Sequence data to settle the taxonomic position of bean common mosaic virus and blackeye cowpea mosaic virus isolates

Jawaid A. Khan,* Dick Lohuis, Rob Goldbach and Jeanne Dijkstra

Department of Virology, Agricultural University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands

The nucleotide sequences of the coat protein genes and 3’ non-translated regions (3’-NTRs) of three isolates of bean common mosaic virus (NL1, NL3 and NY15) and one isolate of blackeye cowpea mosaic virus (W) were determined. Comparison of these sequences revealed that the coat proteins of NL1, NY15 and W were identical in size (287 amino acids) and exhibited an overall sequence similarity (94 to 97 %), and 84 to 98 % in their N-terminal regions. Furthermore, their 3’-NTRs were very similar in length [253 to 256 nucleotides (nt)] and sequence (93 to 96 % similarity). In contrast, the coat protein of NL3 had only 261 amino acids and showed 87 to 89 % similarity with NL1, NY15 and W whereas its N-terminal region revealed only 46 to 61 % similarity. The 3’-NTR of NL3 also displayed appreciable differences, both in length (240 nt) and sequence (56 to 63 % similarity). These results, in combination with earlier serological findings, justify the conclusion that NL1, NY15 and W should be considered strains of the same virus, i.e. bean common mosaic virus, and that NL3 is a strain of a different potyvirus for which the name ‘bean black root virus’ is proposed.

Introduction

Bean common mosaic virus (BCMV) and blackeye cowpea mosaic virus (BICMV), both members of the genus Potyvirus, family Potyviridae (Barnett, 1991), have been informally placed in the bean common mosaic virus subgroup (Dijkstra & Khan, 1992; McKern et al., 1992a; Mink & Silbernagel, 1992). Despite their economic importance, the taxonomic status of BCMV and BICMV has so far remained unclear. Classical taxonomic parameters such as biological properties and conventional serology have not led to a proper distinction of strains or species within the BCMV subgroup. In the past, the main reason for classifying BCMV and BICMV as two distinct viruses was the pathogenicity of BCMV to bean and that of BICMV to cowpea (Drijfhout, 1978; Lana et al., 1988). Mink & Silbernagel (1992), using polyclonal and monoclonal antibodies, stated that despite all the work done in several laboratories, confusion still exists in labelling isolates or strains within the azuki bean mosaic virus/BCMV/BICMV/cowpea aphid-borne mosaic virus cluster of viruses.

Recently, the increased knowledge of coat protein structure has contributed greatly in differentiating between potyviruses and their strains. Extensive comparisons of a large number of potyviruses have revealed that distinct potyviruses show 38 to 71 % (average of 54 %) sequence identity in their coat proteins, whereas this identity is greater than 90 % among strains belonging to the same virus. The N-terminal domain of the coat protein in particular differs markedly between individual viruses, whereas the central core and C-terminal domain are highly conserved (Shukla & Ward, 1988, 1989). Furthermore, it has been shown that the 3’ non-translated regions (3’-NTRs) of different potyviruses also display a high degree of sequence variation (similarity of 39 to 53 % only) whereas this sequence is highly conserved between virus strains (similarity of 83 % and more) (Frenkel et al., 1989). With N-terminal domain-targeted serology it has been possible to distinguish between some strains of BCMV and BICMV (Khan et al., 1990). Using antibodies directed towards the N-terminal peptide domains of the coat proteins in ELISA and electroblot immunoassay (EBIA), BCMV strain NL1 reacted with antibodies to BCMV strain NY15 and BICMV strain W, whereas NY15 did not react with those to W or vice versa. Furthermore, BCMV strain NL3 reacted with none of the antisera to the isolate tested except with that to BCMV strain NL5. Therefore, judging from the N-terminal domains of the coat proteins, BCMV strains NL3 and NY15 and BICMV strain W should all be regarded as different viruses (Shukla et al., 1989; Shukla & Ward, 1989). However, based on HPLC profiles of digested coat proteins it was concluded that BCMV strains NL1 and NY15 and BICMV strain W are strains of one virus, whereas NL3 is a strain of another virus (McKern et al., 1992a, b).
In view of these conflicting results the sequences of the coat protein genes and 3'-NTRs of BCMV strains NL1, NL3 and NY15 and BICMV strain W have now been determined. The implications for the taxonomic status of these viruses and their strains are discussed.

