2004

GENETIC DIVERSITY AND SYMPTOM SEVERITY DETERMINANTS OF BEAN POD MOTTLE VIRUS

Hongcang Gu
University of Kentucky, hgu2@uky.edu

ABSTRACT OF DISSERTATION

Hongcang Gu

The Graduate School

University of Kentucky

2004

GENETIC DIVERSITY AND SYMPTOM SEVERITY DETERMINANTS OF *BEAN POD MOTTLE VIRUS*

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the College of Agriculture
at the University of Kentucky

By

Hongcang Gu

Lexington, Kentucky

Director: Dr. Said A. Ghabrial, Professor of Plant Pathology

Lexington, Kentucky

2004

ABSTRACT OF DISSERTATION

GENETIC DIVERSITY AND SYMPTOM SEVERITY DETERMINANTS OF BEAN POD MOTTLE VIRUS

Bean pod mottle virus (BPMV), a member of the genus *Comovirus* in the family *Comoviridae*, is widespread in the major soybean-growing areas in the United States. Soybean yield losses of 10-40% have been reported as a consequence of BPMV infection. The complete nucleotide sequences of two strains, K-Ha1 and K-Ho1, were determined. Field isolates of BPMV were classified into two distinct subgroups (I and II) based on slot blot hybridization and sequence analyses. Full-length cDNA clones from which infectious transcripts can be produced were constructed for strains K-G7, K-Ho1 and K-Ha1. Whereas strains K-Ha1 and K-G7 induced mild or moderate symptoms in infected soybean plants, strain K-Ho1 produced very severe symptoms. Symptom severity was mapped to RNA1. Chimeric RNA1 constructs were generated by exchanging full or partial coding regions of the five RNA1-encoded mature proteins between the full-length cDNA clones of the three RNA1s and the resultant transcripts were inoculated onto soybean. The results showed that the coding regions of the protease co-factor (Co-pro) and the putative helicase (Hel) are determinants of symptom severity. Although symptom severity correlated well with accumulation of viral RNA, neither the Co-pro nor Hel protein could be demonstrated as a suppressor of RNA silencing. Furthermore, separate expression of the Co-pro or Hel proteins from a PVX vector induced necrosis on the inoculated leaves of *Nicotiana benthamiana*.

Characterization of BPMV K-Ho1 indicated that it is a diploid reassortant, containing two distinct types of RNA1s and one type of RNA2. Examination of field isolates from various locations in the United States and Canada revealed that diploid reassortants are of frequent occurrence in natural populations of BPMV. The vary severe symptoms induced by BPMV K-Ho1 can be mimicked by inoculation of plants with a mixture of RNA1 transcripts from two distinct strain subgroups and RNA2 transcript from either subgroup. Plants inoculated with a mixture of transcripts containing two types of RNA1 from the same strain subgroup did not produce very severe symptoms. These are due to interactions between two distinct types of RNA1s.

At present, no soybean cultivars with resistance to BPMV are commercially available. Therefore, the feasibility of cross protection as an alternative disease management strategy was studied. Two mild strains of BPMV (K-Da1 and K-Ha1), belonging to subgroup II, were tested for their ability to protect infected plants against a severe strain (K-Ho1). Inoculation of the soybean cultivar Essex on the primary leaves with either of the two mild strains conferred complete protection against challenge inoculation with the severe strain K-Ho1, regardless of the timing of challenge inoculation. Cross-protection was evident regardless of whether virions or BPMV-RNA were used as inocula. Cross protection was independent of the soybean cultivar used and method of virus inoculation, sap-inoculation or by the bean leaf beetle, vector of BPMV. Protection was complete and durable.

KEYWORDS: Bean pod mottle virus, diploid reassortments, cross-protection, infectious cDNA clones, symptom severity determinants.

Hongcang Gu

Date: 07/27/04

GENETIC DIVERSITY AND SYMPTOM SEVERITY DETERMINANTS OF *BEAN POD MOTTLE VIRUS*

By

Hongcang Gu

Said A. Ghabrial
Director of Dissertation

Lisa J. Vaillancourt
Director of Graduate Studies

Date: 07/27/04

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctors degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patron is expected to secure the signature of each user.

DISSERTATION

Hongcang Gu

The Graduate School

University of Kentucky

2004

GENETIC DIVERSITY AND SYMPTOM SEVERITY DETERMINANTS OF *BEAN POD MOTTLE VIRUS*

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Agriculture at the University of Kentucky

By

Hongcang Gu

Lexington, Kentucky

Director: Dr. Said A, Ghabrial, Professor of Plant Pathology

Lexington, Kentucky

2004

ACKNOWLEDGEMENTS

I would like to express my great appreciation to my major advisor, Dr. Said A. Ghabrial, for his guidance and financial support. He always encouraged me whenever I met with difficulties in the past five years. He is not only the mentor of my academic research, but also a friend of my personal life. Without his guidance and encouragement, this project could not have been accomplished. Sincere gratitude is also expressed to the members of my Advisory Committee: Dr. Christopher L. Schardl, Dr. Peter D. Nagy, Dr. Arthur G. Hunt, Dr. Lisa J. Vaillancourt and my outside examiner Dr. John C. Snyder. Each individual input their valuable time and thereby provided insights and supports to my work. I am extremely grateful to Dr. David A. Smith for his encouragement of my research work and strong backup when I needed help in my personal life.

I would like to extend my gratitude to Ms. Wendy Havens, for her technical support, great assistance and friendship. I wish to thank Ms. Amy Crume, for her greenhouse work. In addition, I would like to thank K. S. Rajendran, Ana Soldevia, Tianyong Zhao, Renyuan Wang, Chunquan Zhang, and Anat Kritzman**,** for their suggestions and friendship. Many other people in the Department of Plant Pathology, although their names are not mentioned here, helped me in one way or another and to them I am really also grateful.

Special gratitude extends to my wife, Liping and my son Yu Gu, my parents, and my siblings. Without their love and sacrifice, I cannot imagine how I would have finished my work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FILES

Hongcang.pdf ……..13 MB

Chapter I

Literature review

Introduction

The genus *Comovirus* in the family *Comoviridae* includes 15 approved members, *Andean potato mottle virus* (APMoV), *Broad bean stain virus* (BBSV), *Bean rugose mosaic virus* (BRMV), *Broad bean true mosaic virus* (BBTMV), *Bean pod mottle virus* (BPMV), *Cowpea mosaic virus* (CPMV), *Cowpea severe mosaic virus* (CPSMV), *Glycine mosaic virus* (GMV), *Pea green mottle virus* (PGMV), *Pea mild mosaic virus* (PMiMV), *Quail pea mosaic virus* (QPMV), *Radish mosaic virus* (RaMV), *Red clover mottle virus* (RCMV), *Squash mosaic virus* (SqMV) and *Ullucus virus C* (UVC) (Welink, et al., 2000). The genomes of comoviruses are bipartite consisting of two positive sense single-stranded RNA molecules that are separately encapsidated in icosahedral particles (T=1, pseudo T=3) 28-30 nm in diameter. Three sedimenting components, designated top, middle and bottom components are resolved when purified virions are subjected to density gradient centrifugation. The top (T) sedimenting component contains only empty capsids, whereas middle (M) bottom (B) components contain single molecules of RNA2 (M-RNA) and RNA1 (B-RNA), respectively (Welink *et al.*, 2000). The comovirus capsid is composed of 60 copies of two coat proteins. The host range of each member is usually limited to one plant family. For example, the natural hosts of CPSMV and BPMV are limited to species in the family Leguminosae (Fabaceae; Valverde and Fulton, 1996).

Organization and expression of comovirus genomes

The complete nucleotide sequences of BPMV (Di, et al., 1999; MacFarlane, et al., 1991), CPMV (Lomonossoff and Shanks, 1983; van Wezenbeek et al., 1983), CPSMV (Chen and Bruening, 1992a, b), RCMV (Shanks and Lomonossoff, 1992; Shanks et al., 1986), SqMV (Han et al., 2002) have been reported. The sequences of RNA2 and the RdRp gene of APMoV are also published (Shindo et al., 1993; Krengiel et al., 1993). The sizes of RNA1 of comoviruses range from 5.9 to 7.2 kb and those of RNA2 vary from 3.5

to 4.5 kb. Both RNA1 and RNA2 have a genome-linked viral protein (VPg) covalently linked to the 5' terminus and a 3'-terminal poly (A) tail (Welink *et al.*, 2000).

Cowpea mosaic virus (CPMV), the type species of the genus *Comovirus*, has been extensively studied (Pouwels et al., 2002a). Both genome organization, expression strategy, and the conserved sequence motifs of the proteins encoded by CPMV RNA1 show a strong resemblance to those of the animal picornaviruses (Kamer *et al.*, 1984; Franssen *et al.*, 1984). Both RNA1 and RNA2 are expressed via the synthesis and subsequent cleavage of large precursor polyproteins. CPMV RNA1 contains a single large open reading frame (Lomonossoff and Shank, 1983; Van Wezenbeek eat al., 1983). The 200 kDa polyprotein precursor encoded by RNA1 is cleaved into 5 mature proteins (from 5' to 3': protease cofactor [Co-Pro, 32K], NTPM [58K], VPg, Protease [Pro, 24K], Polymerase [Pol or RdRp, 87K]) and 5 intermediate proteins (170K [NTBM+VPg+Pro+Pol], 112K [VPg+Pro+Pol], 110K [Pro+Pol], 60K [NTBM+VPg], 84K [NTBM+VPg+Pro]) (Pouwels *et al.*, 2002a). The Co-Pro is a hydrophobic protein functioning as a cofactor during cleavage of the RNA2 encoded polyprotein precursor (Peter *et al.*, 1992). The Co-Pro is also involved in the regulation of RNA1 polyprotein processing as well as targeting the replication proteins to host membranes where viral RNA replication takes place (Peter *et al.*, 1992; Carette *et al.*, 2002a). When individually expressed from a *Tobacco rattle virus* (TRV) expression vector, the Co-Pro induced necrotic lesions in the inoculated *Nicotiana benthamiana* leaves, suggesting that it is a cytotoxic protein (Carette *et al.*, 2002b). The 58K protein has been designated NTBM, based on the fact that it contains a Walker nucleotide-binding motif, which is characteristic of viral RNA helicases (Gorbalenya *et al.*, 1990). Mutation in the NTBM debilitated its binding capacity to ATP, and as a consequence viral RNAs were not able to replicate in cowpea protoplasts (Peters *et al.*, 1994). The 60K [NTBM+VPg] protein also interacts with three host proteins upon transient expression, two of which are homologs of the VAP family of SNARE-like proteins of animals and they are associated with ER membranes. Another protein, which interacted with the 60K in the yeast two-hybrid system, is the translation elongation factor eEF-1β, a putative component of the replication complex of positive sense RNA viruses (Carette *et al.*, 2002c). The 60K protein (NBTM+VPg) was shown to induce the formation of small membranous

structures in both plant and insect cells by using the TRV and baculovirus expression systems, respectively (Carette *et al.*, 2002; van Bokhoven *et al.*, 1992). The VPg of CPMV contains 28 amino acids and is covalently linked to the 5'-termini of both positive and negative strands of the viral RNAs. It was proposed that VPg may prime viral RNA transcription (Lomonossoff *et al.*, 1985; Pouwels *et al.*, 2002a). The 24K protein is a member of the trypsin-like family of serine proteases and the major player in processing of both RNA1 and RNA2 encoded polyproteins (Verver et al., 1987; Dessens and Lomonossoff, 1991). Although the C-terminal 87K protein has an RNA-dependent RNA polymerase domain, it was proposed that the 110-kDa protein (87K+24K) is the polymerase because it is the only viral protein found in associated with purified viral replication complex (Eggen et al., 1988).

Depending on which start codon (AUG at positions 161 to 163 or positions 512 to 514) is used, RNA2 can be translated into either a 105K or 95K polyprotein (Rezelman et al., 1989). RNA2 polyproteins are processed into 4 mature proteins (from 5' to 3': CR [cofactor of RNA2 replication, 58K], MP [movement protein], LCP [large coat protein] and SCP [small coat protein]) and one intermediate protein (60K [LCP+CP]) (Pouwels et al., 2002a). The CR is the protein translated from the first start codon in RNA2 and is carboxy coterminal with the MP. The N-terminal region of CR (58K) is tightly associated with CPMV RNA2 replication. It was suggested that the N-terminal region of the CR may target RNA2 to the replication complex (van Bokhoven et al., 1993a). The MP of CPMV is capable of inducing the formation of tubular structures, which are deposited into the cell walls at the site where viral particles move from cell-to-cell (Wellink et al., 1993; Pouwels et al., 2002b). Using wild type and mutant forms of MP, Carvalho et al. (2003) showed that the MP binds to intact viral particles and specifically to the LCP. The MP is able to bind to rGTP and ssRNA *in vitro* (Carvalho et al., 2004), and mutational analysis further demonstrated that GTP binding is critical in targeted transport of the MP. The MP is divided into two regions based on their functions. The N-terminal region (1-313) is responsible for the formation of the tubular structures, which protrude from the virus infected cells to the neighboring cells, whereas the C-terminal region (314-331) is required for targeting virus particles to tubules (Lekkerkerker et al., 1996). The LCP is able to bind to MP *in vitro*, but the significance of this interaction is not clear (Carvalho

et al., 2003). Notably, the LCP and SCP can form virus-like particles when expressed separately in insect cells or protoplasts (Wellink et al., 1996, Shank and Lomonossoff, 2000), suggesting that host proteins are not required for virion assembly.

Remarkably, it is not known which of the CPMV genes may serve as the determinants of symptom severity in infected plants even though almost every gene has been studied extensively. The symptoms expressed in diseased plants are of the greatest concern for growers because they are highly correlated with yield reduction. In the case of RCMV, it was shown that both RNA1 and RNA2 contribute to symptom severity in *Pisum sativum* (Oxelfelt et al., 1992). However, which gene in RNA1 and RNA2 specifies symptom severity remains unknown. It is thus important to investigate further the symptom determinants in comoviruses

Like CPMV, RNA1 of all comoviruses so far sequenced also encodes 5 mature proteins (Di et al., 1999). Comparison of RNA1 encoded proteins of 4 species in the genus *Comovirus* indicates that the sizes of the individual proteins are very close to their corresponding proteins of other comoviruses (Di et al., 1999). The percentage identities/similarities scores of the deduced amino acid sequences of the mature proteins between comoviruses vary from 27.5/39.3% to 75.0/89.3%. The corresponding similarity scores of RNA2- encoded proteins range from 21.1 to 56.7% (Haudenshield and Palukaitis, 1998).

Replication of comovirus RNA

The RNA1 of CPMV is capable of self replication in cowpea protoplasts in the absence of RNA2 (Goldbach et al., 1980). In contrast, replication of RNA2 is dependent on RNA1-encoded proteins (van Bokhoven, 1993b). CPMV infection induces a cytopathological structure containing small membranous vesicles in cowpea cells (De Zoeten et al., 1974). It has been proposed that replication of positive-strand RNA viruses is associated with intracellular membranes (Buck, 1996). The notion that CPMV replication takes place in association with host membranes is further supported by the fact the RdRp activity of CPMV is associated with the crude membrane fraction of infected cowpea plants (Eggen et al., 1988). In the GFP transgenic *N. benthamiana* plants, GFP binds to the endoplasmic reticulum (ER) and the Golgi membrane (Haseloff

et al., 1997; Carette et al., 2000). By using the transgenic plants, Carette and coworkers (2000) found that CPMV infection elicited a massive proliferation of the ER, which in turn resulted in the production of small membranous vesicles where viral RNA replicates. In contrast, the Golgi membranes remain intact. These results were confirmed by fluorescent *in situ* hybridization (FISH) assays. CPMV plus-strand RNA is localized in multiple small membrane vesicles distributed throughout the cytoplasm 12 hours post infection. These small bodies merge into large membranous vesicles that are detected adjacent to the nucleus (Carette et al., 2002a). It was further shown that CPMV RNA colocalized with viral replication complex, suggesting that virus replication takes place in the membrane sites. CPMV replication is blocked by an inhibitor of lipid synthesis, implying an inter- connection between lipid biosynthesis and the proliferation of the membranes (Carette et al., 2000).

