Nucleotide sequences of the coat protein genes of two aphid-transmissible strains of soybean mosaic virus

Ch. Jayaram, John H. Hill and W. Allen Miller*

Department of Plant Pathology, Iowa State University, Ames, Iowa 50011, U.S.A.

The nucleotide sequences of the coat protein genes and 3' non-coding regions of two aphid-transmissible isolates (G2 and G7) of soybean mosaic virus (SMV) were determined. The coat protein of the G2 isolate differs from that of the aphid non-transmissible N isolate by a single amino acid at position 12 (aspartic acid in N, glycine in G2 and G7). The G7 isolate differs from G2 and N at three and four amino acid residues, respectively. The nucleotide sequence similarity of the three isolates in the coat protein coding and 3' non-coding regions ranges from 93 to 100%, respectively. In contrast, watermelon mosaic virus 2 (WMV 2) is only 77% to 79% similar to the SMV isolates, suggesting that WMV 2 is a distinct virus and not an isolate of SMV as proposed previously.

Most potyviruses are transmitted by aphids in a non-persistent fashion. However, aphid non-transmissible strains of several potyviruses have been found. Changes in either a non-structural helper component (HC) protein or the coat protein can eliminate aphid transmissibility of tobacco etch virus (Pirone & Thornbury, 1983). By comparing coat protein sequences of several aphid-transmissible and aphid non-transmissible potyviruses, Harrison & Robinson (1988) proposed that a conserved aspartic acid-alanine-glycine (DAG) sequence is required for aphid transmission. This sequence is on an exposed, outer part of the capsid (Allison et al., 1985). Consistent with the above hypothesis, changes in the DAG sequence have been found in aphid non-transmissible strains of tobacco vein mottling virus (TVMV) (to DAE; Atreya et al., 1990), plum pox virus (to DAL; Maiss et al., 1989) and papaya ringspot virus (to DTG; Quemada et al., 1990a). Other differences between coat protein sequences of aphid-transmissible and aphid non-transmissible strains have been found in all but TVMV. Recently, Atreya et al. (1990) clearly demonstrated the requirement for the G of the DAG sequence by showing that a single amino acid change (DAG to DAE) in the coat protein gene in a full-length infectious clone of TVMV completely eliminated aphid transmissibility.

Several strains of soybean mosaic virus (SMV) have been defined on the basis of aphid transmissibility (Lucas & Hill, 1980), symptoms induced on different soybean varieties (Buzzel & Tu, 1984; Cho & Goodman, 1979) and serological relationships (Hill et al., 1989). The isolates identified by phenotypic response on differential soybean lines have been grouped into strains G1 to G7, G7a and C14. The coat protein sequence of an aphid non-transmissible strain (N) of SMV has been published (Eggenberger et al., 1989). It lacks the DAG sequence. Here we present the coat protein sequences of two SMV isolates which were identified as members of the G2 (Ia 75-16-1, unpublished results) and G7 strains as described previously (Hill et al., 1989). For the purposes of this report, they will be referred to as strains G2 and G7. Both of these strains are aphid-transmissible. The aphid transmission efficiencies of strains G2 and G7 (determined as described by Lucas & Hill, 1980), using Myzus persicae and an acquisition access period of 3 min, were 50% and 53%, respectively. Experiments employed 30 Williams soybean plants and 30 virus-free apterous aphids per plant.

The two SMV strains were purified according to Hill & Benner (1980) and RNAs were purified according to Vance & Beachy (1984). cDNA was synthesized by the method of Gubler & Hoffman (1983) using a kit (Pharmacia) and cloned into pGEM3Z(f−) vector (Promega). The cDNA clones were sequenced with Taq polymerase (Promega) by the dideoxynucleotide method (Sanger et al., 1977) using ssDNA produced by super-infecting the transformed cells with the helper phage M13K07 (Vieira & Messing, 1987). Six independent clones spanning portions of the coat protein gene and 3' non-coding region were sequenced for each strain. At least one clone of each strain spanned the entire coat protein gene and 3' non-coding region. Every base was determined from at least two independent clones. No
base differences were detected among independent clones of the same strain.