Methods

**Virus purification and RNA extraction.** BCMV strains NL1, NL3 and NY15 were isolated from leaves of infected Phaseolus vulgaris, and BICMV strain W was isolated from Nicotiana benthamiana, as described earlier (Khan et al., 1990). RNA was extracted by incubation with 1% SDS followed by phenol-chloroform extraction and subsequent ethanol precipitation (Maniatis et al., 1982).

**Primer design.** Initially, potyvirus group-specific degenerate primers UW48 and D341 positioned in the core of the coat protein gene (Langeveld et al., 1991), kindly supplied by Dr Simon Langeveld, were used to determine the sequences from the core and polymerase genes of NL1, NL3 and NY15. Subsequent designs were based on determined partial sequences (Table 1).

**Amplification of coat protein genes including 3'-NTRs of NL1, NL3 and NY15 by PCR.** First-strand cDNA synthesis was performed using 0.5 µg viral RNA and oligo(dT)$_{12-18}$ as primer. For amplification of cDNA representing part of the 3'-NTR, the C-terminal region and the core of the coat protein, primers U341 and oligo(dT) were used. For amplification of cDNA representing the N-terminal region, the core of the coat protein and part of the polymerase gene, primers DW48 and UW19 were used. PCR was performed using 5 µl of first-strand DNA synthesized from viral RNA, 50 pmol of each primer and 1 unit Super Taq polymerase. Thirty-five reaction cycles were performed with 30 s for annealing at 50 °C, 1 min 30 s for synthesis at 72 °C and 1 min for melting at 94 °C. Samples (5 µl) were analysed by agarose gel electrophoresis.

**Cloning and nucleotide sequencing of PCR-amplified fragments.** After electrophoresis, the desired amplified DNA fragments of NL1, NL3 and NY15 from the N-terminal regions of their coat protein genes and 3'-NTRs were isolated by the freeze–squeeze method (Tautz & Renz, 1983). For efficient cloning, a T vector was constructed from plasmid pBlueScript (Stratagene) digested with EcoRV, and incubated with Super Taq polymerase in the presence of dITP (Marchuk et al., 1990). The DNAs were then ligated into the T cloning vector and transformed into competent DH5α cells, recombinants were selected from filter replicas using [³²P]ATP-radiolabelled first-strand DNA as a probe. Two clones, pWCP-13 and pWCP-8 containing inserts of 1.3 and 0.8 kb, respectively, were selected for the purpose of sequencing, which was done as described above.

**Comparison of sequences.** The nucleotide sequences of the protein genes, their deduced amino acids and the nucleotide sequence data for the 3'-NTRs of all four isolates were compiled, analysed and the level of sequence relatedness was compared using the GCG program package from the University of Wisconsin (Devereux et al., 1984) and where necessary they were aligned manually.

Results

**Amplification and sequence analysis of coat protein genes and 3'-NTRs of NL1, NL3, NY15 and W**

The oligonucleotide sequences and positions of downstream and upstream primers are shown in Table 1. Primers UW48 and DW19 amplified DNA fragments of about 600 bp covering part of the polymerase gene and the N-terminal half of the coat protein gene. Primers U341 and oligo(dT) amplified DNA fragments of about 750 bp covering the C-terminal half of the coat protein gene and the 3'-NTR (Fig. 1 and 2).