Proliferation of the ER membranes induced by individually expressed Co-pro or 60K (58Kl+VPg) resembles that elicited by CPMV infection (Carette et al., 2002b). It was proposed that the Co-pro and 60K target the replication complex to ER membranes at the beginning. As a result, interaction between the viral-encoded proteins and ER membranes leads to rearrangement of ER membranes and formation of the small membranous structure in plant cells. Accordingly, Pouwels et al. (2002a) proposed a model about the formation of CPMV replication sites. Upon infection of plant cells by CPMV, the RNA1 encoded polyprotein is processed to the Co-pro (32K) and 170K proteins. Interaction between Co-pro and 60K (Hel+VPg) domain in the 170K prevents further proteolytic cleavage of the 170K. The localization signal in the Co-pro guides the Co-pro/170K complex to ER membranes, from where the 170K is further processed to either NTBM (58K) and 112K (VPg+Pro+Pol) or 60K and 110K (Pro+Pol). Meanwhile, the NTBM or 60K is inserted into the ER membrane. As a consequence, the replication complex binds to the ER membrane. The VPg released from either 60K or 112K may link to viral RNAs shortly. The interaction between Co-pro/NTBM (or 60K) and membranes triggers lipid biosynthesis and proliferation of ER membranes, which further compart-mentalizes the viral RNAs and provides more spaces for RNA replication.

However, how the viral RNAs bind to the replication complex remains a mystery, and it is also not known how the Co-pro/NTBM induces the proliferation of the ER

membranes. Furthermore, no information on the replication of other comoviruses is presently available.

Movement of comoviruses

Following replication, viral RNAs are assembled into virus particles in the cytoplasm. Carette and coworkers (2002a) have shown that viral RNAs accumulate in membranous vesicles adjacent to the nucleus 36 hours post infection. Virus particles, on the other hand, have been shown, based on immunodetection, to accumulate at the periphery of the plasma membranes in the protoplasts. The results suggest that viral assembly may take place in the peripheral structures. However, it is more likely that the viral RNAs are encapsidated at the sites where RNAs are replicated and translated and the virions are then transported to the cell periphery (Carette et al., 2002a). The latter proposal is supported by the evidence that the capsid proteins can assemble into virus-like particles when expressed in insect and protoplast cells (Wellink et al., 1996; Shank and Lomonossoff, 2000). Present evidence suggests that no other viral proteins are required for capsid assembly. How the virions travel from the assembly site to the cell periphery remains uncertain.

CPMV moves from cell-to-cell in the form of viral particles via tubular structures that are pierced through the cell wall. It was shown that tubules containing a single row of virions were produced in plant cells upon infection by CPMV (van Lent et al., 1990). Mutation analysis indicates that the MP of CPMV is the only viral protein required for tubule formation (Kasteel et al. 1993; Wellink et al., 1993). In addition, mutations in the central region of CPMV MP abolish virus ability to induce tubular structures and therefore mutant virions are unable to move from cell-to-cell (Bertens et al., 2000). Expression of MP using a baculovirus vector also induces the formation of similar tubular structures in the insect cell surface (Kasteel et al., 1996). Taken together, these results suggest that either host components are not necessary for forming tubules or that they are highly conserved between plants and animals.

It is not clear how the virion-containing tubules are formed. However, it is certain that both cytoskeleton and the secretory pathway are not necessary in the assembly of those tubular structures because inhibitors of cytoskeleton and secretory pathway cannot

block tubule formation induced by CPMV inoculation (Pouwels et al., 2002a). And also, intact plasmodesmata are not essential since the tubular structures, which resemble that induced by virus infection in plant cells, protruding from the cell surface to medium were observed in the protoplast and insect cells in the presence of CPMV movement protein (Carvalho et al., 2003; Kasteel et al., 1996; ven Lent et al., 1991). Immuno-electron microscopy observation indicates that the CPMV MP is a main ingredient of the tubules (van Lent et al., 1990). Similarly, when the MP fused to GFP (MPfGFP) was expressed, it was observed that the MP accumulated in the peripheral punctate structures and in the long tubules protruding from the protoplast cell surface (Pouwels et al., 2002c). Pouwels and associates (2003) identified the distinct steps during tubule formation by mutational analysis of MP. The C-terminal deletion (deletion of amino acids 252-331) of MP results in that the mutant MPfGFP accumulated mostly in the plasma membrane and rarely in peripheral punctate structure. In contrast, fluorescence from mutant MPfGFP is observed mainly in peripheral punctate structures and in some protoplast cells when protoplast are infected with a MP mutant virus (deletion of amino acids 277-331). Another mutation of MP (deletion of amino acids 297-331) corresponds to mutant MPfGFP accumulates in peripheral punctate structures and occasionally in short tubular structures. The results suggest that MP may move to and accumulate in the plasma membranes in an early stage and gather to the punctate structure where the tubules are assembled later. MP mutations further demonstrate that tubular structures disperse and virions are released in the neighboring cells (Pouwels et al., 2003). Taken together, a working model about CPMV cell-to-cell movement is proposed as follows (Poulwels et al., 2002a; Poulsels et al., 2003): A proportion of MP binds virus particles at viral RNA replication/virion assembly sites and guides the MP-virion complex to plasma membranes without participation of cytoskeleton and secretory pathway (Carvalho et al., 2003). On the other hand, the majority of MP or GTP binding MP diffuses to the cell periphery and are targeted to plasma membranes (Pouwels et al., 2002b). In the plasma membranes, MP and MP-virion complex accumulate via interaction with membrane residing proteins. Tubules initiated in punctate structure form within plasmodesmata through polymerization of MP and MP-virion complex, thereby encaging virus particles (Pouwels et al., 2002a). The growing tubules replace the desmotubules within the plasmodesmata and eventually extend to the

neighboring cells where the tubular structures are disassembled and the virions released. Subsequently, the neighboring cells are infected (Pouwels et al., 2003). Hydrolysis of the MP-binding GTP may provide energy to the whole event (Carvalho et al., 2004).

Silva and colleagues (2002) reported that vascular movement of CPMV is through phloem of the cowpea plant. By using GFP-expressing recombinant virus, it was illustrated that CPMV invade both major and minor veins of the inoculated leaves. However, it is only unloaded from major veins. CPMV is able to replicate in all types of vascular cells except companion cell (CC) and sieve cell (SC). CPMV move from phloem parenchyma cell to CC and from CC to SC in a way other than through tubules because those structures are absent in the plasmodesmata connecting those cells (Silva et al., 2002). The detail of CPMV vascular movement is still enigmatic.

Transmission of comoviruses by vectors

Most comoviruses are transmitted by leaf feeding beetles in the families *Chrysomelidea, Coccinellidae, Curculionidea or Meloidae*, except GMV, PMiMV and UVC whose beetle vectors have not been identified (Gergerich and Scott, 1996). Beetles are able to transmit comoviruses immediately after feeding on infected plants, although prolonged feeding indeed increases transmission frequency. Bean leaf beetle (*Cerotoma trifurcata*), the vector of BPMV, can transmit virus after a single bite of an infected soybean plant (Pitre and Patel, 1975). The virus acquisition/inoculation access periods of *C. ruficornis* of less than 5 min/ 1hr have been documented for CPSMV (Gergerich and Scott, 1996). Comoviruses do not replicate in their beetle vectors (Hull, 2001). Furthermore, a latent period is not necessary for the transmission of comoviruses. The retention time during which a beetle can transmit a virus following feeding acquisition varies from 1 to 10 days among different beetles species. Virus retention time as long as several months has been observed in preliminary reports for SqMV and BPMV in the overwintering beetles (Gergerich and Scott, 1996). The efficiency of comovirus transmission is highly related with the species of beetle vector and the species of host plant. For instance, BRMV is transmitted by *C. ruficornis* at a frequency close to 80%. In contrast, the transmission frequencies of *Diabrotica balteata* and *D. adelpha* are only 20% and 10%, respectively (Gergerich and Scott, 1996).

Seed transmission

Seed transmission has been documented for 6 members in the genus *Comovirus*, and the transmission frequencies range from 1% to 90% (Hull, 2001). Embryo infection by comoviruses derives either from the pollen or the ovule. Seedling infections of SqMV and BBMTV from ovule have been reported. In addition, BBMV was able to infect the seeding through either the virus-containing pollen or ovule (Gergerich and Scott, 1996). The transmission frequency of BPMV through seeds is as low as 0.1%, suggesting that the seed transmission might come from seed coat contamination by this virus (Giesler et al., 2002). Seed transmission plays an important role in the epidemiology of comoviruses. Low percentage of infected seedlings resulting from seed infection will be augmented after secondary spread by a large number and efficient beetles, thereby causing significant yield reductions (Gergerich and Scott, 1996).

Diversity of the genus *Comovirus*

Based on experimental and diagnostic host reactions and/or serological assays, ten species in the genus *Comovirus* have been reported to contain more than one strain (Lomonossoff, 2001, Valverde and Fulton, 1996). In contrast, only one strain was identified in BBTMV, GMV, PMiMV, QPMV and UVC. SqMV contains 6 biotypes based on host range as well as symptomatology. The biotypes are classified into two serological groups by agar double-diffusion serological tests (Nelson and Knuhtsen, 1973). Diversity of SqMV isolates collected in the United States was further investigated by nucleotide sequencing and Northern blot hybridization analysis of RNA2. The results indicate that SqMV has at least two subgroups (Haudenshield and Palukaitis, 1998). One isolate of SqMV collected from Japan (Y-SqMV) was sequenced recently (Han et al., 2002). Sequence comparison among isolatesof SqMV suggests that the isolates Y may belong to a distinct subgroup of strains. In the case of RCMV, cDNA clones of strains O and S were generated and those c-DNAs were used to make probes. The results from the Northern blot hybridization indicated the existence of two subgroups of strains as well as pseudo-recombinants between them (Oxelfelt et al., 1992).

Complete nucleotide sequences of 6 species in the genus *Comovirus* have been reported (Lomonossoff, 2001). However sequence diversity within species other than SqMV has not been studied extensively although yield reductions caused by different isolates vary significantly. Sequence comparisons and molecular genetic analysis of different strains are needed to extend our knowledge about host range, symptomatology and serology, properties that are used to differentiate strains within species.

Cross-protection

Cross-protection among three isolates of SqMV has previously been reported. In contrast, SqMV was unable to protect against infection with *tobacco ringspot nepovirus, cucumber mosaic cucumovirus* and *watermelon mosaic 2 potyvirus* (Demski, *1969)*. Cross-protection was also demonstrated between two isolates of RCMV. Attempts to demonstrate cross-protection between two different members of the genus *Comovirus*, BBSV and BBTMV were not successful. However, incomplete protection was observed between BBMV and RCMV (Valverde and Fulton, 1996). All these studies were conducted in the 60s and 70s at a time when the general knowledge about viruses was limited. With more and more comoviruses characterized at the molecular level, it may now be possible to gain an understanding of the mechanisms underlying cross-protection.

Research objectives and outline

The major objectives of my research are to:

1. decipher the genetic diversity among natural isolates of *Bean pod mottle virus*;
2. identify the viral genes responsible for symptom severity; and
3. investigate whether cross protection can be exploited as a disease management strategy and elucidate the mechanism underlying cross-protection.

The information generated in this study should be useful to the understanding of BPMV-host interaction.

In **Chapter II** , the complete nucleotide sequences of the genomic RNAs of two strains of BPMV are reported. The sequences are compared with those previously reported for BPMV strain K-G7 and other comoviruses. In **Chapter III**, I describe the results of my studies on the diversity among natural isolates of BPMV, which

demonstrate the presence of at least two distinct subgroups of strains. In **Chapter IV**, cross-protection among isolates of BPMV is investigated and the possible mechanisms are discussed. In **Chapter V,** I describe the isolation and molecular characterization of a novel diploid reassortant of BPMV that induces very severe symptoms in soybean. The diploid reassortant contains two types of RNA1 and one type of RNA2. In **chapter VI,** I describe the production of full-length cDNA clones from which infectious transcripts can be generated and present evidence that the Co-pro and Hel coding regions are symptom severity determinants

Chapter II

Complete nucleotide sequences of two strains of BPMV: Sequence comparisons and evolutionary relationships to other Comoviruses

Introduction

Bean pod mottle virus (BPMV) is a member of the genus *Comovirus* in the family *Comoviridae* (Goldbach, et al., 1995). Like other comoviruses, BPMV has a bipartite positive-strand RNA genome consisting of RNA-1 and RNA-2, which are separately encapsidated in isometric particles 28 nm in diameter. The genus *Comovirus* includes 15 approved members (Lomonosssoff and Ghabrial 2001). Of these, the complete nucleotide sequences of BPMV K-G7 (Di, et al., 1999; MacFarlane, et al., 1991), Cowpea mosaic virus (CPMV, Lomonossoff and Shanks, 1983; van Wezenbeek et al., 1983), *Cowpea severe mosaic virus* (CPSMV, Chen and Bruening, 1992a, b), *Radish mosaic virus* (RCMV, Shanks and Lomonossoff, 1992; Shanks et al., 1986), and *Squash mosaic virus* (SqMV, Han et al., 2002) have been reported. The sequences of RNA2 and RdRp gene of *Andean potato mottle virus* (APMoV) have also been published (Shindo et al., 1993; Krengiel et al., 1993).

Symptoms induced by various isolates of BPMV vary from mild, moderate to severe. The strain K-G7 induces mild symptoms and its genome sequence was reported earlier (Di, et al., 1999; MacFarlane, et al., 1991). In this chapter, I report the complete nucleotide sequences and deduced amino acid sequences of two strains of BPMV, of which K-Ho1 induces severe symptoms and K-Ha1 elicits mild symptoms in soybean plants. Sequence comparisons to the previously characterized BPMV strain K-G7 (MacFarlane et al., 1991; Di et al., 1999) as well as to other comoviruses are made.

Materials and Methods

cDNA synthesis and cloning

The BPMV isolates, K-Ho1 and K-Ha1, were collected from Kentucky and maitained in soybean plant (cv. Essex) in green house. Viral RNAs were extracted from

purified virions according the procedures of Peden and Symons (1973). RNA1 and RNA2 were purified from low-melting agrose following electrophoretic separation of the viral RNAs. cDNA synthesis was carried out using the SuperScript choice system (Gibco-BRL). First strand cDNA synthesis was primed with oligo$(dT)_{12-18}$ primers. Following addition of *Eco*RI adapters to the ends of the double stranded cDNA, it was ligated into *Eco*RI-linearized pGEM 3ZF(+) vector (Promega).

Sequence analysis

Multiple independent cDNA clones containing large inserts were selected and used for sequencing. Universal M13 and gene-specific sequencing primers were used to sequence both K-Ho1 and K-Ha1. M13 forward and reverse primers were used to sequence the 5' and 3' terminal nucleotides of the selected cDNA clones and primers based the generated sequences were synthesized for subsequent sequencing (Tables 2.1 and 2.2). The longest cDNA clones missing nucleotides 1-23 were cloned into pGem-Teasy vector (Promega) using the 5' RACE kit. All sequencing was carried out by the dideoxy-termination sequencing using the Rhodamine-Terminator sequencing kit (ABI) and an ABI310 automated sequencer. The complete nucleotide sequences of the genomic RNAs from isolates K-Ha1 and K-Ho1 have been deposited in the GenBank (see Chapter III for the GenBank accession numbers). Sequence analysis was performed using the GAP, PILEUP and PRETTY programs of the University of Wisconsin GCG software package (Genetics Computer Group, Inc, Madison, WI). Phylogenetic trees were created using the neighbor-jointing method, as implemented in the PAUP program 4.0b 2a PPC (Swofford, 2000).

