Molecular characterization of the garlic virus X genome

Sang Ik Song,1 Jong Tae Song,1 Chung Ho Kim,2 Jong Seob Lee3 and Yang Do Choi1

1 Department of Agricultural Chemistry, Seoul National University, Suwon 441-744, Korea
2 Department of Food and Nutrition, Seowon University, Chongju 361-140, Korea
3 Department of Molecular Biology, Seoul National University, Seoul 151-742, Korea

The complete nucleotide sequence of the cDNA genome for garlic virus X (GVX), one of the major viruses infecting garlic plants, was determined. GVX is a single-stranded positive-sense RNA virus consisting of 8106 nucleotides excluding the 3′-end poly(A) tail and contains six open reading frames (ORFs) which encode putative proteins of 174 kDa (ORF1), 26 kDa (ORF2), 12 kDa (ORF3), 32 kDa (ORF4), 26 kDa (ORF5) and 15 kDa (ORF6). The putative viral proteins show similarity to those of carlaviruses and potexviruses but show the highest homology to shallot virus X (ShVX). Even though the GVX genome contains most of the structural elements common to carlaviruses and potexviruses, it is distinguished from them by the presence of an ORF4 which encodes an unusual protein. These results suggest that GVX may belong to an unassigned group of ShVX and GarV-type viruses rather than to the carlaviruses or potexviruses.

Most of the garlic plants cultivated throughout the world are infected with various kinds of viruses which give rise to mosaic or streak symptoms on their leaves and thus reduction of bulb and clove size (Ahlawat, 1974). In attempting to control these viral diseases, the first essential step is to establish the identity of the viruses responsible and to determine their molecular characteristics. There is no clear method for identification of garlic viruses, but garlic plants are usually infected with various viruses including potex-, carla-, GarV-type and potyviruses (Chang et al., 1988; Choi et al., 1995; Kobayashi et al., 1996; Lee et al., 1979; Marys et al., 1994; Nagakubo et al., 1994; Song et al., 1995; Sumi et al., 1993; Tsuneyoshi & Sumi, 1996; Van Dijk, 1993; Yamashita et al., 1996). A potyvirus, termed garlic mosaic virus (GMV), was isolated and their nucleotide sequences determined (Ryabov et al., 1996; Sumi et al., 1993; Yamashita et al., 1996). These partial cDNAs revealed that the GarV-type viruses share structural similarity with shallot virus X (ShVX) (Kanyuka et al., 1992), which belongs to an unassigned new virus group closely related to the carla- and potexviruses (Tsuneyoshi & Sumi, 1996). This new type of virus seems to be responsible for the mosaic symptoms in garlic plants observed in Japan and Europe. In this paper, we have determined the complete nucleotide sequences of cDNAs for one of the major viruses infecting garlic plants showing mosaic or streak symptoms.

To identify this virus, a cDNA library was constructed using viral RNA isolated from garlic plants. The double-stranded cDNAs of the garlic viruses were synthesized with SuperScript II reverse transcriptase according to the GIBCO BRL manual using oligo(dT)12-18 as a primer. Blunt-ended cDNA was ligated into an Smal-digested pUC18 plasmid and transformed into Escherichia coli MC1061. Recombinant colonies were randomly chosen and their nucleotide sequences were determined. One of these clones, GVX53, contained a cDNA insert of about 5500 bp and showed nucleotide sequence homology with ShVX. The virus from which the cDNA clone GVX53 was derived was named garlic virus X (GVX) in this study. When colony hybridization was carried out using GVX53 as a molecular probe, up to 28% of the colonies hybridized with the probe. Northern and immunoblot analyses showed that almost all of the tested garlic plants are infected with GVX (data not shown). These results suggest that GVX is one of the major viruses infecting garlic plants.

