Complete Nucleotide Sequence and Genome Organization of Broad Bean Wilt Virus 2*

Shigeo NAKAMURA**, Takayoshi IWA1**,*** and Ryoso HONKURA**,†

Key words: fabavirus, nucleotide sequence, genome organization.

Broad bean wilt virus (BBWV) is a member of the fabavirus group, which has recently been classified into two species BBWV-1 and BBWV-2. The former is the type member of the Genus Fabavirus⁴). BBWV has a wide host range among dicotyledons and some families of monocotyledons and is widely distributed throughout the world¹³). BBWV has a bipartite genome, consisting of two positive-strand RNA molecules designated as RNA1 (Mr 2×10⁶) and RNA2 (Mr 1.5×10⁶). Both genomic RNAs are separately encapsidated into two species of icosahedral particles with different sedimenting properties, built up by 42 kDa and 26 kDa coat proteins⁵). Fabaviruses have been included in the Family Comoviridae together with the comoviruses and the nepoviruses. Both the comoviruses and the nepoviruses have been well investigated, whereas the fabaviruses have not been characterized at the molecular level⁶). We present the complete nucleotide sequences of both BBWV-2 RNA1 and RNA2, and compare the deduced translation products of BBWV-2 with those of related viruses.

The MB7 isolate of BBWV-2, originally isolated from broad bean (Vicia faba) leaves with showing mosaic symptoms, was propagated in broad bean⁸). Broad bean leaves infected with MB7 were homogenized in two volumes (v/w) of 0.1M phosphate buffer (pH 7.0) containing 10 mM EDTA and 0.5% 2-mercaptoethanol and in one volume (v/w) of carbon tetrachloride. After centrifugation of the homogenate at 9000×g for 10 min, 1% Triton X-100 was added to the aqueous phase. The clarified solution was layered onto a 20% (w/v) sucrose cushion, and centrifuged at 80,000×g for 2 hr. The pellet was dissolved in phosphate buffer and further purified by repeated differential centrifugation.

Virus preparation was treated with SDS and proteinase K followed by SDS-phenol to extract viral RNA⁸). Purified viral RNAs were reverse-transcribed by priming with oligo (dT) or random hexanucleotides using TimeSaver cDNA Synthesis Kit (Pharmacia). The cDNAs were cloned into the EcoRI site of pBluescript SK+ (Stratagene). Cloning of the 5' ends was performed

---

* The nucleotide sequence data reported in this paper will be appear in the DDBJ/EMBL/GeneBank nucleotide sequence databases under the accession numbers AB013615 and AB013616.
** Miyagi Prefecture Agricultural Research Center, Takadate-Kawakami, Natori 981-1243, Japan 宮城県農業センター
*** Present address: National Institute of Agrobiological Resources, Kan-nondai, Tsukuba 305-8602, Japan 現在：農業生物資源研究所
† Present address: Sendai Agricultural Extension Center of Miyagi Prefecture, Tsutsumidori-amamiyamachi, Aoba-ku, Sendai 981-0914, Japan 現在：仙台地域農業改良普及センター
using the 5' RACE System (GIBCO BRL). The cDNA clones covering the complete genome were sequenced and analysed as described previously.

RNA1 is 5957 nucleotides in length, excluding the 3' poly(A) tail and contains a single long open reading frame (ORF) of 3613 nucleotides extending from nucleotide 235 to 5847 (Fig. 1). The protein encoded by the ORF, consisting of 1870 amino acids, has a predicted Mr of 210 kDa. The amino acid sequence of the RNA1-coded 210 kDa protein was compared with that of cowpea mosaic comovirus (CPMV), cowpea severe mosaic comovirus (CPSMV) and red clover mottle comovirus (RCMV), revealing that the size and genetic organization of BBWV-2 RNA1 is similar to that of the comovirus. As shown in Fig. 1, three conserved motifs, the NTP-binding, the protease and the RNA-dependent RNA polymerase, were identified at positions similar to those of comoviruses. The consensus sequences of the motifs were well conserved among BBWV-2 and three comoviruses (Fig. 2). In rabbit reticulocyte lysate, BBWV-2 RNAs directed the large polypeptide of approximately 200 kDa (Fig. 3). These results suggest that the 210 kDa protein encoded by BBWV-2 RNA1 is a polyprotein and is proteolytically cleaved by a virus-encoded protease to yield the mature viral proteins as described for comoviruses. For comoviruses, four dipeptides, Q/S, Q/A, Q/M and Q/G are shown to be cleavage sites of polyproteins by virus-coded protein. In the 210 kDa protein of BBWV-2, a Q/G at position 1171-1172 is the most probable cleavage site of the protease/polymerase junction because of the resulting size of the mature polymerase. The predicted polymerase domain at position 1172-1870 of BBWV-2 shows 39.8%, 40.7%, 41.0% and 29.8% identities with the corresponding regions of CPMV, CPSMV, RCMV and grapevine fanleaf nepovirus (GFLV), respectively, whereas identities between the N-terminal domain at position 1-1171 of BBWV-2 and the corresponding regions of other viruses are 28.6% (CPMV), 28.5% (CPSMV), 27.4% (RCMV) and 21.5% (GFLV). These results indicate that BBWV-2 is closely related to the comovirus rather than the nepovirus.