Results and discussion

Sequence analysis of two strains of BPMV

The complete nucleotide sequences of K-G7 RNA1 and RNA2 have been reported (Di et al., 1999; MacFarlane et al., 1991). To decipher the diversity among isolates of BPMV, multiple cDNA clones representing the entire lengths of the genomic RNAs of two other strains, K-Ho1 and K-Ha1, were sequenced (Figure 2.1-2.4). At least three

clones were used to sequence each nucleotide. The lengths of the coding and noncoding regions of RNA1 and RNA2 were compared with the corresponding regions of strain K-G7 as well as with other comoviruses (Tables 2.3 and 2.4). The percentage nucleotide and deduced amino acid sequence identity of RNA1 and RNA2 between BPMV strains is shown in Figure 2.5. Alignment of the 5'-UTR of RNA1 and RNA2 showed that they share extended regions of sequence identity (Figure 2.11). Percentage sequence identity and similarity of the individual mature proteins encoded by RNA1 and RNA2 between BPMV strains and other comoviruses are shown in Figures 2.7 and 2.10, respectively. Phylogenetic analyses of comovirus RNA1 and RNA2 encoded polyproteins and mature proteins was performed using the neighbor-joining method. The resulting consensus trees of 1000 bootstrap replicates are separately shown in Figures 2.6 and 2.9.

RNA1 and RNA1 encoded proteins

K-Ho1 RNA1 is 5986 nucleotides in length excluding the poly (A) tail (Figure 2.1). The complete nucleotide sequence of K-Ha1 RNA1 is 5989 in length (Figure 2.3). Similar to BPMV strain K-G7 as well as other members of the genus *Comovirus*, sequence analysis of RNA1 of either K-Ho1 and K-Ha1 revealed the presence of a single open reading frame (ORF). The K-Ho1 ORF is predicted to initiate at the AUG at nucleotide positions 367 to 370 and to terminate at the UAG at nucleotide positions 5920 to 5922. The K-Ho1 ORF encodes a polyprotein of 1851 amino acids with calculated molecular mass of 209,491 Da. It is one amino acid longer than that of K-G7 (Di, et al., 1999). The RNA1 ORF of K-Ha1 is predicted to initiate at the AUG at nucleotide positions 370 to 373 and to terminate at the UAG at nucleotide positions 5923-5925. The RNA1 ORF codes for a polyprotein with calculated molecular mass of 209,309 Da. The sequence context (ACAAC**AUG**AA) surrounding the start codon in the RNA1 ORFs of K-Ho1 and K-Ha1 RNA1 are identical to those of K-G7. It is not in optimal context (AA<u>C</u>A**AUG**<u>G</u>C) for plant mRNA (Lütcke, 1987). Sequence comparisons indicated that K-Ho1 RNA1 share very high nucleotide and amino acid sequence identities with K-G7 RNA1 (98.1% and 98.0%, respectively; Figure 2.5A). The percentage of nucleotide sequence identity scores between K-Ha1 and K-Ho1 or between K-Ha1 and K-G7 RNA1 were similar (85.9% and 85.5%, respectively). Relatively higher identity scores,

however, were obtained for the deduced amino acid sequences (95.9% and 97.3%, respectively). These results are consistent with those of slot blot hybridization, which placed K-Ho1 RNA1 and K-G7 RNA1 in subgroup I and K-Ha1 in subgroup II (Gu, et al., 2002).

In order to demonstrate the relationship among BPMV strains and other members of the genus *Comovirus*, the percentage deduced amino acid identity/similarity between the individual proteins encoded by RNA1 were generated by the GAP program in the UWGCG package. Phylogenetic trees were also constructed based multiple alignments of the amino acid sequences of the RNA-encoded polyproteins and mature proteins. All five proteins encoded by RNA1s of K-Ho1 and K-Ha1 were of similar sizes to those encoded by K-G7 RNA1 except that the RdRps from K-Ho1 and K-Ha1 have one more amino acid residue than that of K-G7. With the exception of VPg, the sizes of the other four mature proteins varied among the different species in the genus *Comovirus*.

The percentage amino acid sequence identities/similarities of the Co-pro between BPMV strains in the same subgroup (99.7/99.7%) are higher than those between strains in different subgroups (96.7/97.7%). In contrast, the percentage identities/similarities of the Co-pro between different species in the genus *Comovirus* are significantly lower (23.3-39.2%/39.2-52.0%; Figure 2.7). The lowest identities/similarities scores of Co-pro were between RCMV and SqMV, whereas the highest interspecific scores were between BPMV (strain K-Ha1) and CPSMV. The percentage amino acid sequence identities/similarities of the helicase between strains of BPMV are also very high, ranging from 96.0/97.8% to 98.8/99.0%; whereas the interspecific identities/ similarities scores are low and vary from 39.6/49.7% to 52.6/62.9%. Sequence comparisons of the Co-pro and Hel indicated that K-G7 and K-Ho1 are more similar to each other than either is to K-Ha1. These results support the conclusion that BPMV is more closely related to CPSMV than to other comoviruses. Likewise, CPMV is more closely related to RCMV than to other members of comovirus (Di et al., 1999). The VPg is composed of 28 amino acid residues and is the most conserved protein of comoviruses. Fifteen out of 28 amino acid residues are identical (Figure 2.8A) among all comoviruses sequenced so far. The amino acid sequence VPg is identical among all BPMV strains.

Interestingly, it was found that the highest amino acid sequence similarity/identity scores of the protease is between strains K-Ho1 (subgroup I) and K-Ha1 (subgroup II) instead of between members of the same strain subgroup (Figure 4.7). Similar observations were also made in comparisons of the RdRp sequence among BPMV strains. The percentage identities/similarities scores of protease between different comoviruses range from 43.8/54.8% to 54.8/63.5%. The catalytic triad of CPMV protease, His 40, Glu76 and Cys166, was also found be absolutely conserved in the proteases of all three strains of BPMV as well as of other comoviruses (Dessens, et al., 1991; Di, et al., Figure 2.8 B). Sequence alignments indicated that the repeated Cys-Trp-Asp tripeptide in the protease of CPMV and RCMV show variation in one amino acid position in one repeat or another in BMPV, SqMV and CPSMV. Differences between BPMV (strain K-G7) and CPMV in the amino acids comprising the substrate-binding pocket of the protease reported by Di, et al. (1999) were also found in the other two strains of BPMV and SqMV (Figure 2.8. B). Comparisons of deduced amino acid sequences of comovirus RdRps demonstrated that RdRp is the second most conserved protein encoded by comovirus genomes. The highest interspecific identity/similarity scores of RdRp are between CPMV and RCMV (61.6/69.2%). In contrast, the lowest interspecific identity/similarity scores (49.3/57.9%) are between APMoV to CPSMV.

The results from protein sequence comparisons are in agreement with those inferred from phylogenetic analysis (Figure 2.6). In the phylogenic trees generated based on the complete polyprotein sequences, BPMV was shown to cluster with CPSMV, whereas CPMV clusters with RCMV (Figure 4.6A, Di et al., 1999). In addition, the BPMV strains K-Ho1 and K-G7 were found to be more closely related to each other than either one to K-Ha1. These same conclusions can also inferred from phylogenetic analysis of the Co-pro and Hel proteins. In contrast, phylogenetic trees created using VPg, protease and RdRp varied slightly from those generated based on the complete polyprotein sequences. CPSMV did not cluster with any of the other comoviruses when comovirus protease sequences were subjected to phylogenic analysis (Figure 2.6 E). Apparently, APMoV is distantly related to other comoviruses sequenced as far as can be inferred from the phylogenetic analysis of comovirus RdRp.

RNA2 and RNA2-encoded proteins

The complete nucleotide sequences of RNA2s of K-Ho1 and K-Ha1 are shown in Figure 2.2 and 2.4, respectively. The RNA2 of K-Ho1 consists of 3674 nucleotides, and it is one and twelve nucleotides longer than those of K-Ha1 and K-G7 RNA2, respectively. It was proposed that the translation of CPMV produces two carboxy coterminal polyproteins depending on which start codon is used (Holness, 1989). The ORFs of K-Ho1 or K-Ha1 RNA2s code for two large polyproteins, the larger of which have molecular masses of 113,508 and 113,517 Da, respectively. The larger RNA2 ORF in both K-Ho1 and K-Ha1 is predicted to initiate translation at the AUG at nucleotide positions 467 to 469 and to terminate at the UGA at nucleotide positions 3521 to 3523. The smaller ORFs are predicted to start at the AUG at nucleotide positions 773 to 775 and to terminate at the same UGA as the larger ORF. The second in-frame AUGs of K-Ha1 and K-Ho1 RNA2, like that of K-G7, are in optimal sequence contexts (GAAAG**AUG**GA) for plant mRNA (MacFarlane, et al., 1991). Nucleotide sequence alignment of RNA2 of BPMV strains showed that the RNA2s of strains K-Ho1 and K-Ha1 have the highest percentage sequence identity (98.8%). The percentage nucleotide sequence identities of RNA2 between K-G7 and K-Ho1 or K-G7 and K-Ha1 are 87.25 and 86.9%, respectively (Figure 2.5). These results are in good agreement with those of slot blot hybridization analysis, which placed K-Ho1 and K-Ha1 RNA2s in the same subgroup (II) and placed K-G7 RNA2 in a separate subgroup (subgroup I).

The sizes of the mature proteins encoded by BPMV RNA2, CR, MP, LCP and SCP, are 466, 364, 374 and 198 amino acid residues, respectively (Figure 2.2 and 2.4). Values for percentage of deduced amino acid sequence identities/similarities of the three proteins CR, LCP and SCP between BPMV strains and other comovirus, as determined by the GAP program, are shown in Figure 2.10. Since the coding region of MP is in frame with that of CR, the inter- and intraspecific identities and similarities scores were very similar to those of CR and thus were not included here. The values for percentage identities/similarities of CR were highest (>94.6/96.4%) between strains of BPMV with values as high as of 99.3/99.3% determined between the RNA2s of strains K-Ha1 and K-Ho1 (both RNA2s belong to subgroup II). Pair-wise alignments of other mature proteins encoded by BPMV RNA2 yielded similar results. Significant differences, however, were

observed when comparing the percentage amino acid sequence identities/similarities of CR between BPMV and other comoviruses. Pair-wise sequence alignments indicated that BPMV is most closely related to CPSMV (percentage identity/similarity of 42.0-42.8/52.4-52.9%). The lowest percentage identity score of CR (28.6%.) was obtained between APMoV and CPMV. These results are consistent with those of Haudenshield and Palukaitis (1998). It is noteworthy that the percentage amino acid sequence identities/similarities scores of LCP between BPMV strains are very close to each other, ranging from 98.9 to 99.7%. Based on identities/similarities scores of LCP between BPMV and other comoviruses, BPMV is most closely related to RCMV, but not CPSMV as is the case with other BPMV proteins. The lowest percentage identities/similarities scores of LCP between comoviruses were obtained between APMoV and CPSMV (30.3/42.9%). Pair-wise sequence alignments of SCP from BPMV strains K-Ho1 and K-Ha1 indicated that they are identical. Whereas, the percentage identity/similarity scores of SCP between BPMV subgroups I and II strains were 95% and 98%, respectively, the corresponding values of SCP between APMoV and RCMV (28.7%/36.2%) were the lowest among comoviruses. In this regard, the highest interspecific scores were obtained between CPMV and RCMV (49%; Figure 2.10).

The sequence comparison results were confirmed by phylogenetic analysis of comovirus RNA2-encoded polyproteins and individual mature proteins. The phylogenetic trees created by using the complete polyprotein sequences as well as sequences of the individual mature protein were essentially indistinguishable expect for that of SCP. Within BPMV, strain K-Ho1 RNA2 (subgroup II) is most closely related to K-Ha1 RNA2 (subgroup II) and the two cluster with strain K-G7 RNA2 (subgroup I). Among the comoviruses, BPMV is most closely related to CPSMV. On the other hand, APMoV is distantly related to other comovirues (Figure 2.10; Haudenshield and Palukaitis, 1998).

5' and 3' UTR

Sequencing data demonstrated that the lengths of the 5'- and 3'- UTRs of RNA1s and RNA2s are similar among BPMV strains. In comparing the 5'-UTRs of BPMV RNA1 and RNA2, it was found that the 5' UTRs of RNA2 are about 100 nucleotides longer compared to those of RNA1. Interestingly, the 5' UTRs of RNA1 and RNA2 were

found to share extended regions of sequence identity (Figure 2.11). Only one out of the first 5'-terminal 91 nucleotides is different in the 5' UTRs in the BPMV genomic RNAs. This characteristic has also been reported for RCMV and some nepoviruses (Shank and Lomonossoff, 1992; Greif *et al*., 1988; Ritzenthaler *et al*., 1991; Rott *et al*., 1991*)*. The lengths of the 3' UTRs in BPMV RNA2s are twice as large as those of RNA1s. No highly conserved regions were found upstream of the poly(A) tail, with the exception of the putative polyadenylation signal (AAUAAA).

Table 2.1. Primers used in K-Ho1 cDNA sequencing.

Name	5'→3'	Position
1. Forward primer for RNA1 cDNA		
BH-RNA1-5'	TACTGAAGTCCTCGCTCGTTTG	475-496
2BM-RNA1-5'	TGCATTTTGGCTAGCTCC	1034-1051
4HOPKINS-RNA1-5'	TGAAGGCTTACATGTCGC	1569-1586
5HOPKINS-RNA1-5'	AACATGGCTGGTTTGGAG	2021-2038
6HOPKINS-RNA1-5'	TGAACAAGGAAGGCGAGTG	2526-2544
2. Reverse primer for RNA1 cDNA		
BH-RNA1-3'	TCCAAGCAGTAGTAGGTAAAC	5652-5672
2BM-RNA1-3'	TCACACTCTGACAACTTTC	5091-5109
4HOPKINS-RNA1-3	ATTCCAACTTGGCAACTC	4537-4554
5HOPKINS-RNA1-3	TATCCTTCATGCTCTGTGC	3990-4008
6HOPKINS-RNA1-3	AGAAAAGGTCCCAAGCAG	3435-3452
HOP-RNA1-GSP1	CAGACGAGGATTACACT	501-507
HOP-RNA1-GSP2	GCATCTAACCTATTGGCTGTCTCA	424-447
3. Forward primer for RNA2 cDNA		
BH-RNA2-5'	TAGGACTTCGTGGGTAGAC	350-368
2BH-RNA2-5'	TCCTGCTGTTGACAAGTTG	874-892
4HOPKINS-RNA2-5'	AGATCTCGCAATGGTTAAAG	1399-1418
4. Reverse primer for RNA2 cDNA		
BH-RNA2-3'	ACTGTAGACTGTTTGGGATTG	3382-3402
2BH-RNA2-3'	ACCTGGTATTGTAGACACTGAAC	2802-2824
4HOPKINS-RNA2-3	TGTAACCTGAACATCCTGC	2287-2305
HAN-RNA2-GSP1	GCAGGAATGCCTCTAT	789-804
HOP-RNA2-GSP2	ACAGATGCCACCTAATGTAATGC	592-614

Table 2.2. Primers used in K-Ha1 cDNA sequencing.

