vector. The virus was purified from infected pea leaves by the method described for FBNYV (Katul et al.,
Replicative-form (RF) dsDNA was isolated from infected pea leaves by a modification of the method of Sunter et al. (1984). Leaf material was pulsed in liquid nitrogen and ground in 2 vols 50 mM Tris–HCl, pH 7.6 containing 100 mM NaCl. The homogenate was mixed with 2 vols 0.2 M NaOH–1% SDS and then with 1.5 vols 3 M potassium acetate, pH 4.8. After incubation at 0 °C for 1 h, the mixture was centrifuged at 10,000 g for 10 min. The supernatant was extracted once with phenol–chloroform (1:1), and then with chloroform. The aqueous phase was removed and mixed with 2 vols ethanol, and then centrifuged. The crude nucleic acid pellet was dissolved in TE (10 mM Tris–HCl, pH 8.0; 1 mM EDTA) containing 10 µg/ml RNase A (Pharmacia), incubated at 37 °C for 30 min, and electrophoresed on a 1.2% agarose gel containing 0.5 µg/ml ethidium bromide. The viral dsDNA band was excised from the gel, electroeluted and precipitated with ethanol. The pellet was washed with 70% ethanol and finally resuspended in sterile distilled water.

Characterization of ten ssDNA components

Ten circular ssDNA components (C1–C10) associated with the MDV N isolate were identified and sequenced. Analysis of the sequences revealed that each DNA contained only one potential open reading frame (ORF) that encoded a protein larger than 10 kDa. The sizes, putative motifs and functions of the single ORFs, together with possible gene regulation signals, are shown in Fig. 1. All the DNA components had a region of 29–34 bases capable of forming a stem–loop structure, as has been identified consistently in the genome structure, as has been identified consistently in the genome of other nanoviruses, as well as in those of geminiviruses and PCV (Meehan et al., 1997). Furthermore, C4–C9 also had certain similar sequences in the non-coding regions. The lengths of the common regions varied from 47 bases in C8 to 437 bases in C4, C5 and C7, in which more than 93% of the nucleotides were identical. C1, C2, C3 and C10 did not contain the equivalent regions due to the large sizes of their coding regions. Analysis of the predicted amino acid sequences revealed that C1, C2, C3 and C10 all encoded putative replication-associated (Rep) proteins. The major ORFs of these DNA fragments were similar in size (281–284 amino acids), and each contained a NTP-binding motif (Gorbalenya et al., 1990) as well as three other motifs typical of replication-initiator proteins involved in rolling-circle DNA replication (RCR-I, II and III; Koonin & Ilyina, 1992) (Fig. 1). However, they had only limited amino acid identity (42–58%) to each other.

Potential TATA boxes and AATAAA-like polyadenylation signals were appropriately located for all the major ORFs of the ten DNA components. Although C3 and C10 did not contain well-defined TATA boxes at appropriate positions, AATAATA analogues were located 59 and 57 bases, respectively, upstream of the start codons of the predicted ORFs (Fig. 1, marked by grey triangles). All of the proposed poly(A) signals were preceded 19–41 bases upstream by at least one (A/T)TGTAA motif (Sanfaçon, 1994; Rothnie et al., 1994), and were followed downstream by a GT-rich stretch (Gil & Proudfoot, 1984; Conway & Wickens, 1985) containing...
Components of the milk vetch dwarf virus genome

Fig. 1. Diagrammatic representation of the proposed organization of the ssDNA genome of MDV. The single ORF (open box) on the virus-sense strand of each DNA component is shown clockwise.

Fig. 2. Alignment of the nucleotide sequences of the putative stem–loop structures of the genome components of MDV, FBNYV, SCSV, BBTV-A (Australian isolate), BBTV-T (Taiwanese isolate), CFDV, PCV, MSV, BCTV and BGMV. The highly conserved nonanucleotide sequences of the loop-forming domain are boxed; asterisks indicate the nucleotides conserved in all the components. The related stem-forming sequences of six MDV components and those of FBNYV, SCSV and BBTV-T are shaded.

5' stem loop stem 3'

The TTG sequence within 47 bases. In addition to the AATAAA signal, the (A/T)TGTA element and the downstream GT-rich sequence have been shown to be involved in correct 3'-end processing of CFDV Rep gene transcripts (Merits et al., 1995).