Table 1. Oligonucleotide primer sequences

<table>
<thead>
<tr>
<th>Primer*</th>
<th>Sequence</th>
<th>Position†</th>
</tr>
</thead>
<tbody>
<tr>
<td>UW18</td>
<td>5'C GGACTACTT CGGAATTTTC 3'</td>
<td>609-626</td>
</tr>
<tr>
<td>DW19</td>
<td>5'C TTGCTCAC TCCATCCTAC 3'</td>
<td>459-478</td>
</tr>
<tr>
<td>UW48</td>
<td>5'C CAGCAT AAAAAAT CTTTAACAC 3'</td>
<td></td>
</tr>
<tr>
<td>DW49</td>
<td>5'C CTTTCAC CCAGCGCCAGT 3'</td>
<td>294-314</td>
</tr>
<tr>
<td>U341</td>
<td>5'C CGGAAT CATTGTTGTTYGAYTIGAAAYGGG 3'</td>
<td>477-501</td>
</tr>
<tr>
<td>D341</td>
<td>5'C CGCGATC CCGCGTTTT TCA TTCGTTTGC 3'</td>
<td>697-722</td>
</tr>
<tr>
<td>U1000</td>
<td>5'C ACIGTIGAYAAYYS IY AYGGG 3'</td>
<td></td>
</tr>
</tbody>
</table>

* U and D are upstream and downstream primers, respectively.
† Positions of the primers are based on the nucleotide sequence of BCMV strain NY15.
Fig. 1. (a) General potyviral genome map; (b) positions of the different PCR primers are shown by arrows. (c) N-ter is PCR-amplified DNA fragment containing part of the polymerase (POL), the N-terminal region and the core of coat protein (CP) of BCMV strains NL1, NL3 and NY15 and BICMV strain W; 3'-NTR is a PCR-amplified fragment containing part of the core, the C-terminal region of coat protein and the 3'-NTRs of BCMV strains NL1, NL3 and NY15 and BICMV strain W. (d, e) cDNA clones pWCp-13 and pWCp-8, respectively, used to determine the nucleotide sequence of part of the polymerase gene (POL), the coat protein gene and the 3'-NTRs.

Fig. 2. PCR amplification of fragments of BCMV strains NL1 (lane 1), NL3 (lane 3) and NY15 (lane 15) and BICMV strain W (lane W) containing their 3'-NTRs and the coat protein N-terminal (N-ter) regions. A 5 µl aliquot was analysed on an agarose gel containing ethidium bromide.

Fig. 3. Comparison of the nucleotide sequences of the polymerase (partly) and the coat protein genes of BCMV strains NL1, NL3 and NY15 and BICMV strain W. Nucleotides coding for N-terminal regions of the coat proteins are underlined and positions of stop codons are shown by asterisks.
were predicted. The putative proteolytic cleavage site W, respectively [excluding the poly(A) tail].

From the nucleotide sequences, the amino acid sequences (~) between the NIb protein and the coat protein genes (§) and 3'-NTRs (¶).

Fig. 4. Comparison of amino acid sequences of coat proteins of BCMV strains NL1, NL3 and NY15 and B1CMV strain W. Identical amino acids are indicated by dashes and gaps are indicated by dots. The putative cleavage site is indicated by the vertical arrow. The locations of the N and C termini of the coat proteins are predicted trypsin-sensitive sites.

Table 2. Amino acid and nucleotide sequence similarities (%) between the coat proteins and 3'-NTRs of NL1, NL3, NY15 and W.

<table>
<thead>
<tr>
<th>Strain</th>
<th>NL1</th>
<th>NL3</th>
<th>NY15</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>NL1</td>
<td>53*</td>
<td>98</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>NL3</td>
<td>746</td>
<td>97</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>NY15</td>
<td>96</td>
<td>86</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>90</td>
<td>90</td>
<td>93</td>
<td></td>
</tr>
</tbody>
</table>

* Above the diagonal. Amino acid sequence similarity (%) between the N-terminal domains (*), whole coat proteins (†) and cores (‡). § Below the diagonal. Nucleotide sequence similarity (%) between the coat protein genes ($) and 3'-NTRs (¶).

240, 256 and 254 nucleotides for NL1, NL3, NY15 and W, respectively [excluding the poly(A) tail].