Name	5'→3'	Positions
1. Forward primer for RNA1 cDNA		
HAN-RNA1-FOR-2nd	TGTGGCTGTGAAGAGGATACTGAAG	461-485
HAN-RNA1-FOR-3rd	TTGAAGAGGCTGCGAAGG	969-986
HAN-RNA1-FOR-4th	AGGAAACTTTTGGGATTGG	1460-1478
HAN-RNA1-FOR-5th	ATTTCAAGTGCTCCATACC	2000-2018
HAN-RNA1-FOR-6th	TCTCTGCTGGAAGGAAGGAC	2490-2509
2. Reverse primer for RNA1 cDNA		
HAN-RNA1-REV-1st	TGCTTCTGAAGGTAATTGAC	5598-5616
HAN-RNA1-REV-2nd	AGAACCACTAAAGTAAGGC	4996-5104
HAN-RNA1-REV-3rd	TCCATAGGGAGAATAGTGAAGC	4439-4470
HAN-RNA1-REV-4th	TGTTCTGTGCCCACCAATC	3984-4002
HAN-RNA1-REV-5th	TCACTGTCCCAGCAAAAAAG	3439-3468
HAN-RNA1-GSP1	CGAGGGTTACATTGCT	501-506
HAN-RNA1-GSP2	CCTCTTCACAGCCACAGGCAAAATA	452-476
3. Forward primer for RNA2 cDNA		
HAN-RNA1-FOR-2nd	TGTGGCTGTGAAGAGGATACTGAA	461-485
2BH-RNA2-5'	TCCTGCTGTTGACAAGTTG	874-892
4HOPKINS-RNA2-5'	AGATCTCGCAATGGTTAAAG	1399-1418
4. Reverse primer for RNA2 cDNA		
HAN-RNA2-REV-1st	TCCCAGGAATGTGGTTCTG	3261-3279
HAN-RNA2-REV-2nd	ACACCCATCAGCCAACAG	2756-2773
4HOPKINS-RNA2-3	TGTAACCTGAACATCCTGC	2287-2305
HAN-RNA2-GSP1	GCAGGAATGCCTCTAT	789-804
HAN-RNA2-GSP2	AACAACCAAAATGTCCAAATCTCT	530-553

Table 2.3. Lengths of RNA1 coding and noncoding regions in comovirus.

Virus	5'UTR	Co-pro	Hel	VPg	Pro	RdRp	3'UTR	Full length	Reference
K-Ha1	370	924	1788	84	624	2133	63	5989	This study
K-Ho1	367	924	1788	84	624	2133	63	5986	This study
K-G7	367	924	1788	84	624	2130	63	5983	Di et al, 1999
CPMV	206	978	1779	84	624	2133	82	5889	Lomonossoff and Shank, 1983
CPSMV	256	939	1785	84	630	2136	124	5957	Chen and Bruening, 1992a
RCMV	269	945	1800	84	624	2139	169	6033	Shank and Lomonossoff, 1992
SqMV	235	936	1800	84	630	2133	53	5865	Han et al., 2002

Table 2.4. Lengths of RNA2 coding and noncoding regions in comovirus.

Virus	5'UTR	CR	LCP	SCP	3'UTR	Full length	Reference
K-Ha1	466	1338	1122	594	150	3763	This study
K-Ho1	466	1338	1122	594	150	3674	This study
K-G7	454	1338	1122	594	151	3662	MacFarlane et al., 1991
APMoV	193	1239	1461	591	484	3671	Shindo et al., 1993
CPMV	511	1377	1122	639	180	3481	van Wezenbeek et al., 1983
CPSMV	254	1299	1122	585	469	3732	Chen and Bruening, 1992b
RCMV	290	1218	1128	642	262	3543	Shank et al., 1986
SqMV	164	1353	1122	552	124	3354	Han et al., 2002

```
UAUUAAAAUUUUCAUAAGAUUUGAAAUUUUGAUAAACCGCGAUCAUAGGUUGCCGCACCU 60

UAAAACCGGAAACAAAAGCAAUCGUUACUUGAUUUCAAAGACUUCUCAAUUUCUCUCUAC 120

AUUUCUUGUAUACAGCUUUCAAAGUGAAAGAAAAUCACUCUCUGUGCUGGUCACAGACUU 180

CGUGAAUCAUUUUCUUUCUGUUCUCAGUUCAUUUGCUGAACACUCUCCUAUUUGAUAUAG 240

GACUUCGUGUCAGAUUUGAACUUCUCCUAUCUUUCUUUCUCGGUUCUUCAUUUGAUUUCA 300

AAUUUCUCUGAAAUUUAAAUUUCUUUUGACAUUUUGAACUUUGUGUUGGCUCCAUUUGAA 360
```

Co-pro

			▼M	K	F	Y	P	G	Q	N	I	S	E	I	V	13
AAACAAC	AUG	AAG	UUC	UAU	CCU	GGU	CAA	AAU	AUU	UCU	GAA	AUU	GUU	406
Y	H	F	Q	S	N	E	T	A	N	R	L	D	A	Y		28
UAC	CAC	UUU	CAG	AGU	AAU	GAG	ACA	GCC	AAU	AGG	UUA	GAU	GCA	UAU		451
F	A	C	G	C	E	E	D	T	E	V	L	A	R	L		43
UUU	GCC	UGU	GGC	UGU	GAG	GAG	GAU	ACU	GAA	GUC	CUC	GCU	CGU	UUG		496
K	Q	C	N	P	R	L	L	H	L	S	Y	A	A	F		58
AAG	CAG	UGU	AAU	CCU	CGU	CUG	CUU	CAU	UUG	UCA	UAU	GCU	GCU	UUU		541
C	L	E	M	G	S	H	S	I	E	E	M	E	Y	D		73
UGU	UUG	GAA	AUG	GGC	AGU	CAU	UCA	AUA	GAG	GAA	AUG	GAA	UAU	GAU		586
D	G	E	L	I	F	S	Y	F	Q	N	F	L	L	S		88
GAU	GGG	GAA	UUA	AUU	UUU	UCC	UAU	UUU	CAA	AAU	UUU	UUG	CUU	UCC		631
I	V	S	N	S	S	K	T	T	K	L	R	A	Y	I		103
AUC	GUU	UCC	AAU	UCU	UCU	AAA	ACA	ACC	AAA	UUG	AGA	GCA	UAC	AUU		676
R	S	A	F	A	Y	H	F	Q	H	F	V	E	F	D		118
CGU	UCA	GCA	UUU	GCA	UAU	CAU	UUU	CAG	CAU	UUU	GUU	GAA	UUU	GAU		721
Q	Y	T	N	D	S	L	N	T	V	D	T	S	V	S		133
CAA	UAU	ACA	AAU	GAU	UCU	CUC	AAU	ACU	GUA	GAU	ACA	AGU	GUA	UCA		766
A	Q	G	I	A	D	L	A	L	S	M	V	R	W	I		148
GCC	CAA	GGG	AUA	GCA	GAC	UUG	GCU	CUC	UCU	AUG	GUU	AGA	UGG	AUA		811
P	T	Q	I	K	K	V	V	N	F	G	V	G	S	V		163
CCC	ACU	CAG	AUU	AAA	AAA	GUU	GUU	AAU	UUU	GGU	GUG	GGA	UCU	GUU		856
I	E	S	F	S	E	H	F	N	K	L	L	M	Q	Y		178
AUA	GAG	UCU	UUU	UCA	GAG	CAU	UUU	AAU	AAG	CUC	UUG	AUG	CAA	UAU		901
C	P	I	V	F	Q	A	F	S	W	V	N	N	I	W		193
UGU	CCA	AUA	GUU	UUU	CAA	GCU	UUC	AGC	UGG	GUU	AAC	AAU	AUU	UGG		946
T	M	V	K	E	W	I	E	E	A	A	K	E	I	S		208
ACA	AUG	GUC	AAA	GAA	UGG	AUA	GAA	GAA	GCU	GCG	AAA	GAA	AUU	UCA		991
W	F	L	Q	G	C	K	E	L	L	A	W	G	M	C		223
UGG	UUC	UUG	CAA	GGA	UGU	AAA	GAG	CUG	CUA	GCC	UGG	GGA	AUG	UGC		1036

Figure 2.1. The complete nucleotide sequence of BPMV-K-Ho1 RNA1. The deduced amino acid sequence of the polyprotein encoded by RNA1 is indicated in the one-letter code below the nucleotide sequence. The positions of the proteolytic cleavage sites are indicated with arrowheads with the names of the cleavage products indicated at the start of the coding regions of each of the mature proteins.

(Figure 2.1 continued)

```
I   L   A   S   S   C   A   L   G   L   V   E   K   C   L    238
AUU UUG GCU AGC UCC UGU GCU CUA GGA UUG GUU GAA AAA UGC CUU  1081
I   S   L   G   M   I   S   E   S   F   D   L   V   G   L    253
AUC UCU UUG GGC AUG AUU UCU GAA UCU UUU GAU UUG GUU GGU UUG  1126
F   V   R   S   A   I   V   G   A   F   C   V   S   I   K    268
UUU GUU CGA UCU GCC AUU GUG GGA GCU UUC UGU GUU UCC AUA AAA  1071
T   G   K   F   V   T   N   S   E   L   I   T   C   A   T    283
ACU GGU AAG UUC GUC ACG AAC AGU GAA UUG AUC ACU UGU GCU ACC  1216
I   A   V   S   T   I   A   T   V   M   S   Q   A   F   K    298
AUU GCA GUU UCU ACA AUA GCA ACU GUA AUG UCU CAG GCU UUU AAG  116
                                      Hel
P   S   E   E   I   K   G   Q   F   Q  ▼A   L   S   V   L    313
CCU UCC GAA GAG AUU AAG GGA CAG UUC CAA GCC CUU UCA GUU CUA  1306
E   G   L   A   T   Q   L   T   S   F   C   D   T   S   L    328
GAA GGG UUG GCA ACA CAG CUC ACU UCA UUU UGU GAC ACG UCU UUA  1251
V   A   M   G   K   T   C   T   A   F   N   Q   I   C   T    343
GUU GCU AUG GGA AAA ACC UGC ACA GCU UUU AAU CAA AUU UGC ACU  1396
A   G   K   N   V   K   V   I   A   G   R   L   L   E   V    358
GCU GGC AAA AAU GUU AAG GUG AUU GCA GGU AGG UUG CUA GAA GUU  1341
V   S   N   F   V   R   K   L   L   G   L   D   S   A   F    373
GUU UCU AAU UUU GUC AGA AAA UUA UUA GGA UUG GAU AGU GCU UUU  1486
L   R   D   A   A   L   I   F   S   Q   D   V   D   G   W    388
CUC AGA GAU GCU GCA CUC AUU UUU UCU CAA GAU GUG GAU GGA UGG  1531
L   R   N   I   S   W   C   Q   E   Q   F   L   L   K   A    403
UUG CGU AAC AUC AGU UGG UGC CAA GAA CAG UUU UUG UUG AAG GCU  1576
Y   M   S   Q   D   D   L   I   V   L   R   S   L   V   V    418
UAC AUG UCG CAA GAU GAU CUU AUU GUC CUG CGC UCC UUA GUU GUC  1621
K   G   E   R   M   R   E   Q   M   L   E   G   E   V   K    433
AAA GGU GAA AGA AUG AGG GAA CAG AUG CUU GAA GGA GAA GUU AAG  1666
V   S   P   S   V   C   N   L   I   V   K   G   C   E   E    448
GUG UCU CCA AGU GUU UGC AAC CUU AUU GUC AAA GGC UGU GAA GAA  1711
A   N   K   L   M   R   E   S   A   L   H   C   S   K   T    463
GCA AAU AAA UUG AUG CGU GAG AGC GCA CUU CAU UGU UCA AAA ACA  1756
I   R   K   I   P   F   V   I   F   A   H   G   E   S   R    478
AUU AGG AAG AUU CCU UUU GUU AUU UUU GCU CAC GGU GAA UCC CGG  1801
V   G   K   S   L   L   V   D   R   L   I   T   D   F   C    493
GUU GGG AAA UCU CUG CUG GUU GAU AGG CUA AUC ACA GAU UUC UGU  1846
D   H   L   E   I   G   E   D   A   V   Y   S   R   N   P    508
GAU CAU UUG GAA AUU GGA GAA GAU GCU GUG UAC UCA AGG AAU CCA  1891
S   D   P   F   W   S   G   Y   R   R   Q   P   I   V   T    523
UCA GAU CCU UUC UGG AGU GGA UAU AGA AGG CAG CCA AUU GUU ACU  1936
I   D   D   F   A   A   V   V   S   E   P   S   A   E   A    538
AUU GAU GAU UUU GCU GCU GUU GUU UCG GAG CCA UCU GCU GAA GCU  1981
Q   L   I   P   L   V   S   S   A   P   Y   P   L   N   M    553
CAG UUA AUU CCA UUA GUU UCA AGU GCU CCU UAU CCA UUA AAC AUG  2026
A   G   L   E   E   K   G   M   H   F   D   S   Q   I   M    568
GCU GGU UUG GAG GAA AAG GGA AUG CAC UUU GAU UCC CAG AUC AUG  2071
M   C   S   S   N   F   L   E   P   S   P   E   A   K   I    583
AUG UGU UCU UCA AAU UUC UUA GAG CCG UCU CCU GAA GCU AAA AUU  2116
R   D   D   M   A   F   R   N   R   R   H   V   L   I   T    598
AGA GAU GAU AUG GCU UUU AGA AAU CGA AGA CAU GUG CUG AUC ACA  2061
```