To isolate cDNA clones for the entire genome of GVX, overlapping cDNA clones were isolated from a cDNA library prepared with viral RNA from a single garlic plant collected in Wonju, Korea. The size and relationship of the selected

Author for correspondence: Yang Do Choi.
Fax +82 331 291 7011; e-mail choyngd@plaza.snu.ac.kr

The GenBank accession number of the sequence reported in this paper is U89243.
recombinant cDNA clones for GVX were determined by cross-hybridization and restriction enzyme analysis.

To elucidate the genome structure of GVX RNA, nucleotide sequences of the overlapping clones pGVX37, pGVX53 and pGVX46 were determined. Dideoxynucleotide sequencing of both strands of dsDNA templates was carried out using [32P]dATP and modified T7 DNA polymerase (Sequenase, US Biochemical). The extreme 5’ end of GVX RNA was determined by primer extension using oligonucleotide GVX5N (5’ AGTATCCCTATCAAGTAGTG 3’), complementary to the sequence 46 to 65 residues from the 5’ end (Sambrook et al., 1989). The nucleotide sequences of the overlapping cDNA clones, along with that of the extreme 5’ end of GVX RNA as determined by primer extension, indicated that the full size of the genome for GVX is 8106 nucleotides excluding the 3’-end poly(A) tail.

Analysis of the nucleotide sequence for GVX RNA based on the cDNA clones showed the presence of six open reading frames (ORFs). ORF1 encodes a 174 kDa polypeptide of 1543 residues, ORF2 encodes a 26 kDa polypeptide of 234 residues, ORF3 encodes a 12 kDa polypeptide of 106 residues and ORF4 encodes a 32 kDa polypeptide of 288 residues. There is an additional putative ORF between ORF3 and ORF4, albeit without an initiation codon. ORF5 encodes a 26 kDa polypeptide of 243 residues and ORF6 encodes a 15 kDa polypeptide of 127 residues. The organization of ORFs encoding the putative viral proteins is similar to that of carlavirus and potexviruses but more specifically to ShVX (Fig. 1).

The amino acid sequence of the 174 kDa polypeptide encoded by GVX ORF1 showed high similarity in three extensive regions with that of virus-specific RNA replicases of potexviruses, carlaviruses, tymoviruses and closteroviruses (Kanyuka et al., 1992). Significant overall similarity was observed between GVX ORF1 and RNA replicases of ShVX and potexviruses but not with those of carlaviruses (Table 1).

The sizes of putative virus-specific RNA replicases of potexviruses, carlaviruses and tymoviruses are 147–186 kDa.

Table 1. Deduced amino acid sequence similarity between GVX encoded proteins and the corresponding proteins of ShVX, GarV-types, carlaviruses and potexviruses

<table>
<thead>
<tr>
<th>Virus</th>
<th>ORF1</th>
<th>ORF2</th>
<th>ORF3</th>
<th>ORF4</th>
<th>CP</th>
<th>ORF6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unassigned virus</td>
<td>57±8</td>
<td>40±6</td>
<td>68±5</td>
<td>55±1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GarV-A</td>
<td>NA</td>
<td>NA</td>
<td>57±8</td>
<td>40±6</td>
<td>68±5</td>
<td>55±1</td>
</tr>
<tr>
<td>GarV-B</td>
<td>NA</td>
<td>NA</td>
<td>68±9</td>
<td>50±5</td>
<td>84±4</td>
<td>79±4</td>
</tr>
<tr>
<td>GarV-C</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>69±9</td>
<td>69±3</td>
<td></td>
</tr>
<tr>
<td>GarV-D</td>
<td>NA</td>
<td>NA</td>
<td>63±6</td>
<td>59±1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ShVX</td>
<td>66±0</td>
<td>59±4</td>
<td>55±4</td>
<td>34±0</td>
<td>63±2</td>
<td>63±8</td>
</tr>
<tr>
<td>Carlavirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVB</td>
<td>30±8</td>
<td>41±9</td>
<td>26±9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSV</td>
<td>30±2</td>
<td>36±2</td>
<td>24±2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVM</td>
<td>28±0</td>
<td>36±5</td>
<td>28±7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVS</td>
<td>30±0</td>
<td>33±0</td>
<td>25±0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potexvirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMV</td>
<td>37±0</td>
<td>23±2</td>
<td>48±4</td>
<td>23±7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LVX</td>
<td>31±0</td>
<td>38±3</td>
<td>32±2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVX (X3)</td>
<td>37±0</td>
<td>33±4</td>
<td>35±8</td>
<td>30±4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMYEAV</td>
<td>40±0</td>
<td>26±3</td>
<td>37±0</td>
<td>29±7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 2. Putative ORF3-4 of GVX between ORF3 and ORF4. The deduced amino acid sequences of ORF3 and ORF4 are shown under the nucleotide sequence. The putative ORF3-4 is shown lightly shaded. Terminator codons are denoted by the asterisks (*). Two possible alternative initiation codons (UUG, AUU) are underlined. The conserved amino acid sequence (C-V-I-X-I-T-G-E-S-X2-I-X-G-C) in ORF4 of carla- and potexviruses (Foster & Mills, 1991) is shown in the box beneath the amino acid sequence.