Fig. 2 shows an alignment of the putative stem–loop sequences of the ten MDV components, together with those identified in the ssDNA components of FBNYV (Katul et al., 1995, 1997, 1998), SCSV (Boevink et al., 1995), BBTV (Harding et al., 1993; Wu et al., 1994; Burns et al., 1995), CFDV (Rohde et al., 1990), PCV and three representatives of the geminivirus genera, MSV, BCTV and BGMV. In addition to the highly conserved loop sequences, particular components of MDV and other nanoviruses had extended sequence similarity within their stem-forming domains (Fig. 2, indicated by grey shading). Most evidently, six MDV components (C4–C9) and seven FBNYV components (C2–C6, C8 and C10) were identical within the stem–loop sequence over at least 30 bases. In contrast, the sequences of MDV-C1, C2, C3 and C10, all of which encoded putative Rep proteins, were less well conserved and had many nucleotide substitutions in the stems and in the 5'-proximal regions of the loop domains. It has been demonstrated that, upon initiation of rolling-circle DNA replication, the geminivirus Rep protein cleaves the RF templates and introduces a nick within a conserved nonanucleotide (5'TAATATT$AC3') on the virus-sense strand (Heyraud et al., 1993; Laufs et al., 1995; Stanley, 1995). The Rep protein of BBTV-C1 was also shown to cleave ssDNA in vitro between positions +7 and +8 of a nonanucleotide (5'TAATATT$AC3') (Hafner et al., 1997). The conserved nonanucleotides of the MDV components are thus likely to be involved in initiation of virus DNA synthesis.

Analysis of virus-sense polarity

The polarity of the viral ssDNA was examined with strand-specific RNA probes transcribed from a full-length cDNA clone of MDV-C6. The complementary-sense RNA hybridized with virus-sense DNA present in crude RF DNA, purified
virions and viruliferous aphids (Fig. 3b, lanes 1, 2 and 3, respectively), whereas the virus-sense probe hybridized only with the RF DNA sample (Fig. 3c, lane 1), indicating that only the coding strand was encapsidated in particles. The results also indicate that MDV does not multiply in the aphid vector, since complementary-sense DNA could not be detected in viruliferous aphids. The polarities of other components, as shown in Fig. 1, were deduced by analogy to C6: all the MDV components had a conserved nonanucleotide within the putative loop domain (Fig. 2), and six, including C6, also shared sequences surrounding the stem–loop domains.

Predicted genome organization and relationships with other nanoviruses

The sequences of the putative genome products encoded by the major ORFs of MDV-C1 to -C10 were aligned and compared with those identified in the DNA components of FBNYV, SCSV, BBTV and CFDV. Calculated amino acid identities of the putative Rep and non-Rep proteins are shown in Tables 1 and 2, respectively. Three of the four putative MDV Rep proteins showed striking similarity to particular Rep proteins of FBNYV and SCSV: the Rep proteins encoded by MDV-C10, C2 and C3 had 89-8, 82-3 and 73-0% amino acid identity to those encoded by FBNYV-C7, SCSV-C2 and SCSV-C6, respectively (Table 1). Comparisons also indicated at least 35-7% amino acid identity between the putative Rep proteins of MDV and other nanoviruses. When the sequences were compared with the putative Rep protein of PCV (Meehan et al., 1997) or with those of MSV, BCTV and BGMV, the four MDV Rep proteins showed only weak similarity (17-7–24-5% amino acid identity, not shown).

In addition to the Rep proteins, both MDV and FBNYV have six putative gene products, whereas only five and four putative non-Rep proteins have been identified for SCSV and BBTV, respectively. MDV-C7 and FBNYV-C6 each encode a protein with a molecular mass of 15 kDa, which has no counterpart in the putative proteins of SCSV and BBTV. Likewise, a homologue of the proteins encoded by FBNYV-C3 and SCSV-C7, which are similar to that encoded by MDV-C5, has not been found in the BBTV components sequenced to date. While six DNA components have been identified in an Australian isolate of BBTV, the one named BBTV-C2 does not contain a relevant ORF (Burns et al., 1995). In addition, BBTV-C1 has a second, small ORF within the major ORF (Beetham et al., 1997). However, we did not find an equivalent small ORF in the MDV components. The proposed CFDV genome differs from other nanoviruses in that the single ssDNA component of 1291 bases potentially encodes six ORFs, which are thought to be transcribed bi-directionally (Rohde et al., 1990). Comparison between CFDV and MDV of all putative proteins did not show