(Figure 2.1 continued)

```
V   E   L   K   P   G   V   E   Y   D   E   S   D   F   T    613
GUU GAA CUC AAA CCU GGG GUU GAA UAU GAU GAG AGU GAU UUU ACU  2206
K   N   Q   R   Y   L   L   K   T   W   F   H   D   H   Y    628
AAA AAU CAG CGA UAU UUG CUG AAA ACU UGG UUU CAU GAU CAU UAU  2151
V   V   D   Q   T   F   E   S   Y   A   D   L   L   A   H    643
GUU GUA GAC CAA ACU UUU GAG UCU UAU GCU GAU CUG CUG GCA CAU  2296
C   F   T   K   W   E   R   H   V   K   E   Q   E   S   N    658
UGU UUU ACU AAG UGG GAG AGA CAU GUU AAG GAG CAA GAA UCA AAU  2341
L   S   Q   I   K   G   K   K   N   E   S   G   H   F   N    673
CUG UCU CAA AUC AAG GGC AAG AAA AAU GAA AGU GGU CAU UUC AAU  238
N   F   Q   Q   L   M   D   L   A   V   S   W   N   L   S    688
AAC UUU CAA CAA CUU AUG GAU UUG GCU GUU UCA UGG AAU CUU AGU  2431
A   D   I   M   K   N   R   I   K   A   E   R   N   D   M    703
GCA GAU AUC AUG AAA AAC AGG AUC AAG GCU GAG AGA AAU GAC AUG  2376
V   Y   V   F   S   A   G   R   K   D   K   I   F   H   C    718
GUU UAU GUU UUU UCU GCA GGG AGG AAG GAU AAA AUU UUU CAU UGU  2521
F   L   N   K   E   G   E   C   T   V   R   P   D   S   I    733
UUU CUG AAC AAG GAA GGC GAG UGC ACG GUU CGU CCU GAU UCA AUA  2566
D   D   P   E   A   Q   A   L   L   K   A   S   E   T   M    748
GAU GAU CCU GAA GCG CAA GCU UUG CUC AAA GCU UCA GAG ACA AUG  2611
L   M   K   A   Y   A   F   L   K   Y   N   N   A   T   N    763
CUC AUG AAA GCC UAU GCC UUC CUC AAA UAU AAU AAU GCA ACA AAU  2656
L   I   V   R   T   H   L   A   E   L   V   N   E   D   F    778
UUG AUU GUC AGA ACC CAU UUG GCA GAA CUG GUG AAU GAA GAU UUC  2701
Y   D   E   K   F   N   F   I   G   T   I   G   T   P   A    793
UAU GAU GAG AAA UUC AAU UUC AUU GGA ACA AUU GGA ACA CCG GCU  2746
F   H   R   Q   I   A   A   H   L   E   K   M   P   L   W    808
UUU CAU CGC CAA AUA GCU GCA CAU UUG GAA AAG AUG CCA UUG UGG  2791
Q   K   A   I   L   C   G   M   G   H   C   L   S   R   K    823
CAA AAA GCA AUU UUG UGU GGA AUG GGA CAU UGU UUG UCU CGG AAA  2836
S   K   E   T   W   Y   T   G   M   K   E   K   F   V   Q    838
AGC AAA GAA ACC UGG UAU ACU GGU AUG AAG GAG AAA UUU GUG CAG  2871
M   M   K   S   I   Y   E   T   E   V   T   D   W   P   V    853
AUG AUG AAA AGC AUC UAU GAA ACU GAA GUC ACA GAU UGG CCA GUG  2926
P   L   K   I   I   S   G   T   I   L   A   T   I   L   G    868
CCA UUG AAA AUC AUU UCU GGU ACU AUU CUA GCC ACC AUU UUG GGA  2961
T   T   F   W   K   L   F   S   F   L   R   D   A   G   N    883
ACA ACU UUU UGG AAG UUA UUU UCC UUU UUA AGG GAU GCU GGU AAU  3016
G   G   V   F   V   G   N   V   A   S   A   F   T   T   S    898
GGA GGU GUU UUU GUU GGU AAU GUU GCU UCA GCA UUU ACU ACA UCA  3051
                      VPg
S   V   L   E   A   Q  ▼S   R   K   P   N   R   Y   E   V    913
AGU GUG CUC GAG GCG CAA AGC CGA AAA CCU AAC AGA UAU GAG GUC  3106
S   Q   Y   R   Y   R   N   V   P   I   K   R   R   A   W    928
UCU CAA UAU AGG UAU CGC AAU GUG CCA AUA AAG CGC AGA GCG UGG  3141
                  Pro
V   E   G   Q  ▼M   S   F   D   Q   S   V   V   A   I   M    943
GUU GAG GGC CAA AUG UCU UUU GAU CAA UCA GUG GUA GCA AUU AUG  3196
S   K   C   K   A   S   M   R   M   G   N   T   D   A   Q    958
UCA AAA UGU AAA GCC AGU AUG AGA AUG GGA AAC ACU GAU GCU CAA  3241
I   L   M   V   P   G   R   R   F   I   A   H   G   H   F    973
AUU UUG AUG GUU CCA GGG CGU AGA UUC AUU GCA CAU GGU CAU UUU  3286
```

(Figure 2.1 continued)

```
F   K   N   L   T   Q   K   V   R   V   Q   I   V   T   S    988
UUC AAG AAU CUC ACC CAA AAA GUU AGA GUC CAA AUU GUU ACU UCU  3331
E   K   T   Y   W   H   V   Y   D   P   D   K   F   Q   M    1003
GAG AAA ACC UAU UGG CAU GUG UAU GAU CCU GAU AAA UUU CAA AUG  3376
F   D   N   S   E   I   G   L   Y   T   N   P   T   L   E    1018
UUU GAU AAC AGU GAA AUC GGG UUG UAU ACA AAU CCA ACU UUG GAG  3421
D   I   P   H   S   A   W   D   L   F   C   W   D   S   E    1033
GAC AUC CCA CAU UCU GCU UGG GAC CUU UUC UGC UGG GAC AGU GAG  3466
K   T   L   P   N   N   F   S   A   E   L   L   S   C   K    1048
AAA ACU CUG CCA AAU AAU UUU UCU GCU GAA UUG CUU UCC UGU AAA  3511
```

(Figure 2.1. continued)

```
L   D   T   V   T   G   Q   Y   Y   P   E   W   A   P   I    1063
UUG GAC ACU GUU ACG GGA CAG UAU UAC CCA GAA UGG GCU CCA AUA  3556
N   C   R   V   H   R   Q   P   I   H   I   T   E   G   N    1078
AAU UGU CGA GUA CAU CGG CAA CCA AUU CAC AUA ACU GAA GGG AAU  3601
Y   V   R   K   Q   D   V   S   I   E   Y   D   A   C   T    1093
UAU GUU AGG AAA CAA GAU GUA AGC AUC GAA UAU GAU GCC UGC ACA  3646
I   P   N   D   C   G   S   L   V   V   A   K   V   G   N    1108
AUU CCU AAU GAU UGU GGA UCU CUG GUG GUU GCU AAG GUC GGA AAU  3691
H   K   Q   I   V   G   F   H   V   A   G   S   K   G   R    1123
CAC AAG CAA AUU GUU GGU UUU CAU GUU GCU GGA AGC AAA GGA AGA  3736
L   G   Y   A   S   L   I   P   Y   V   E   P   V   V   Q    1138
UUG GGC UAU GCU UCA UUA AUA CCA UAU GUU GAG CCU GUG GUA CAA  3781
```

RdRp

```
A   Q   ▼S   A   E   V   Y   F   D   F   F   P   V   E   V   1153
GCC CAA AGU GCU GAA GUC UAU UUU GAC UUC UUU CCU GUG GAA GUU  3826
D   S   Q   E   G   V   A   H   I   G   E   L   K   S   G    1168
GAU AGU CAA GAG GGA GUU GCU CAU AUU GGU GAA CUC AAA UCU GGA  3871
V   Y   V   P   L   P   T   K   T   N   L   V   E   T   P    1183
GUU UAU GUA CCA UUG CCC ACA AAA ACU AAU CUU GUG GAA ACU CCC  3916
K   E   W   Q   L   D   L   P   C   D   K   I   P   S   V    1198
AAA GAA UGG CAG UUG GAU UUG CCU UGU GAU AAG AUU CCA AGU GUG  3961
L   T   T   T   D   E   R   L   V   G   T   E   H   E   G    1213
UUA ACC ACU ACU GAU GAG AGA UUG GUU GGC ACA GAG CAU GAA GGA  4006
Y   D   P   F   L   G   G   I   Q   K   Y   A   T   P   M    1228
UAU GAC CCA UUU CUU GGU GGU AUU CAA AAA UAU GCA ACU CCC AUG  4051
M   P   L   D   E   E   I   L   S   K   V   A   Q   D   M    1243
AUG CCU CUU GAU GAG GAG AUU CUU UCC AAA GUU GCA CAA GAC AUG  4096
V   E   E   W   F   D   C   V   D   E   E   D   T   F   E    1258
GUU GAA GAA UGG UUU GAU UGU GUU GAU GAG GAG GAU ACA UUU GAA  4141
E   V   S   L   S   A   A   L   N   G   V   E   G   L   D    1273
GAA GUU UCU UUG AGU GCU GCA CUC AAU GGU GUU GAA GGU UUG GAU  4186
Y   M   E   R   I   P   L   A   T   S   E   G   F   P   H    1288
UAC AUG GAA CGC AUU CCU CUU GCC ACU UCA GAG GGU UUU CCU CAU  4231
V   L   S   R   K   N   G   E   K   G   K   R   R   F   V    1303
GUU CUG UCC AGG AAA AAU GGU GAA AAA GGC AAG AGA AGA UUU GUC  4276
T   G   D   G   E   E   M   S   L   I   P   G   T   S   V    1318
ACU GGA GAU GGU GAA GAA AUG UCA CUA AUU CCU GGU ACC AGU GUU  4321
E   E   A   Y   N   K   L   T   V   E   L   E   K   C   V    1333
GAA GAA GCA UAC AAU AAA UUG ACU GUU GAA CUA GAA AAG UGU GUU  4366
P   T   L   V   G   I   E   C   P   K   D   E   K   L   P    1348
CCA ACA UUG GUU GGC AUA GAA UGU CCC AAG GAU GAA AAA CUU CCC  4411
```

(Figure 2.1 continued)

```
R   R   K   I   F   D   K   P   K   T   R   C   F   T   I    1363
CGU CGC AAA AUU UUU GAU AAA CCC AAG ACG CGC UGC UUC ACC AUA  4456
L   P   M   E   F   N   L   V   V   R   Q   K   F   L   N    1378
CUU CCU AUG GAA UUU AAU CUA GUG GUG CGU CAA AAA UUC UUG AAU  4501
F   V   R   F   I   M   K   K   R   D   K   L   S   C   Q    1393
UUU GUG CGA UUC AUU AUG AAG AAA AGG GAC AAA UUG AGU UGC CAA  4546
V   G   I   N   P   Y   S   M   E   W   T   G   L   A   N    1408
GUU GGA AUC AAU CCA UAU UCU AUG GAG UGG ACU GGU UUG GCA AAU  4591
R   L   L   S   K   G   N   D   I   L   C   C   D   Y   A    1423
AGA CUG UUG AGC AAG GGA AAU GAC AUU UUG UGU UGU GAC UAU GCU  4636
S   F   D   G   L   I   T   K   Q   V   M   S   K   M   A    1438
AGU UUU GAU GGU CUG AUA ACU AAG CAA GUC AUG AGC AAG AUG GCA  4681
E   M   I   N   S   L   C   G   G   D   E   K   L   M   R    1453
GAA AUG AUA AAC AGU CUU UGU GGU GGA GAU GAG AAA CUG AUG CGU  4726
E   R   T   H   L   L   L   A   C   C   S   R   M   A   I    1468
GAG AGA ACA CAU CUU CUG UUA GCU UGU UGC UCC AGG AUG GCA AUC  4771
C   K   K   D   V   W   R   V   E   C   G   I   P   S   G    1483
UGU AAA AAA GAU GUU UGG AGA GUU GAG UGU GGU AUC CCU UCU GGA  4816
F   P   L   T   V   I   C   N   S   I   F   N   E   M   L    1498
UUU CCA CUC ACU GUU AUC UGU AAU AGC AUU UUC AAU GAG AUG CUU  4861
I   R   Y   S   Y   E   K   L   L   R   Q   A   K   A   P    1513
AUC AGA UAU AGU UAU GAA AAG UUG UUG CGU CAA GCU AAG GCU CCU  4906
S   M   F   L   Q   S   F   K   N   F   I   S   L   C   V    1528
AGU AUG UUU CUC CAG UCU UUU AAA AAU UUU AUU UCU UUG UGU GUU  4951
Y   G   D   D   N   L   I   S   V   H   E   Y   V   K   P    1543
UAU GGA GAU GAU AAU UUA AUU AGU GUU CAU GAG UAU GUU AAG CCA  4996
Y   F   S   G   S   K   L   K   S   F   L   A   S   H   N    1558
UAU UUU AGU GGU UCU AAA UUA AAA AGU UUC CUA GCU AGU CAU AAC  5041
I   T   I   T   D   G   I   D   K   T   S   A   T   L   Q    1573
AUC ACC AUU ACU GAU GGA AUU GAC AAA ACU AGU GCA ACU UUA CAG  5086
F   R   K   L   S   E   C   D   F   L   K   R   N   F   K    1588
UUU AGA AAG UUG UCA GAG UGU GAU UUU CUU AAA AGA AAU UUC AAG  5131
Q   M   S   N   V   L   W   V   A   P   E   D   K   A   S    1603
CAA AUG UCC AAU GUU UUG UGG GUA GCU CCU GAA GAC AAA GCU AGU  5176
L   W   S   Q   L   H   Y   V   S   C   N   N   L   E   M    1618
UUG UGG UCA CAA UUA CAC UAU GUU UCA UGU AAC AAU UUG GAA AUG  5221
Q   E   A   Y   L   V   N   L   V   N   V   L   R   E   L    1633
CAA GAA GCU UAU CUU GUU AAC UUG GUU AAU GUG UUG CGU GAG UUG  5266
Y   L   H   S   P   E   E   A   C   Q   L   R   R   R   A    1648
UAC CUG CAC AGU CCA GAA GAA GCU UGU CAG UUG AGA AGA AGG GCU  5311
L   S   R   I   E   W   L   Q   K   A   D   V   P   T   I    1663
CUC UCU CGU AUU GAG UGG UUG CAA AAA GCU GAU GUG CCC ACC AUA  5356
A   Q   I   E   E   F   H   S   M   Q   R   I   M   N   A    1678
GCA CAA AUU GAA GAA UUU CAU UCA AUG CAG AGG AUU AUG AAU GCU  5401
P   D   S   N   D   N   I   D   L   L   L   S   I   D   L    1693
CCU GAU UCA AAU GAU AAU AUU GAU CUU UUG UUG AGC AUC GAC UUG  5446
L   G   L   Q   G   A   G   K   A   F   P   N   K   I   V    1708
UUG GGU CUU CAG GGU GCA GGC AAG GCC UUC CCA AAU AAG AUU GUG  5491
F   D   D   K   L   V   L   A   N   T   Q   E   F   F   D    1723
UUU GAU GAU AAA UUG GUA UUG GCA AAU ACA CAA GAA UUU UUU GAU  5536
G   N   F   P   T   D   S   W   L   P   I   F   V   N   C    1738
GGA AAU UUU CCA ACA GAU UCU UGG UUA CCA AUA UUU GUU AAU UGU  5581
```

(Figure 2.1 continued)

```
L   Y   P   V   S   Q   L   P   A   E   A   V   I   V   N    1753
CUU UAC CCU GUG AGU CAA UUG CCC GCA GAA GCU GUC AUU GUU AAU  5626
V   V   C   G   S   G   R   G   G   L   P   T   T   A   W    1768
GUU GUC UGU GGG AGU GGG CGC GGU GGU UUA CCU ACU ACU GCU UGG  5671
I   S   S   A   V   N   N   R   S   S   D   I   N   K   K    1783
AUU AGU UCU GCA GUU AAC AAU CGC UCC UCA GAU AUC AAU AAG AAA  5716
I   R   T   A   L   G   K   G   K   K   I   V   F   L   T    1798
AUU CGG ACA GCG CUU GGA AAA GGU AAG AAA AUU GUC UUU UUG ACU  5761
R   V   D   P   F   P   V   A   L   L   A   V   L   F   G    1813
AGA GUU GAU CCU UUU CCU GUG GCC UUG UUA GCU GUU CUU UUU GGU  5806
```