219–223 kDa and 194–210 kDa, respectively. The size of the putative RNA replicase of closteroviruses is 217 kDa. The size of GVX ORF1 was about the same as those of RNA replicases of smaller potexviruses. It is concluded, therefore, that GVX ORF1 encodes a virus-specific RNA replicase which is more closely related to those of potexviruses than carlaviruses.

ORF2 encodes a 26 kDa polypeptide showing similarity to the first ORF of the triple gene block found in all carla- and potexviruses. A highly conserved motif, G-C-G-K-S, which is conserved in NTP-dependent DNA helicases including those of carla- and potexviruses, was found (Zimmern, 1987). The same motif was also identified within the 174 kDa replicase (ORF1). The duplication of this motif has been reported for members of the carla- and potexviruses (Forster et al., 1988; Zavriev et al., 1991).

ORF3 encodes a 12 kDa polypeptide homologous to the second ORF of the triple gene block proteins of all carla- and potexviruses. The putative proteins encoded by the triple gene block of white clover mosaic virus (WCIMV) were shown to be required for virus transport in plant cells (Beck et al., 1991). The amino acid sequence similarities in ORF3s are about 50–60% within carlaviruses, about 40% within potexviruses, and about 30–40% between carla- and potexviruses. As shown in Table 1, the protein encoded by GVX ORF3 shows about 33–48% similarity with those of carla- and potexviruses, and 55, 58 and 69% similarity with the corresponding protein of ShVX, GarV-A and GarV-B, respectively. Partial cDNA clones for GarV-A to -D were isolated from virus-diseased garlic plants in Japan and showed close homology with those from ShVX (Sumi et al., 1993). Recently, GarV-A to -D were classified (Tsuneyoshi & Sumi, 1996) under ‘unassigned viruses’ in the ICTV Sixth Report. This result indicates that ORF3s of GVX, ShVX, GarV-A and GarV-B are rather different from those of carla- and potexviruses.

The polypeptide encoded by GVX ORF4 was much larger than those of ORF4, the third ORF in the triple gene block, of carla- and potexviruses, and showed no similarity to them. It showed 34–50% similarity only to the proteins encoded by ORF4s of ShVX, GarV-A or GarV-B (Table 1). This putative polypeptide is extremely rich in serine (9±4%) and threonine (10±4%) and seems to be the hallmark of GVX, ShVX, GarV-A and GarV-B. The structural features of ORF4 distinguish these viruses from carla- and potexviruses.