Table 1. Comparison of the four putative Rep proteins of MDV with those of related ssDNA viruses

<table>
<thead>
<tr>
<th>Percentage identity</th>
<th>MDV</th>
<th>FBNYV</th>
<th>SCSV</th>
<th>BBTV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C2</td>
<td>C3</td>
<td>C10</td>
<td>C1</td>
</tr>
<tr>
<td>MDV C1</td>
<td>42-1</td>
<td>49-3</td>
<td>56-0</td>
<td>52-8</td>
</tr>
<tr>
<td>(56-7)</td>
<td>(59-3)</td>
<td>(60-9)</td>
<td>(61-8)</td>
<td>(55-1)</td>
</tr>
<tr>
<td>MDV C2</td>
<td>43-9</td>
<td>44-4</td>
<td>62-4</td>
<td>63-7</td>
</tr>
<tr>
<td>–</td>
<td>(56-6)</td>
<td>(50-3)</td>
<td>(64-4)</td>
<td>(51-4)</td>
</tr>
<tr>
<td>MDV C3</td>
<td>57-9</td>
<td>57-9</td>
<td>48-1</td>
<td>39-7</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
<td>(63-3)</td>
<td>(58-5)</td>
<td>(53-4)</td>
</tr>
<tr>
<td>MDV C10</td>
<td>49-3</td>
<td>39-7</td>
<td>89-8</td>
<td>89-8</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
<td>–</td>
<td>(58-7)</td>
<td>(53-7)</td>
</tr>
</tbody>
</table>
The six putative proteins of MDV encoded by C4–C9 were most identical to the related proteins of FBNYV (52.0–90.9%, Table 2). The percentage amino acid identities calculated between the corresponding five proteins of MDV and SCSV were intermediate (44.8–66.7%), and similar to the levels of identity calculated between the corresponding five Rep proteins of SCSV and FBNYV (43.0–66.0%, not shown). In contrast, the amino acid sequence identities calculated between the corresponding four non-Rep proteins of MDV and BBTV were considerably lower (21.1–46.8%). Of the six putative non-Rep proteins, that encoded by MDV-C6 and the related proteins encoded by FBNYV-C8, SCSV-C4 and BBTV-C6 showed higher levels of amino acid identity than other proteins; for example, those encoded by MDV-C6 and FBNYV-C8 were 90.9% identical. In contrast, the proteins encoded by MDV-C7 and FBNYV-C6, which have been found in only these two viruses, shared only 52.0% identity. The ten DNA components identified in this study may represent the entire genome of the N isolate of MDV. The predicted genome organization of MDV is in agreement with that of FBNYV, and we did not detect the presence of additional components. However, due to the small size of the DNA components, it cannot be excluded that other genome components had no appropriate restriction sites, and were therefore missed.

The putative protein encoded by MDV-C9 had a molecular mass of 19.3 kDa, and showed 83.1, 55.5 and 23.7% amino acid identity to the capsid proteins (CPs) encoded by FBNYV-C5 (Katul et al., 1997), SCSV-C5 (Chu et al., 1993) and BBTV-C3 (Wanitchakorn et al., 1997), respectively. From this result, and the observation that MDV and FBNYV share common epitopes (Franz et al., 1996), it is likely that the 19.3 kDa protein encoded by MDV-C9 is the virus CP.

The putative 12.7 kDa protein encoded by MDV-C8 had 76.3, 48.3 and 21.8% amino acid identity to those encoded by FBNYV-C4, SCSV-C1 and BBTV-C4, respectively. Based on structural similarity to the movement protein (MP) of MSV (Boulton et al., 1993; Dickinson et al., 1996), these 13 kDa proteins have been suggested to be movement proteins (Burns et al., 1995; Katul et al., 1997). Although none of these proteins has significant sequence similarity to the MSV MP, they all have a hydrophobic domain of 25–30 amino acids in the N-terminal region (not shown), which is also present in the MSV MP.