(Figure 2.1. continued)

```
V   K   N   E   I   L   S   S   N   A   T   N   P   M   L    1828
GUU AAG AAC GAA AUU CUG AGU UCU AAU GCC ACA AAU CCA AUG UUG  5851
T   R   L   L   E   N   C   K   S   L   K   Y   L   V   D    1843
ACA AGG CUU CUU GAG AAC UGC AAG AGU CUU AAA UAU UUG GUU GAU  5896
E   C   P   F   A   F   V   N   *                            1852
GAG UGU CCU UUU GCA UUU GUU AAC UAG UUUGUAAUAUUUUGUUCACUUAAA 5947

UAAAGCGCAUUACUAUGUGCAAUGAGUGUGUUUAAAUAU                     5986
```

```
UAUUAAAAUUUUCAUAAGAUUUGAAAUUUUGAUAAACCGCGAUCACAGGUUGCCGCACCU 60

UAAAACCGGAAACAAAAGCAAUCGUUACUUGAUUUCAAGAAUUUCUCAAUUUCUUCCUAC 120

UUCCUUGUGUACGAUUUCUUAAGGGAAAGAAAAUCACUCUCUGUGCUGGCCACAGACUUC 180

GUGAAUCAUUUUCUUUUUCAAUCUUAGUUUAUUUGCUGAACACUCUCCUAUUUGAUAUAG 240

GACUUCGUGUCAGAUUUAAACUUCUUCUGUUUCUCUCUCAGUUCUCUGUUUAAUUUCAAG 300

UUCAAGCUGGUGAAAUCUUGGAUUAGUGCUCCCACUCUCCUAUCUGGUAUAGGACUUCGU 360

GGGUAGACUUUUCUAUUUCUCUCCUUUCUUUCACUCUCUUCUUCUCACUGAUCCGCAUUG 420
                                                        CR
                                                      ▼M   F   A     3
CCGUUCAAAGUGGUCUUAUUUGAAAAACACUUGGGCGUUGGUGCAA AUG UUU GCU   475
S   L   I   F   S   G   D   N   R   L   T   E   K   T   I    18
UCG UUA AUU UUC UCU GGA GAC AAC AGG CUC ACU GAG AAA ACA AUU  520
F   T   C   R   D   L   D   I   L   V   V   Y   Y   T   I    33
UUU ACU UGC AGA GAU UUG GAC AUC UUG GUU GUU UAU UAU ACA AUA  565
A   T   Q   F   R   K   F   L   P   H   Y   I   R   W   H    48
GCA ACU CAA UUU AGA AAA UUU CUA CCG CAU UAC AUU AGG UGG CAU  610
L   Y   T   L   L   I   Y   I   L   P   S   F   L   T   A    63
CUG UAU ACC UUG UUG AUC UAC AUU CUC CCA UCU UUU CUC ACU GCU  655
E   I   K   Y   K   R   N   L   S   N   I   H   I   S   G    78
GAA AUU AAA UAU AAG CGG AAU CUG AGU AAU AUU CAU AUU UCC GGC  700
L   F   Y   D   G   R   Y   K   F   W   T   K   H   E   K    93
UUA UUU UAC GAC GGC AGA UAC AAA UUC UGG ACU AAA CAC GAG AAA  745
                                    MP
N   L   A   L   T   E   E   E   K ▼M   E   V   I   R   N   108
AAU CUU GCU UUG ACA GAA GAG GAA AAG AUG GAA GUG AUU AGA AAU  790
R   G   I   P   A   D   V   L   A   K   R   A   H   E   F   123
AGA GGC AUU CCU GCU GAU GUU CUU GCA AAG CGA GCU CAU GAA UUU  835
E   K   H   V   A   H   E   S   L   K   D   Q   I   P   A   138
GAA AAA CAU GUU GCU CAU GAA AGC CUC AAG GAU CAA AUU CCU GCU  880
V   D   K   L   Y   S   T   K   V   N   K   F   A   K   I   153
GUU GAC AAG UUG UAU UCC ACU AAG GUU AAU AAG UUU GCA AAA AUU  925
M   N   L   R   Q   S   V   V   G   D   L   K   L   L   T   168
AUG AAC CUU AGA CAA AGU GUU GUU GGU GAU CUU AAA CUU CUU ACU  970
D   G   K   L   Y   E   G   K   H   I   P   V   S   N   I   183
GAU GGG AAG UUG UAU GAG GGU AAG CAU AUU CCU GUA UCU AAU AUU  1015
```

Figure 2.2. The complete nucleotide sequence of BPMV K-Ho1 RNA2. The deduced amino acid sequence of the polyprotein encoded by RNA2 is indicated in the one-letter code below the nucleotide sequence. The positions of the proteolytic cleavage sites are indicated with arrowheads with the names of the cleavage products indicated at the start of the coding regions of each of the mature proteins.

(Figure 2.2 continued)

```
S   A   G   E   N   H   V   V   Q   I   P   L   M   A   Q    198
AGU GCA GGG GAA AAU CAU GUA GUU CAA AUA CCC UUA AUG GCA CAG  1060
E   E   I   L   S   S   S   A   S   D   F   R   T   A   M    213
GAG GAA AUU CUG UCU UCU AGU GCA AGU GAU UUC AGA ACU GCA AUG  1105
V   S   K   N   S   K   P   Q   A   T   A   M   H   V   G    228
GUG AGU AAA AAU AGC AAG CCU CAA GCU ACU GCA AUG CAU GUG GGA  1150
A   I   E   I   I   I   D   S   F   A   S   P   D   C   N    243
GCU AUA GAA AUU AUC AUU GAU AGU UUC GCA AGU CCU GAC UGC AAC  1195
I   V   G   A   M   L   L   V   D   T   Y   H   T   N   P    258
AUA GUU GGU GCA AUG CUU UUG GUU GAU ACU UAU CAU ACC AAU CCU  1240
E   N   A   V   R   S   I   F   V   A   P   F   R   G   G    273
GAA AAU GCA GUU CGU AGU AUU UUU GUU GCG CCU UUC AGA GGC GGA  1285
R   P   I   R   V   V   T   F   P   N   T   I   V   Q   I    288
AGG CCC AUU CGG GUG GUU ACA UUU CCG AAU ACC AUU GUG CAG AUU  1330
E   P   D   M   N   S   R   F   Q   L   L   S   T   T   T    303
GAA CCA GAC AUG AAU UCA AGG UUU CAG CUU UUG AGU ACC ACU ACC  1375
N   G   D   F   V   Q   G   K   D   L   A   M   V   K   V    318
AAU GGU GAU UUU GUU CAA GGA AAA GAU CUC GCA AUG GUU AAA GUU  1420
N   V   A   C   A   A   V   G   L   T   S   S   Y   T   P    333
AAU GUA GCA UGU GCU GCU GUU GGC UUG ACA UCA AGU UAC ACU CCA  1465
T   P   L   L   E   S   G   L   Q   K   D   R   G   L   I    348
ACU CCA UUG UUG GAA UCU GGU UUG CAA AAA GAC AGA GGG UUA AUU  1510
V   E   Y   F   G   R   M   S   Y   V   A   H   N   V   N    363
GUG GAG UAU UUU GGA AGG AUG UCU UAC GUU GCU CAU AAC GUC AAU  1555
Q   P   Q   E   K   D   L   L   E   G   N   F   S   F   D    378
CAG CCC CAA GAG AAA GAU UUG UUG GAG GGA AAU UUU UCC UUU GAU  1600
I   K   S   R   S   R   L   E   K   V   S   S   T   K   A    393
AUU AAA UCU CGC UCU AGA UUG GAA AAA GUU UCC UCU ACA AAA GCA  1645
Q   F   V   S   G   K   T   F   K   Y   D   I   I   G   A    408
CAG UUU GUU AGU GGA AAA ACC UUC AAA UAU GAU AUA AUU GGU GCU  1690
G   S   H   S   S   E   D   F   P   E   K   E   D   Q   G    423
GGU UCA CAU UCU UCA GAA GAU UUU CCU GAA AAG GAA GAU CAA GGA  1735
K   P   K   K   I   D   A   R   L   R   Q   R   I   D   P    438
AAA CCC AAA AAG AUU GAU GCC AGA UUG AGA CAA AGA AUA GAU CCU  1780
                          L-CP
Q   Y   N   E   V   Q   A   Q  ▼M   E   T   N   L   F   K    453
CAA UAC AAU GAG GUU CAG GCU CAG AUG GAA ACA AAU UUG UUU AAA  1825
L   S   L   D   D   V   E   T   P   K   G   S   M   L   D    468
UUG UCU CUU GAU GAU GUU GAA ACU CCU AAA GGU UCC AUG UUG GAU  1970
L   K   I   S   Q   S   K   I   A   L   P   K   N   T   V    483
CUU AAA AUU UCU CAA UCU AAA AUU GCA CUU CCC AAG AAU ACA GUU  1915
G   G   T   I   L   R   S   D   L   L   A   N   F   L   T    498
GGA GGA ACC AUU UUG CGU AGU GAU CUA UUG GCA AAU UUU UUG ACA  2060
E   G   N   F   R   A   S   V   D   L   Q   R   T   H   R    513
GAG GGC AAU UUU AGA GCA AGU GUU GAU UUG CAG CGC ACU CAU CGU  2105
I   K   G   M   I   K   M   V   A   T   V   G   I   P   E    528
AUU AAA GGA AUG AUU AAA AUG GUG GCC ACA GUU GGU AUU CCU GAA  2150
N   T   G   I   A   L   A   C   A   M   N   S   S   I   R    543
AAU ACA GGU AUA GCA UUG GCC UGU GCU AUG AAU AGU UCU AUU AGG  2195
G   R   A   S   S   D   I   Y   T   I   C   S   Q   D   C    558
GGG CGC GCC AGU UCU GAU AUU UAC ACC AUC UGU UCU CAA GAC UGU  2140
```

(Figure 2.2 continued)

```
E   L   W   N   P   A   C   T   K   A   M   T   M   S   F    573
GAA UUA UGG AAU CCU GCU UGC ACA AAA GCA AUG ACC AUG UCA UUU  2185
N   P   N   P   C   S   D   A   W   S   L   E   F   L   K    588
AAU CCA AAC CCG UGU UCU GAU GCA UGG AGU UUG GAA UUU CUG AAG  2230
R   T   G   F   H   C   D   I   I   C   V   T   G   W   T    603
CGU ACC GGA UUU CAU UGU GAU AUC AUU UGU GUC ACU GGA UGG ACU  2275
A   T   P   M   Q   D   V   Q   V   T   I   D   W   F   I    618
GCC ACC CCA AUG CAG GAU GUU CAG GUU ACA AUU GAU UGG UUU AUU  2320
S   S   Q   E   C   V   P   R   T   Y   C   V   L   N   P    633
UCC UCU CAG GAA UGU GUU CCC AGG ACC UAU UGU GUU UUA AAU CCA  2365
Q   N   P   F   V   L   N   R   W   M   G   K   L   T   F    648
CAA AAU CCC UUU GUG UUA AAU AGG UGG AUG GGA AAA CUG ACU UUC  2410
P   Q   G   T   S   R   S   V   K   R   M   P   L   S   I    663
CCC CAG GGC ACU UCC CGA AGU GUU AAG AGA AUG CCU CUU UCU AUA  2455
G   G   G   A   G   A   K   N   A   I   L   M   N   M   P    678
GGG GGA GGA GCU GGU GCA AAG AAU GCU AUU CUC AUG AAU AUG CCA  2500
N   A   V   L   S   M   W   R   Y   F   V   G   D   L   V    693
AAU GCU GUU CUU UCA AUG UGG AGA UAU UUU GUU GGA GAU CUC GUC  2545
F   E   V   S   K   M   T   S   P   Y   I   K   C   T   V    708
UUU GAA GUU UCU AAG AUG ACU UCU CCC UAC AUU AAA UGU ACA GUC  2590
S   F   F   I   A   F   G   N   L   A   D   D   T   I   N    723
UCU UUC UUC AUA GCA UUU GGA AAU UUG GCU GAU GAU ACC AUC AAU  2635
F   E   A   F   P   H   K   L   V   Q   F   G   E   I   Q    738
UUU GAG GCU UUU CCC CAC AAG CUG GUG CAG UUU GGA GAG AUU CAG  2680
E   K   V   V   L   K   F   S   Q   E   E   F   L   T   A    753
GAA AAA GUU GUA UUG AAA UUU UCA CAA GAG GAA UUU CUU ACA GCU  2725
W   S   T   Q   V   R   P   A   T   T   L   L   A   D   G    768
UGG UCA ACU CAG GUG CGA CCU GCA ACA ACU CUG UUG GCU GAU GGG  2770
C   P   Y   L   Y   A   M   V   H   D   S   S   V   S   T    783
UGU CCA UAU UUG UAU GCU AUG GUA CAU GAU AGU UCA GUG UCU ACA  2815
I   P   G   D   F   V   I   G   V   K   L   T   T   I   N    798
AUA CCA GGU GAU UUU GUC AUU GGU GUU AAG UUG ACA ACC AUA AAC  2860
N   M   C   A   Y   G   L   N   P   G   I   S   G   S   R    813
AAU AUG UGU GCA UAU GGG CUC AAU CCU GGU AUU UCA GGU UCU CGU  2905
                          S-CP
L   L   G   T   I   P   Q  ▼S   I   S   Q   Q   T   V   W    828
CUU UUG GGC ACC AUU CCU CAG UCC AUU UCA CAG CAA ACU GUU UGG  2950
N   Q   M   A   T   V   R   T   P   L   N   F   D   S   S    843
AAU CAA AUG GCA ACA GUG AGA ACA CCA UUG AAU UUU GAU UCU AGC  2995
K   Q   S   F   C   Q   F   S   I   D   L   L   G   G   G    858
AAG CAG AGC UUU UGU CAA UUU UCU AUU GAC CUU CUC GGU GGA GGA  3040
I   L   V   D   K   T   G   D   W   I   T   L   I   Q   N    873
AUU UUA GUG GAC AAA ACU GGA GAU UGG AUC ACA CUU AUA CAA AAU  3085
S   P   I   S   N   L   L   R   V   A   A   W   K   K   G    888
UCU CCA AUU AGU AAC UUG UUG AGA GUU GCU GCU UGG AAG AAA GGC  3130
C   L   M   V   K   I   V   M   S   G   N   A   A   V   K    903
UGU UUA AUG GUU AAG AUU GUG AUG UCU GGG AAU GCA GCA GUC AAA  3175
R   S   D   W   A   S   L   V   Q   V   F   L   T   N   S    918
AGG AGU GAU UGG GCC UCA UUG GUA CAA GUG UUU UUA ACA AAC AGC  3220
N   S   T   E   H   F   D   A   C   K   W   T   K   S   E    933
AAC AGU ACA GAG CAU UUU GAU GCA UGU AAG UGG ACA AAA UCA GAA  3265
```

(Figure 2.2 continued)

```
P    H    S    W    E    L    I    F    P    I    E    V    C    G    P     948
CCA CAU UCC UGG GAA UUG AUC UUU CCA AUA GAG GUA UGU GGU CCU  3310
N    N    G    F    E    M    W    S    S    E    W    A    N    Q    T     963
AAU AAU GGU UUU GAA AUG UGG AGU UCU GAG UGG GCA AAU CAA ACU  3355
S    W    H    L    S    F    L    I    D    N    P    K    Q    S    T     978
UCA UGG CAU UUG AGU UUC CUU AUU GAC AAU CCC AAA CAG UCU ACA  3400
V    F    D    I    L    L    G    I    S    Q    D    F    E    I    A     993
GUU UUU GAU AUU CUC UUG GGA AUC UCU CAA GAU UUU GAA AUU GCU  3445
G    N    T    L    M    P    A    F    S    V    P    Q    A    T    A    1008
GGU AAU ACU CUU AUG CCA GCU UUU UCU GUU CCA CAG GCU ACU GCC  3490
R    S    S    E    N    A    E    S    S    A    *                        1019
AGA UCU UCU GAA AAU GCG GAA UCU UCU GCA UGA UCUGGAAUUUGUGUU  3538

UUCUUUCGCUUGUUCGUUUGUUUAAUUCAAUAAAGGAAAUUAGGCAUGACCCUCUUGUUG 3598

AGUAUGCUCUGCCUAUUUGAAAAUUUCCACACCUCUUUUAAUUGUCGUAAUGAUGUGUGA 3658

AGUGUGUGUUAUUUU                                                3673
```

```
UAUUAAAAUUUUCAUAAGAUUUGAAAUUUUGAUAAACCGCGAUCACAGGUUGCCGCACCU 60

UAAAACCGGAAACAAAAGCAAUCGUUACUUGUUCUCAAGAAUCUUCAAACUUUCUUCUGU 120

UUUCUCUGCACUCGGUUUUUCAGGAAAAGAAAAUCACUCUCUGUAUUGAUUACAGACUUC 180

GUGAAUCAUUUUCUUUUUCAACUUUCAGUUCACUUGCUGAACACUCUCCUAUAAAUAUAU 240

AGGACUUCGUGUCAGAUUUGAACUUUUCCUGUCUCUUUCUCGGUUUUCUUUCUUAUUCUC 300

AUCUUCUUUAAAUUUUAAGGCUCGUAUUUUGUUUCUUUAAACUUUCUGUUGGACUCAUUU 360
```