Comparison of coat protein sequences

From the nucleotide sequences, the amino acid sequences were predicted. The putative proteolytic cleavage site (ESVXXQ/S) between the NIb protein and the coat protein (Fig. 4) was in agreement with M, determinations and sequence comparisons with other isolates of BCMV (Vetten et al., 1992). Fig. 4 shows the multiple alignment of the predicted amino acid sequence data for the coat proteins of NL1, NL3, NY15 and W. The coat proteins of NL1, NY15 and W isolates were identical in size (287 amino acids), whereas that of NL3 was shorter (261 amino acids).

Among isolates NL1, NY15 and W, the N-terminal region of the coat proteins also appeared highly conserved, both in length (52 amino acids) and in sequence (similarity 84 to 98%). In contrast, the N-terminal domain of NL3 was significantly shorter (26 amino acids) and had a distinct sequence (similarity to NL1, NY15 and W was only 46 to 61%). Furthermore, the core and C-terminal regions of NL1, NL3, NY15 and W coat proteins were all fairly conserved in length and in sequence (Fig. 4, Table 2). Thus, the overall sequence similarity of the coat proteins of NL1, NY15 and W was quite high (94 to 97%) and comparable to identities observed between strains of a single potyvirus species (Shukla & Ward, 1988; Ward et al., 1992). On the other hand, since the N-terminal part of the NL3 coat protein showed only limited similarity with those of NL1, NY15 and W, this isolate may be considered a distinct virus.

Reference has to be made here regarding the presence of a conserved motif DAG in the N-terminal region of potyvirial coat proteins required for aphid transmission.
Taxonomy of BCMV and BICMV isolates

Comparison of the 3'-NTRs

The 3'-NTRs of NL1, NY15 and W were all similar in length [253, 256 and 254 nucleotides (nt), respectively] whereas that of NL3 was only 240 nt long. For strains NL1, NY15 and W, the sequence similarity in this region ranged from 93 to 96% (Fig. 5; Table 2). This high percentage was in line with the high sequence identity of their coat proteins. However, the 3'-NTR of the isolate NL3 revealed a lower similarity (56 to 63%) to those of the other three isolates.

Discussion

The availability of the sequences of both the coat proteins and 3'-NTRs of potyvirus isolates NL1, NL3, NY15 and W allowed a reliable establishment of their taxonomic status. The sequence homology found between both the coat protein genes and the 3'-NTRs of NL1, NY15 and W suggested that these isolates are strains of the same potyvirus. The coat proteins of NL1, NY15 and W were identical in size (287 amino acids) and displayed high sequence similarity. Furthermore, the number and sequence of nucleotides of their 3'-NTRs were almost identical. The coat protein sequence of NL3 showed an overall sequence similarity of 87 to 89% with those of NL1, NY15 and W, values which lie between those of different species (38 to 71%) and strains of the same species (greater than 90%). However, the sequence differences between NL3 and the other isolates were located in the N-terminal region (sequence similarity of 46 to 61%). Furthermore, both the sequence and size of the 3'-NTR of NL3 differ greatly from those of NL1, NY15 and W, indicating that NL3 should be considered a distinct virus. In a sequence comparison of complete coat proteins of NL1, NY15 and W with that of isolates of peanut stripe virus (serologically closely related to BCMV isolates) and BCMV strain NL4, a similarity of 92 to 95% was found and their 3'-NTRs showed a similarity of more than 85 to 98% (McKern et al., 1991; Vetten et al., 1992; Cassidy et al., 1993). Furthermore, NL3 showed high sequence similarities of 98 and 92% at the coat protein and 3'-NTR levels, respectively, with the sequence of BCMV strain NL8 (Vetten et al., 1992).

The present results corroborate earlier findings that NL3 differs greatly in its biological properties from the other BCMV strains and that NL1, NY15 and W are rather closely related. According to the classification of Drijfhout (1978), NL3 belongs, together with NL5 and NL8, to the pathogenicity group of BCMV strains that cause systemic necrosis (black root) in certain bean cultivars.

Remarkably, the above assignments also correlate well with previous information on serological and other coat protein properties of strains of these viruses (Mink & Silbernagel, 1992; Vetten et al., 1992).