The coat protein genes of both SMV G2 and SMV G7 are the 3'-terminal 795 nucleotides of a larger open reading frame (ORF). The coat protein is 265 residues long, starting from the serine residue at position 1. Eggenberger et al. (1989) located the coat protein amino terminus by direct amino acid sequencing of the N strain. The ORFs in G2 and G7 were followed by a 259 nucleotide non-coding region [excluding the poly(A) tail]. The nucleotide sequence of the coat protein gene and the 3' non-coding regions of SMV G2, and the nucleotides and amino acids which differ in SMV G7 and SMV N are aligned in Fig. 1. The SMV G2 and SMV N strains are highly similar, having only three nucleotide differences (all in the coat protein gene) and one amino acid difference. Comparison of SMV G7 with SMV N shows several more nucleotide substitutions and four amino acid differences.

The amino acid common to SMV G2 and SMV G7, but not to SMV N, is a glycine at position 12 which is replaced by aspartic acid in SMV N. This glycine represents the third amino acid of the tripeptide (DAG) that is thought to be involved in aphid transmission (Harrison & Robinson, 1988). Thus, the results presented here are consistent with their model and the evidence of Atreya et al. (1990) in that at least the G of the DAG is correlated with aphid transmissibility. The results presented here are strictly correlative. Many differences elsewhere in the genome, in the HC gene in particular, may also affect aphid transmissibility.

The very high conservation of coat protein gene sequences of the three SMV strains is remarkable in view of their differing biological properties. The three strains have been shown to differ with respect to aphid transmissibility. In addition, strains G2 and G7 can be differentiated by immunoreactive patterns of peptide maps of the virion coat protein (Hill et al., 1989) and by symptoms induced on various soybean cultivars (Cho & Goodman, 1979). The presence of only three amino acid differences between G2 and G7 is of interest because only G7 can infect soybeans containing the Rsv1 resistance gene (Buss et al., 1988). The elucidation of the role of the amino acid differences between N and the aphid-transmissible isolates, and between G2 and G7 awaits construction of full-length clones from which infectious transcripts containing defined sequence differences can be obtained (e.g. Atreya et al., 1990).
Table 1. Percentage nucleotide sequence similarity between the coat protein coding (above diagonal) and 3′ non-coding (below diagonal) regions of SMV G2, G7, N and WMV 2

<table>
<thead>
<tr>
<th></th>
<th>SMV G7</th>
<th>SMV G2</th>
<th>SMV N*</th>
<th>WMV 2†</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMV G7</td>
<td>–</td>
<td>96.1</td>
<td>96.2</td>
<td>77.8</td>
</tr>
<tr>
<td>SMV G2</td>
<td>96.0</td>
<td>–</td>
<td>99.6</td>
<td>79.0</td>
</tr>
<tr>
<td>SMV N*</td>
<td>93.0</td>
<td>100.0</td>
<td>–</td>
<td>79.5</td>
</tr>
<tr>
<td>WMV 2†</td>
<td>77.6</td>
<td>79.1</td>
<td>79.5</td>
<td>–</td>
</tr>
</tbody>
</table>

* Eggenberger et al. (1989).
† Frenkel et al. (1989).

Frenkel et al. (1989) and Quemada et al. (1990b) suggested that watermelon mosaic virus 2 (WMV 2) and SMV N may be strains of the same potyvirus because the coat protein and 3′ non-coding regions were more similar than other pairwise comparisons of different potyviruses. However, the coat protein gene and the 3′ non-coding regions of SMV G2, SMV G7 and SMV N are much more similar to each other than to the analogous regions of WMV 2 (Table 1). We propose that WMV 2 be classified as a virus distinct from SMV. This is supported by biological properties such as major differences in host range. The known isolates of SMV are reported to infect species in only five plant families. In contrast, isolates of WMV 2 are reported to infect species in 22 different plant families (Edwardson & Christie, 1986). Accurate classification awaits determination of the nucleotide sequences of the complete genomes of these viruses and a thorough comparison of their biological properties.

The authors thank Lou Mansky for helpful suggestions and discussion. This research was supported in part by Pioneer Hi-Bred International, Inc., Johnston, Iowa. Journal Paper No. J-14179 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Project No. 2428.

References


(Received 6 September 1990; Accepted 19 December 1990)