Co-pro

```
          ▼M   K   F   Y   P   G   Q   N   V   S   E   I     12
GAAAUACAAC AUG AAG UUU UAU CCA GGA CAA AAU GUC UCU GAA AUU   406
V   Y   H   F   Q   S   N   E   T   A   N   R   L   D   A    27
GUU UAU CAU UUU CAG AGU AAU GAG ACA GCU AAU AGG CUU GAU GCU  451
Y   F   A   C   G   C   E   E   D   T   E   V   L   A   R    42
UAU UUU GCC UGU GGC UGU GAA GAG GAU ACU GAA GUC CUC GCU CGC  496
L   K   Q   C   N   P   R   L   L   H   L   S   Y   A   A    57
UUG AAG CAA UGU AAC CCU CGG UUA CUG CAU UUA UCU UAU GCU GCU  541
F   C   L   E   M   G   S   H   S   V   E   E   I   E   Y    72
UUC UGU UUG GAA AUG GGC AGU CAU UCU GUU GAA GAA AUA GAA UAU  586
D   D   G   E   L   V   F   L   Y   F   Q   N   F   L   L    87
GAU GAU GGA GAG UUG GUU UUC UUG UAU UUC CAA AAU UUU CUA CUC  631
S   I   V   S   N   S   S   K   T   A   N   L   R   A   Y   102
UCC AUA GUG UCC AAU UCU UCC AAA ACA GCA AAU CUG AGA GCA UAC  676
I   R   S   A   F   A   Y   H   F   Q   H   F   V   E   F   117
AUA CGU UCA GCU UUU GCA UAU CAU UUU CAG CAU UUU GUU GAA UUU  721
D   Q   Y   T   N   D   S   L   N   V   M   D   T   S   V   132
GAU CAA UAU ACA AAU GAU UCU CUC AAU GUG AUG GAU ACA AGC GUA  766
S   A   Q   G   I   A   D   L   A   L   S   M   V   R   W   147
UCU GCU CAA GGA AUU GCA GAU UUG GCU CUG UCC AUG GUC AGA UGG  811
I   P   T   Q   I   K   K   V   V   N   F   G   V   G   S   162
AUU CCU ACU CAA AUU AAA AAA GUU GUG AAU UUU GGG GUA GGA UCG  856
V   I   E   S   F   S   E   H   F   N   K   L   I   M   Q   177
GUC AUA GAA UCC UUU UCA GAA CAU UUC AAU AAG CUC AUA AUG CAA  901
Y   C   P   I   V   F   Q   A   F   S   W   V   N   N   I   192
UAU UGU CCA AUA GUG UUC CAA GCU UUU AGC UGG GUC AAC AAU AUU  956
W   T   M   V   K   E   W   I   E   E   A   A   K   E   I   207
UGG ACC AUG GUU AAA GAA UGG AUU GAA GAG GCU GCG AAG GAA AUU 1001
```

Figure 2.3. The complete nucleotide sequence of BPMV K-Ha1 RNA1. The deduced amino acid sequence of the polyprotein encoded by RNA1 is indicated in the one-letter code below the nucleotide sequence. The positions of the proteolytic cleavage sites are indicated with arrowheads with the names of the cleavage products indicated at the start of the coding regions of each of the mature proteins.

(Figure 2.3. continued)

```
S   W   F   L   Q   G   C   K   E   L   L   A   W   G   M    222
UCU UGG UUC CUG CAG GGU UGU AAG GAA UUA UUA GCU UGG GGA AUG  1046
C   I   L   A   S   S   C   A   L   G   L   V   E   K   C    237
UGU AUU CUG GCU AGU UCC UGU GCU UUG GGA UUG GUU GAA AAA UGU  1091
L   I   S   L   G   M   I   S   E   S   F   D   L   V   G    252
CUC AUU UCU UUA GGC AUG AUU UCU GAA UCU UUU GAU UUG GUU GGU  1136
L   F   V   R   S   A   I   V   G   A   F   C   V   S   I    267
UUG UUU GUU CGA UCA GCC AUU GUU GGG GCC UUC UGU GUU UCU AUC  1181
K   T   G   K   F   V   S   N   S   E   L   I   T   C   A    282
AAG ACG GGC AAG UUU GUU UCA AAU AGU GAG UUG AUC ACA UGU GCU  1226
T   I   A   V   S   T   I   A   T   V   M   S   Q   A   F    297
ACC AUU GCA GUU UCU ACA AUU GCA ACU GUU AUG UCU CAA GCU UUC  1261
                                          Hel
K   P   S   E   E   I   K   G   Q   F   Q  ▼A   L   S   V    312
AAA CCU UCU GAA GAA AUU AAA GGG CAA UUC CAG GCU CUU UCU GUU  1306
L   E   G   L   A   T   Q   L   T   S   F   C   D   T   S    327
UUA GAG GGA UUG GCA ACA CAA CUC ACU UCA UUU UGU GAC ACA UCU  1351
L   I   A   M   G   K   T   C   T   A   F   N   Q   I   C    342
UUG AUU GCC AUG GGA AAA ACC UGC ACA GCA UUU AAU CAA AUU UGU  1396
T   A   G   K   N   V   K   V   I   A   G   R   L   L   D    357
ACU GCU GGG AAA AAU GUU AAA GUG AUU GCU GGC AGA UUG UUG GAU  1441
V   V   S   N   F   V   R   K   L   L   G   L   D   S   A    372
GUA GUU UCC AAU UUU GUA AGG AAA CUU UUG GGA UUG GAU AGU GCU  1486
F   L   R   D   A   A   L   I   F   S   Q   D   V   D   G    387
UUU CUU AGA GAU GCA GCG CUU AUU UUC UCU CAA GAU GUU GAC GGU  1531
W   L   R   N   I   S   W   C   Q   E   Q   F   L   L   K    402
UGG UUG CGC AAU AUC AGC UGG UGU CAG GAA CAA UUC CUA CUG AAA  1576
A   Y   M   S   Q   D   D   L   I   V   L   R   S   L   V    417
GCA UAC AUG UCU CAA GAU GAU CUU AUU GUC UUG CGU UCC UUA GUU  1621
V   K   G   E   R   M   R   E   Q   M   L   E   G   E   V    432
GUC AAA GGU GAA AGA AUG AGA GAG CAA AUG CUG GAG GGA GAG GUU  1666
K   V   S   P   S   V   C   N   L   I   V   K   G   C   E    447
AAA GUG UCU CCU AGU GUU UGU AAU CUU AUU GUA AAA GGU UGU GAA  1711
E   A   S   K   L   M   R   E   S   V   L   H   C   S   K    462
GAA GCA AGU AAG UUG AUG AGA GAA AGU GUG CUA CAU UGU UCA AAG  1756
T   V   R   K   I   P   F   V   I   F   A   H   G   D   S    477
ACU GUA CGA AAA AUU CCA UUU GUU AUU UUU GCA CAC GGU GAU UCU  1801
R   V   G   K   S   L   L   V   D   R   L   I   T   D   F    492
CGU GUU GGA AAA UCU UUG CUA GUU GAU AGA CUU AUC ACA GAU UUU  1846
C   D   H   L   E   I   G   E   D   A   V   Y   S   R   N    507
UGU GAU CAU CUA GAA AUU GGG GAG GAU GCU GUU UAU UCA AGG AAU  1891
P   S   D   P   F   W   S   G   Y   R   R   Q   P   I   V    522
CCU UCG GAU CCU UUC UGG AGU GGG UAU AGG AGA CAA CCA AUC GUC  1936
T   I   D   D   F   A   A   V   V   S   E   P   S   A   E    537
ACU AUU GAU GAU UUU GCU GCU GUU GUA UCA GAG CCA UCU GCU GAG  1981
A   Q   L   I   P   L   I   S   S   A   P   Y   P   L   N    552
GCU CAA UUG AUU CCA UUA AUU UCA AGU GCU CCA UAC CCA UUG AAC  2026
M   A   S   L   E   E   K   G   M   H   F   D   S   Q   I    567
AUG GCA AGU UUA GAG GAA AAG GGA AUG CAU UUU GAU UCU CAG AUC  2071
M   M   C   S   S   N   F   L   E   P   S   P   E   A   K    582
AUG AUG UGC UCU UCA AAU UUU UUG GAA CCU UCU CCU GAA GCC AAA  2116
```

(Figure 2.3. continued)

```
I   R   D   D   M   A   F   R   N   R   R   H   V   L   I    597
AUU AGA GAU GAC AUG GCU UUU AGA AAU AGA AGA CAU GUU UUG AUA  2161
T   V   E   L   K   P   G   V   E   Y   D   E   S   D   F    612
ACA GUU GAG CUU AAA CCU GGA GUG GAG UAU GAU GAG AGC GAU UUU  2206
T   K   N   Q   R   Y   L   L   K   T   W   F   H   D   H    627
ACC AAA AAU CAG CGG UAU UUA CUC AAA ACU UGG UUU CAU GAU CAU  2251
Y   V   V   D   Q   T   F   E   S   Y   A   D   L   L   A    642
UAU GUU GUA GAC CAA ACU UUU GAA UCU UAU GCU GAU CUU UUG GCA  2296
Y   C   F   T   K   W   E   R   H   V   K   E   Q   E   S    657
UAU UGC UUC ACU AAA UGG GAG AGA CAU GUG AAG GAA CAA GAG UCU  2341
N   L   S   Q   I   K   G   K   K   S   E   S   G   H   F    672
AAU UUG UCU CAA AUU AAA GGC AAG AAA AGU GAG AGU GGC CAU UUC  2386
N   N   F   Q   Q   L   M   D   L   A   V   S   W   N   L    687
AAU AAU UUU CAA CAA CUU AUG GAU UUG GCA GUU UCA UGG AAU CUC  2431
S   A   N   I   M   K   E   R   I   K   A   D   K   S   D    702
AGU GCA AAU AUC AUG AAG GAA CGA AUC AAA GCU GAU AAA AGU GAU  2476
M   V   Y   V   F   S   A   G   R   K   D   K   I   V   H    717
AUG GUU UAU GUC UUC UCU GCU GGA AGG AAG GAC AAA AUU GUA CAU  2521
C   F   L   N   K   E   G   E   C   S   I   R   P   D   S    732
UGC UUU UUG AAC AAA GAA GGC GAA UGU AGU AUA CGU CCU GAU UCA  2566
I   E   D   P   E   A   Q   L   L   L   K   A   S   E   T    747
AUA GAA GAU CCA GAA GCA CAA CUC UUG CUC AAA GCU UCA GAA ACU  2611
M   L   M   K   A   Y   A   F   L   K   Y   N   N   A   T    762
AUG CUC AUG AAA GCU UAU GCU UUU UUG AAG UAC AAC AAU GCG ACC  2656
N   L   I   V   R   T   H   L   A   E   L   V   N   E   D    777
AAU UUG AUU GUC AGG ACC CAU UUG GCA GAA UUG GUC AAU GAA GAC  2701
F   Y   D   E   K   F   N   F   I   G   T   I   G   T   P    792
UUU UAU GAU GAA AAG UUU AAU UUU AUU GGC ACA AUU GGG ACU CCA  2746
A   F   H   R   Q   I   A   A   H   L   E   K   M   P   L    807
GCU UUU CAU CGA CAA AUU GCU GCA CAC UUG GAG AAA AUG CCA UUA  2791
W   Q   K   A   I   L   C   G   M   G   H   C   L   S   R    822
UGG CAA AAA GCA AUU UUG UGU GGA AUG GGA CAU UGU UUG UCU CGG  2836
K   S   K   E   T   W   Y   S   G   M   K   E   K   F   V    837
AAA AGC AAA GAG ACA UGG UAU UCU GGG AUG AAG GAA AAA UUU GUA  2881
Q   M   M   K   S   I   Y   E   T   E   V   T   D   W   P    852
CAA AUG AUG AAG AGC AUC UAU GAA ACA GAA GUU ACA GAU UGG CCA  2926
V   P   L   K   I   I   S   G   T   I   L   A   T   I   L    867
GUA CCA CUG AAA AUC AUU UCU GGA ACC AUU CUU GCA ACA AUU UUG  2971
G   T   T   F   W   K   L   F   S   F   L   R   D   A   G    882
GGA ACA ACC UUC UGG AAA CUC UUU UCC UUU CUU AGA GAU GCU GGU  3016
N   G   G   V   F   V   G   N   V   A   S   A   F   T   T    897
AAU GGG GGA GUU UUU GUU GGU AAU GUU GCU UCA GCA UUC ACA ACU  3061
                          VPg
S   S   V   L   E   A   Q  ▼S   R   K   P   N   R   Y   E    912
UCA AGU GUG CUU GAG GCC CAA AGU AGA AAG CCC AAC AGA UAU GAA  3106
V   S   Q   Y   R   Y   R   N   V   P   I   K   R   R   A    927
GUU UCC CAA UAC AGA UAU CGC AAU GUG CCA AUA AAG CGC AGA GCA  3151
                  Pro
W   V   E   G   Q  ▼M   S   F   D   Q   S   V   V   A   I    942
UGG GUU GAA GGC CAG AUG UCU UUC GAU CAG UCU GUA GUA GCA AUU  3196
M   S   K   C   K   A   S   M   R   M   G   N   T   D   A    957
AUG UCU AAA UGC AAA GCC AGC AUG AGA AUG GGA AAC ACU GAU GCU  3241
```