RNA2 is 3589 nucleotides in length excluding the 3' poly(A) tail. A single long ORF of 3195 nucleotides extending from nucleotide 222 to 3416 is present (Fig. 1). The protein encoded by the ORF is 1064 amino acids in length with a predicted Mr of 119 kDa. All comovirus RNA2s sequenced to date have the second in-phase AUG codon producing smaller polyprotein. In the BBWV-2 RNA2, such a second in-phase AUG is also present at position 607-609. Whether the second AUG serves as the initiation of the smaller polyprotein is unknown. In rabbit reticulocyte lysate, however, the 105 kDa polypeptide, which agrees with the calculated Mr of the second AUG product, is favored over the 120 kDa polypeptide, which is expected for the first AUG product (Fig. 3). Two proteolytic processing sites, a Q/G at position 465-466 and a Q/A at position 867-868, within the 119 kDa protein have been deduced by comparison with that of patchouli mild mosaic fabavirus. These results suggest that the BBWV-2 RNA2 is translated into a polyprotein and is divided into three mature proteins. Of the resulting three proteins, central (44 kDa) and C-terminal (22 kDa) proteins are predicted to be the

---

**a) NTP-binding domain**

| BBWV-2 | 515 GASGCTGKTSVQVRVAI--20aa--PFWSGYRHAHVVTYDDFGAV |
| CPMV  | 494 GKSRTGKSLSNQTR--20aa--QYWSGYRQFPLMDFDAV |
| CPSMV | 482 GARSQKGTVINROLE--20aa--DYWSGYRQFPIV1DDFAAV |
| RCMV  | 483 GNRSQGKSL1GKLIS--20aa--THWSQGRQFIV1DDFAAV |

**b) Cysteine protease domain**

| BBWV-2 | 1003 H--35aa--E--77aa--LYEASTVAGDCSLILAEIDGK1KLVG |
| CPMV  | 987 H--35aa--E--78aa--LYEAPITIPEDCDSLVIHAGGK1KVVG |
| CPSMV | 976 H--36aa--E--79aa--QVDVSTAEDCSLIIAT1DGK1K11G |
| RCMV  | 983 H--35aa--E--78aa--IRYEA1TVMSDCSM1TINVGK1TVVG |

**c) Polymerase domain**

| BBWV-2 | 1450 CDYSRFDGFL--49aa--GIPSFQPLTVNVS--33aa--YGDN11ISV |
| CPMV  | 1434 CDY1SFDGGL--49aa--GIPSFQ1PTMNVS--36aa--YGD111111SV |
| CPSMV | 1424 CDY1SFDGGL--49aa--GIPSFQ1PTMNVS--36aa--YGD111111SV |
| RCMV  | 1431 CDY1SFDGGL--49aa--GIPSFQ1PTMNVS--36aa--YGD111111SV |

---

Fig. 2. Alignment of the four domains identifying the NTP-binding protein (a), the cysteine protease (b) and the RNA-dependent RNA polymerase (c) between BBWV-2, CPMV, CPSMV and RCMV. Asterisks indicate identical residues in all four sequences. Numbers to the left of the sequence refer to amino acid residue position in each viral polyprotein.
coat proteins, because the molecular weights calculated from amino acid sequences well agree with the values estimated by SDS-PAGE of purified virion, 40 kDa and 24 kDa, respectively. The overall amino acid identities of the BBWV-2 119 kDa protein compared with those of comoviruses are 21.1% (CPMV14)), 21.7% (CPSMV1)), 20.0% (RCMV11)) and 21.9% (bean pod mottle comovirus, BPMV7)). In this report, the 5' ends of both BBWV-2 RNA1 and RNA2 were cloned by the RACE method, in which the 1st strand cDNAs were tailed with dC by terminal deoxynucleotidyl transferase. Therefore, the extreme 5' G residue(s) may be missing, and in the case of G: C rich cDNA, non-specific priming of second strand cDNA synthesis with the deoxyinosine-containing anchor primer may result in truncated products. To determine the sequences of the extreme 5' ends of the RNAs, the dA-tailing RACE and/or the direct RNA sequencing should also be employed. The genome organization of BBWV-2 RNA1- and RNA2-encoded proteins shown in Fig. 1 is hypothetical. Because none of the potential cleavage sites have been confirmed experimentally, the exact sizes and proteolytic cleavage sites of viral gene products need to be confirmed both in vitro and in vivo.

We thank Drs. Y. Honda and Y. Mikoshiba of the National Agricultural Research Center for helpful advice in classifying the MB7 isolate. We are also grateful to Dr. T. Kuroda of Iwate Biotechnology Research Center for useful suggestions and discussion.

Literature cited
和文摘要

中村茂雄・岩井孝尚・本澤直三：ソラマメウイルス2 RNAの全塩基配列

宮城県内のモザイク症状ソラマメから分離したソラマメウイルス2（BBWV-2）MB7株のゲノム RNAをクローニングし、その全塩基配列を決定した。RNA1はポリ（A）鎖を除いて5847塩基から成り、210 kDa（1870 アミノ酸残基）のタンパク質をコード可能な一つの大きなORFが存在した。このORF産物はコモウイルス、次いでネボウイルスのRNA1翻訳産物と同様性が高く、NTP結合モチーフ、プロテアーゼモチーフ、ポリメラーゼモチーフがコモウイルスのそれとほぼ同じ位置に見いだされた。ポリメラーゼ領域の同様性はコモウイルスと39.8～41.0%であった。RNA2はポリ（A）鎖を除いて3589塩基から成り、119 kDa（1064 アミノ酸残基）のタンパク質をコードすることが可能な一つの大きなORFが存在した。このORF産物はコモウイルスのRNA2翻訳産物との同様性は20.0～21.9%と低かった。

(Received May 20, 1988; Accepted July 23, 1998)