Between ORF3 and ORF4 of the GVX genome, however, a putative ORF (named ORF3-4) was detected, even though no initiation codon was found (Fig. 2). It codes for a polypeptide of 114 amino acids containing a highly conserved sequence, C-V-I-X-I-T-G-E-S-X2-I-X-G-C, present in the 7–8 kDa proteins of ORF4 in all of the carla- and potexviruses (Foster & Mills, 1991). Similar structural features were also found in the genomes of lily virus X (LVX; Memelink et al., 1990) and ShVX. The ORF3-4 polypeptides of GVX, ShVX and LVX may be initiated through the use of an unusual initiation codon, if they are expressed at all. Alternative initiation codons (AUU, CUG, UUG) are known to be used in plants (Gordon et al., 1992; Jelkmann et al., 1992; Petty & Jackson, 1990). As shown in Fig. 2, there are five possible alternative initiation codons in the putative ORF3-4 of GVX. On the basis of sequence alignment with ORF4s of carla- and potexviruses, the possible initiation codon could be UAU (49th codon) or AUU (51st codon), which would give rise to a polypeptide of about 7 kDa. AUU was assigned as a possible initiation codon for the ORF2 protein of strawberry mild yellow edge-associated virus, a potexvirus (SMYEAV; Jelkmann et al., 1992). These
results suggest that ORF3-4 may be a homologue of the third gene in the triple gene blocks of carla- and potexviruses and the third gene in the triple gene block of carla- and potexviruses, ORF4, seems to have been transformed into the putative ORF3-4 of GVX.

ORF5 encodes the coat protein (CP) of GVX, which has 63–84% amino acid sequence similarity with the CP of ShVX and GarV-A to -D (Table 1). Even though GVX CP shows closest similarity (84%) with that of GarV-B, the total similarity over ORF3 to ORF6 between the two viruses (where the nucleotide sequence of GarV-B was available) was only 55% for the nucleotide sequence and 68% for the amino acid sequence. This suggests that GVX is a separate member of the unassigned virus group that includes ShVX and GarV-A to -D, rather than a Korean isolate of GarV-B. This conclusion is supported by the recent reports of Ryabov et al. (1996) and Yamashita et al. (1996). They determined the nucleotide sequence of a partial cDNA clone from mite-borne filamentous virus (MbFV) of garlic. It showed 95% identity to GarV-D over the region ORF4 to ORF6. They concluded that MbFV was a European isolate of Japanese GarV-D.

The ORF6 polypeptide of GVX contained a highly conserved region comprising a basic arginine-rich domain and a putative zinc-finger motif. This is a typical feature of polypeptides encoded by ORF6 in all carlaviruses. The 11 kDa polypeptide of carlaviruses is speculated to be involved in regulation of host gene transcription and/or viral RNA replication through binding to nucleic acids (Gramsat et al., 1990). The presence of the ORF6 11 kDa polypeptide between the CP gene and the poly(A) tail sets the carlavirus group apart from the potexvirus group (Foster, 1992). However, no significant similarity was observed between the ORF6 protein of GVX and those of carlaviruses, although a high degree of similarity was maintained among ORF6 proteins of GVX, ShVX and GarV-A to D (Table 1). These results also argue that GVX, ShVX and GarV-A to D are different from carla- and potexviruses.

We conclude that GVX does not belong to the carla- or potexviruses but is a separate member of a new and unassigned group of plant viruses that includes ShVX and GarV-type viruses. This is the first report of a complete nucleotide sequence of a cDNA for one of the major viruses infecting garlic plants but further studies are necessary to understand the pathogenicity of the virus.

The present investigation was supported by grants from the Korea Science and Engineering Foundation through the Research Center for New Biomaterials in Agriculture to Y.D.C., and through the Research Center for Cell Differentiation to J.S.L. and in part by a grant from the Ministry of Science and Technology, Korea.

References


Nagakubo, T., Kubo, M. & Oeda, K. (1994). Nucleotide sequence of the 3′ regions of two major viruses from mosaic-diseased garlic: molecular


Received 30 July 1997; Accepted 24 September 1997