Database searches with the amino acid sequences of the putative MDV proteins revealed no significant matches other than with related nanovirus proteins, and the functions of the proteins encoded by MDV-C4, C5, C6 and C7 remain unknown. However, the 19 kDa proteins encoded by MDV-C4, FBNYV-C10 (sequence not shown), SCSV-C3 and BBTV-C5 all contain the consensus retinoblastoma (Rb)-binding motif (LXCXE) at equivalent positions, suggesting that they may be involved in controlling the host cell cycle (Fig. 4a, boxed). No other Rep or non-Rep protein of MDV, FBNYV, SCSV or BBTV contains the LXCXE sequence. The Rb tumour suppressor is the key regulatory factor of cell cycle progression at the G1/S boundary. Tumour virus oncoproteins are known to inactivate Rb by forming a stable complex through the LXCXE motif, thereby driving the host cell cycle into S phase, where the cellular environment is suitable for replication and transcription of viral DNA. The Rep protein of wheat dwarf virus (WDV), a monopartite geminivirus, has been shown to interact with the human (Xie et al., 1995; Collin et al., 1996) and maize Rb proteins (Xie et al., 1996) through its LXCXE motif.

The proposed LXCXE motifs are surrounded by many acidic amino acids (Fig. 4b, shown in bold), as is commonly observed in viral and cellular proteins which interact with Rb. Also, the 19 kDa proteins each contain a basic amino acid cluster (R or K) in the N-terminal domain (Fig. 4a, boxed), which closely resembles the bipartite motif, a nuclear targeting signal (rXrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr
been suggested to be satellite DNAs (Horser et al., personal communication). Two of the MDV Rep proteins (M23022) and wheat dwarf virus (WDV) (X02869), Digitaria streak virus (DSV) protein (SWISS-PROT P03129), maize streak virus (MSV) (X01633) antigen (T-ag) (J02400), human papilloma virus type 16 (HPV-16) E7 oncoproteins, geminivirus Rep proteins and plant D-type cyclin (Soni et al., 1995). The key residues of the motif (L, C and E) are shown in white on the right. Amino acids constituting putative motifs (see text) are marked by shaded boxes. (a) Comparison of the LXXE motif in the 19 kDa proteins of MDV-C4, BBTV-C5 and SCSV-C3, with those of animal DNA virus genomes. The possible origins of the Rep components presumably resulted from recombination events. Based on the overall sequence similarities, MDV and FBNYV are more closely related to each other than to SCSV. SCSV was first reported in the early 1950s as a serious pathogen of subterranean clover (Trifolium subterraneum L.), and subsequently spread throughout Australia (Chu et al., 1995). The relatively low sequence similarity of SCSV to the other two viruses may reflect the geographical or evolutionary isolation of SCSV from MDV and FBNYV.

Analysis of the sequence variability of BBTV-C1 and -C6 reveals that BBTV isolates from more than eight different countries can be separated into two large groups, the South Pacific and the Asian groups (Karan et al., 1994, 1997). The sequence variability between the two groups is 9±6% (nucleotide) and 5±6% (amino acid) in C1, and 14±5% (nucleotide) and 6±7% (amino acid) in C6. Results from a previous serological study led to the suggestion that MDV and FBNYV are strains of the same virus (Franz et al., 1996). The putative MDV CP encoded by C9 had 78±9% nucleotide and 83±1% amino acid identity to FBNYV-C5. We feel that these levels of similarity are too low for MDV and FBNYV to be different strains of the same virus. Furthermore, none of the corresponding DNA components of MDV and FBNYV have nucleotide sequence identities over 90%. Hence, based on the overall nucleotide and amino acid sequence similarities, we propose that MDV and FBNYV should be regarded as separate species. Similar taxonomic criteria have been proposed for classifying the strains and species of geminiviruses (Padidam et al., 1995; Hong & Harrison, 1995). A certain amount of serial or genomic variability has been identified within isolates of SCSV (Chu et al., 1995) and FBNYV (Franz et al., 1996). For a more accurate definition of strains and species, it will be necessary to examine the sequence variability among different isolates of MDV, FBNYV and SCSV.

We wish to thank Drs L. Katul and H. J. Vetten for supplying sequence information prior to publication. We also thank Dr J. L. Dale for valuable suggestions, T. Fujita for his technical assistance, and the late Dr N. Ogasawara (Plant Biological Defence System Lab) for providing DNA sequencing facilities. This work was supported in part by a grant-in-aid from the Ministry of Education, Science, Sports and Culture, Japan.
Components of the milk vetch dwarf virus genome

References


Received 27 April 1998; Accepted 10 August 1998