Conventional serological studies showed that NL1, NY15 and W are closely related to each other (Lana et al., 1988). In contrast, NL3 showed only a distant relationship with NL1, NY15 and W. Such distant relationships may be due to the presence of antibodies directed towards the core region of the coat proteins (Shukla & Ward, 1989). Interestingly, when antibodies directed towards the N-terminal region of NL3 coat protein were used in double-antibody sandwich (DAS)-ELISA and EBIA, they did not cross-react with NL1, NY15 or W but only with the necrotic strain NL5 (Khan et al., 1990).

In contrast to the results with N-terminal serology, comparison of coat protein HPLC peptide profiles revealed a large degree of similarity between the coat proteins of NY15 and W suggesting they are strains of the same virus (McKern et al., 1992b). It is worth mentioning that N-terminal serology may sometimes give erratic results (Khan et al., 1990; Shukla et al., 1989) due to unexpected paired relationships between distinct viruses or failed cross-reactions between some strains (Shukla et al., 1992). Indeed, the N-terminal regions of coat proteins of NL1, NY15 and W are highly similar but despite this, both in DAS-ELISA and EBIA, NY15 and W did not cross-react whereas NL1 did so with antisera to NY15 and W. Presumably, only few potential amino acid residues in an epitope are necessary for antibody binding or, alternatively, the conformation of the epitope involved is changed after removal of the N-terminal region by lysyl endopeptidase (Geyser et al., 1984, 1987, 1988; Shukla et al., 1989). A comparison of amino acid sequences of the N-terminal domains of NL1, NY15 and W showed that the W sequence differed from those of the other strains by six amino acid residues [Thr(30), Gly(41), Asn(53), Ser(64), Val(68), Thr(75)]. These differences may be responsible for the serological reaction of W with NL1, but not with NY15 (Fig. 4).

Secondary structures of the N-terminal domains of all these isolates, as determined by the Peptide Structure Program of GCG, showed the same folding pattern. However, striking differences were observed in the secondary structure of the NL3 N-terminal domain.
(results not shown). Epitope mapping of the mutated N-terminal regions expressed in heterologous systems may give an insight regarding the complexity of these serological results.

In summary, the following conclusions may be drawn with regard to the taxonomic position of NL1, NL3, NY15 and W. Based on sequences of both the coat protein gene and the 3'-NTR, and serological and biological properties, NL3 should no longer be considered a strain of BCMV but a strain of a different potyvirus, as suggested by Dijkstra & Khan (1992). McKern et al. (1992a, b) and Vetten et al. (1992). For this virus, McKern et al. (1992a) and Vetten et al. (1992) proposed the names ‘bean necrosis mosaic virus’ and ‘bean necrotic mosaic virus’, respectively, indicating that in some bean cultivars it induces systemic necrosis whereas in others it causes mosaic. However, we would favour a name bringing out the most characteristic feature of this virus, i.e. its ability to cause severe vascular necrosis in certain bean cultivars, called black root (Grogan & Walker, 1948). We consider the combination ‘necrosis (necrotic) mosaic’ an undesirable one as it is confusing and therefore we propose this virus should be named ‘bean black root virus’. The fact that not all bean cultivars infected with this virus show black root disease need not be an objection. There are many similar examples, such as tobacco mosaic virus which does not induce mosaic in all tobacco cultivars.

We agree with others (McKern et al., 1992a, b) that BCCMV and the non-necrosis-inducing strains of BCMV, such as NL1 and NY15, should be called strains of one virus, BCMV. However, for relevance to applied virology the strain name must remain recognizable by plant pathologists and breeders. Following the principle of nomenclature proposed by Shukla & Ward (1989), we recommend the addition of ‘blackeye cowpea’ (BIC) for a particular strain of BCMV, adding a suffix to indicate the place of isolation, characteristic symptoms on certain plants, etc. For example, the present BCCMV strain W would then be referred to as BCMV strain B1C/W.

We thank Dr Richard Kormelink for making the cDNA clones to BCCMV W and Dr Simon Langeveld for supplying the degenerate primers.

References


Taxonomy of BCMV and BlCMV isolates


(Received 22 March 1993; Accepted 8 June 1993)