(Figure 2.3 continued)

```
Q   I   L   M   V   P   G   R   R   F   I   A   H   G   H    972
CAA AUC UUG AUG GUU CCA GGG CGC AGA UUC AUA GCU CAU GGA CAU  3286
F   F   K   N   L   T   Q   K   V   R   V   Q   I   V   T    987
UUU UUC AAA AAU CUG ACU CAG AAA GUG CGA GUC CAG AUU GUU ACA  3331
S   E   K   S   Y   W   H   V   Y   D   P   D   K   F   Q    1002
UCU GAA AAG AGU UAC UGG CAU GUG UAU GAU CCU GAC AAG UUU CAG  3376
M   F   D   N   S   E   I   G   L   Y   S   N   P   T   L    1017
AUG UUU GAC AAC AGU GAA AUA GGU CUU UAU UCU AAU CCC ACU UUG  3421
E   D   I   P   H   S   A   W   D   L   F   C   W   D   S    1032
GAA GAU AUC CCA CAU UCA GCU UGG GAC CUU UUU UGC UGG GAC AGU  3466
E   K   T   L   P   N   N   F   S   A   E   L   L   S   C    1047
GAG AAA ACU UUG CCA AAU AAU UUU UCU GCA GAA UUG CUC UCU UGC  3511
K   L   D   T   V   T   G   Q   Y   Y   P   E   W   A   P    1062
AAA UUG GAU ACU GUU ACU GGC CAA UAC UAU CCU GAG UGG GCU CCA  3556
I   N   C   R   V   H   R   Q   P   I   H   I   T   E   G    1077
AUU AAU UGU CGA GUU CAU CGA CAA CCA AUU CAC AUC ACU GAA GGA  3601
N   Y   V   R   K   Q   D   V   S   I   E   Y   D   A   C    1092
AAU UAU GUC AGA AAG CAA GAU GUC AGU AUU GAA UAU GAU GCA UGU  3646
T   I   P   N   D   C   G   S   L   V   V   A   K   V   G    1107
ACA AUU CCA AAU GAU UGU GGU UCA UUG GUU GUU GCC AAG GUU GGA  3691
N   H   K   Q   I   V   G   F   H   V   A   G   S   K   G    1122
AAU CAC AAA CAA AUU GUU GGU UUC CAU GUU GCU GGA AGC AAA GGA  3736
R   L   G   Y   A   S   L   I   P   Y   V   E   P   V   V    1137
AGA CUG GGA UAU GCU UCA UUG AUA CCA UAU GUU GAG CCA GUC GUG  3781
                RdRp
Q   A   Q  ▼S   A   E   V   Y   F   D   F   F   P   V   E   1152
CAA GCU CAA AGU GCU GAA GUU UAC UUU GAU UUC UUC CCU GUG GAG  3826
V   D   S   Q   E   G   V   A   H   I   G   E   L   K   S    1167
GUU GAU AGU CAA GAG GGA GUU GCU CAC AUU GGU GAA UUG AAA UCU  3871
G   V   Y   V   P   L   P   T   K   T   N   L   V   E   T    1182
GGU GUC UAU GUU CCA CUG CCU ACA AAA ACU AAU UUG GUG GAA ACU  3916
P   K   E   W   Q   L   D   L   P   C   D   K   I   P   S    1197
CCC AAG GAA UGG CAA CUG GAU CUA CCU UGU GAU AAA AUU CCU AGU  3951
V   L   T   T   T   D   E   R   L   V   G   T   E   H   E    1212
GUC UUG ACU ACA ACU GAU GAG AGA UUG GUG GGC ACA GAA CAU GAG  3996
G   Y   D   P   F   L   G   G   I   Q   K   Y   A   T   P    1227
GGA UAU GAU CCU UUU CUU GGU GGA AUU CAA AAA UAU GCC ACU CCC  4041
M   M   P   L   D   E   E   I   L   S   K   V   A   Q   D    1242
AUG AUG CCC CUA GAU GAA GAA AUU CUU UCU AAG GUU GCA CAA GAU  4086
M   V   E   E   W   F   D   C   V   D   E   E   D   S   F    1257
AUG GUU GAG GAA UGG UUU GAU UGU GUU GAU GAG GAG GAU UCC UUU  4141
E   E   V   S   L   S   A   A   L   N   G   V   E   G   L    1272
GAG GAA GUU UCU UUA AGU GCA GCA CUC AAU GGU GUU GAG GGU UUG  4186
D   Y   M   E   R   I   P   L   A   T   S   E   G   F   P    1287
GAC UAU AUG GAA AGA AUU CCU CUU GCC ACA UCU GAG GGU UUU CCU  4231
H   V   L   S   R   K   N   G   E   K   G   K   R   R   F    1302
CAU GUG CUU UCA CGC AAA AAU GGU GAA AAA GGC AAA AGG AGA UUU  4276
V   S   G   D   G   E   E   M   T   L   I   P   G   T   S    1317
GUU UCU GGG GAU GGU GAA GAG AUG ACA UUG AUC CCG GGA ACC AGU  4321
V   E   E   A   Y   N   K   L   I   V   E   L   E   K   S    1332
GUU GAA GAG GCU UAC AAC AAG CUA AUA GUU GAA CUU GAA AAA AGU  4366
V   P   T   L   V   G   I   E   C   P   K   D   E   K   L    1347
```

(Figure 2.3 continued)

```
GUU CCU ACA UUA GUU GGC AUU GAA UGU CCC AAG GAU GAG AAA CUU  4411
P   R   R   K   I   F   D   K   P   K   T   R   C   F   T    1362
CCU CGU CGC AAA AUU UUU GAC AAA CCU AAG ACG CGC UGC UUC ACU  4456
I   L   P   M   E   F   N   L   V   V   R   Q   K   F   L    1377
AUU CUC CCU AUG GAG UUU AAU CUU GUG GUU CGU CAA AAG UUU UUA  4501
N   F   V   R   F   I   M   K   K   R   D   K   L   S   C    1392
AAU UUU GUG AGG UUC AUU AUG AAG AAA AGG GAC AAG CUU AGU UGU  4546
Q   V   G   I   N   P   Y   S   M   E   W   T   G   L   A    1407
CAA GUC GGA AUC AAC CCA UAC UCC AUG GAA UGG ACU GGU UUG GCC  4591
N   R   L   L   S   K   G   N   D   I   L   C   C   D   Y    1422
AAU AGA UUA UUG AGU AAG GGC AAC GAU AUU CUG UGU UGC GAU UAU  4636
A   S   F   D   G   L   I   T   K   Q   V   M   S   K   M    1437
GCU AGU UUU GAU GGU UUG AUU ACC AAG CAA GUU AUG AGC AAG AUG  4681
A   E   M   I   N   S   L   C   G   G   D   E   K   L   M    1452
GCA GAA AUG AUA AAC AGU CUU UGU GGU GGU GAU GAA AAA UUG AUG  4726
R   E   R   T   H   L   L   L   A   C   C   S   R   M   A    1467
CGU GAA AGG ACA CAC CUA CUG UUG GCU UGU UGU UCA AGG AUG GCA  4771
I   C   K   K   D   V   W   R   V   E   C   G   I   P   S    1482
AUU UGU AAG AAA GAU GUU UGG AGG GUU GAA UGU GGA AUC CCU UCU  4816
G   F   P   L   T   V   I   C   N   S   I   F   N   E   M    1497
GGA UUU CCG CUC ACA GUU AUU UGC AAU AGC AUU UUU AAU GAG AUG  4861
L   I   R   Y   S   Y   E   K   L   L   R   Q   A   K   A    1512
CUC AUU AGA UAU AGC UAU GAA AAA UUA CUG CGU CAG GCU AAA GCU  4906
P   S   M   F   L   Q   S   F   K   N   F   V   S   L   C    1527
CCA AGU AUG UUU UUA CAA UCC UUC AAA AAU UUU GUU UCU UUG UGU  4951
V   Y   G   D   D   N   L   I   S   V   H   E   Y   V   K    1542
GUG UAU GGU GAU GAC AAC UUA AUU AGU GUU CAU GAA UAU GUC AAG  4996
P   Y   F   S   G   S   K   L   K   S   F   L   A   G   H    1557
CCU UAC UUU AGU GGU UCU AAA UUG AAA AGC UUU CUA GCU GGU CAU  5041
N   I   T   I   T   D   G   I   D   K   T   S   A   T   L    1572
AAU AUC ACC AUU ACU GAU GGU AUU GAC AAA ACU AGU GCA ACU UUG  5086
Q   F   R   K   L   A   D   C   D   F   L   K   R   N   F    1587
CAA UUU AGA AAG UUA GCA GAU UGU GAU UUU CUC AAG AGA AAC UUU  5131
K   Q   M   S   N   V   L   W   V   A   P   E   D   K   A    1602
AAG CAG AUG UCU AAU GUU UUG UGG GUG GCU CCU GAG GAC AAG GCG  5176
S   L   W   S   Q   L   H   Y   V   S   C   N   N   L   E    1617
AGU UUA UGG UCA CAA CUU CAU UAU GUU UCG UGU AAU AAU CUG GAA  5221
M   Q   E   A   Y   L   V   N   L   V   N   V   L   R   E    1632
AUG CAA GAA GCU UAU CUU GUC AAU CUC GUA AAU GUG UUA CGA GAA  5266
L   Y   L   H   S   P   E   E   A   R   Q   L   R   R   K    1647
UUG UAU UUG CAC AGU CCA GAG GAA GCU CGC CAA UUG AGA AGG AAA  5311
A   L   S   R   I   E   W   L   Q   K   A   D   V   P   T    1662
GCU CUC UCU CGU AUC GAA UGG CUG CAA AAA GCU GAU GUG CCU ACC  5356
I   A   Q   I   E   E   F   H   S   M   Q   R   M   M   N    1677
AUA GCA CAG AUU GAA GAG UUC CAU UCG AUG CAG AGG AUG AUG AAU  5401
A   P   D   S   N   D   N   I   D   L   L   L   S   I   D    1692
GCU CCU GAU UCA AAU GAU AAU AUU GAC CUA CUG UUG AGC AUU GAU  5446
L   L   G   L   Q   G   A   G   K   A   F   P   N   K   I    1707
UUG UUG GGU UUA CAA GGA GCA GGU AAA GCU UUU CCA AAC AAG AUU  5491
V   F   D   D   K   L   V   L   A   N   T   Q   E   F   F    1722
GUU UUU GAU GAU AAG CUU GUG UUG GCU AAC ACA CAA GAA UUC UUU  5536
```

(Figure 2.3 continued)

```
D   G   N   F   P   V   D   S   W   L   P   I   F   V   N    1737
GAU GGA AAU UUU CCA GUG GAU UCU UGG UUA CCA AUU UUU GUG AAU  5581
C   L   Y   P   V   S   Q   L   P   S   E   A   V   V   V    1752
UGU CUU UAU CCU GUA AGU CAA UUA CCU UCA GAA GCU GUU GUU GUA  5626
N   V   T   C   G   S   G   R   G   G   L   P   T   T   A    1767
AAU GUC ACA UGU GGU AGU GGA CGU GGU GGU UUA CCC ACC ACU GCU  5671
W   I   S   S   A   V   N   N   R   S   S   D   I   N   K    1782
UGG AUU AGU UCU GCA GUU AAC AAU CGC UCC UCA GAU AUC AAC AAA  5716
K   I   R   T   A   L   G   K   G   K   K   I   V   F   L    1797
AAG AUU CGC ACA GCA CUU GGG AAA GGU AAG AAA AUU GUU UUU CUU  5761
T   R   V   D   P   F   P   V   A   L   L   A   V   L   F    1812
ACU AGA GUU GAU CCU UUU CCA GUA GCU UUA CUG GCU GUU CUU UUU  5806
G   V   K   N   E   I   L   S   S   N   A   T   N   P   M    1827
GGC GUU AAG AAU GAA AUU CUG AGU UCU AAC GCC ACU AAC CCU AUG  5851
L   T   R   L   L   E   N   C   K   S   L   K   Y   L   V    1842
CUG ACG CGA CUU CUU GAG AAC UGU AAG AGU CUC AAA UAU CUG GUU  5896
D   E   C   P   F   A   F   V   N   *                        1851
GAU GAG UGU CCU UUU GCA UUU GUG AAC UAG UAUGUAAUAUUUUAUUCAC  5955

UUAAAUAAAGCGCAUUACUAUGUGCAAUGAGUGUGUUUAAAUAU               5989
```

```
UAUUAAAAUUUUCAUAAGAUUUGAAAUUUUGAUAAACCGCGAUCACAGGUUGCCGCACCU 60

UAAAACCGGAAACAAAAGCAAUCGUUACUUGAUUUUAAGAAUUUCUCAAUUUCUUCCUAC 120

UUCUUUGUGUACGAUUUCUCAAGGGAAAGAAAAUCACUCUCUGUGCUGGUCACAGACUUC 180

GUGAAUCAUUUUCUUUUCCACUCUUAGUUUAUUUGCUGAACACUCUCCUAUUUGAUAUAG 240

GACUUCGUGUCAGAUUUAAACUUCUUCUGUUUCUUUCUCAGUUCUCUGUUUAAUUUCAAG 300

UUCAAGCUGGUGAAAUUUUGGAUUAGUGCUCCCACUCUCCUAUCUGGUAUAGGACUUCGU 360

GGGUAGACUUUUCUAUUUCUCUCCUUUCUUUCACUCUCUUCUUCUCACUGAUCCGCAUUG 420
                                                          CR
                                                        ▼M   F   A     3
CCGUUCAAAGUGGUCUUAUUUGAAAAACACUUGGGCGUUGGUGCAA AUG UUU GCU   475
S   L   I   F   S   G   D   N   R   L   T   E   K   T   I    18
UCG UUA AUU UUC UCU GGA GAU AAC AGG CUC ACU GAG AAA ACA AUU  520
F   T   C   R   D   L   D   I   L   V   V   Y   Y   T   I    33
UUU ACU UGC AGA GAU UUG GAC AUU UUG GUU GUU UAU UAU ACA AUA  565
A   T   Q   F   R   K   F   L   P   H   Y   I   R   W   H    48
GCA ACC CAA UUU AGA AAA UUU CUA CCG CAU UAU AUU AGG UGG CAU  610
L   Y   T   L   L   I   Y   I   L   P   S   F   L   T   A    63
CUG UAU ACC UUG UUG AUC UAC AUU CUC CCA UCU UUU CUC ACU GCU  655
E   I   K   Y   K   R   N   L   S   N   I   H   I   S   G    78
GAA AUU AAA UAU AAG CGG AAU CUG AGU AAU AUU CAU AUU UCC GGC  700
L   F   Y   D   G   R   Y   K   F   W   T   K   H   E   K    93
UUA UUU UAC GAC GGC AGA UAC AAA UUC UGG ACU AAA CAC GAG AAA  745
                                    MP
N   L   A   L   T   E   E   E   K ▼M   E   V   I   R   N   108
AAU CUU GCU UUG ACA GAA GAG GAA AAG AUG GAA GUG AUU AGA AAU  790
R   G   I   P   A   D   V   L   A   K   R   A   H   E   F   123
AGA GGC AUU CCU GCU GAU GUU CUU GCA AAG CGA GCU CAU GAA UUU  835
E   K   H   V   A   H   E   S   L   K   D   Q   I   P   A   138
GAA AAA CAU GUU GCU CAU GAA AGC CUC AAG GAU CAA AUU CCU GCU  880
V   D   K   L   Y   S   T   K   V   N   K   F   A   K   I   153
GUU GAC AAG UUG UAU UCU ACU AAG GUU AAU AAG UUU GCA AAA AUU  925
M   N   L   R   Q   S   V   V   G   D   L   K   L   L   T   168
AUG AAC CUU AGA CAA AGU GUU GUU GGU GAU CUU AAA CUU CUU ACU  970
D   G   K   L   Y   E   G   K   H   I   P   V   S   N   I   183
GAU GGG AAG UUG UAU GAG GGU AAG CAU AUC CCU GUA UCU AAU AUU  1015
```

Figure 2.4. The complete nucleotide sequence of BPMV K-Ha1RNA2. The deduced amino acid sequence of the polyprotein encoded by RNA2 is indicated in the one-letter code below the nucleotide sequence. The positions of the proteolytic cleavage sites are indicated with arrowheads with the names of the cleavage products indicated at the start of the coding regions of each of the mature proteins.

(Figure 2.4 continued)

```
S   A   G   E   N   H   V   V   Q   I   P   L   M   A   Q    198
AGU GCA GGG GAA AAU CAU GUA GUU CAA AUA CCC UUA AUG GCA CAG  1060
E   E   I   L   S   S   S   A   S   D   F   R   T   A   M    213
GAG GAA AUU CUG UCU UCU AGU GCA AGC GAU UUC AGA ACU GCA AUG  1105
V   S   K   N   S   K   P   Q   A   T   A   M   H   V   G    228
GUG AGU AAA AAU AGC AAG CCU CAA GCU ACU GCA AUG CAU GUG GGA  1150
A   I   E   I   I   I   D   S   F   A   S   P   D   C   N    243
GCU AUA GAA AUU AUC AUU GAU AGU UUC GCA AGU CCU GAC UGC AAC  1195
I   V   G   A   M   L   L   V   D   T   Y   H   T   N   P    258
AUA GUU GGU GCA AUG CUU UUG GUU GAU ACU UAU CAU ACC AAU CCU  1240
E   N   A   V   R   S   I   F   V   A   P   F   R   G   G    273
GAA AAU GCA GUU CGU AGU AUU UUU GUU GCG CCU UUC AGA GGC GGA  1285
R   P   I   R   V   V   T   F   P   N   T   I   V   Q   I    288
AGG CCC AUU CGG GUG GUU ACA UUU CCG AAU ACC AUU GUG CAG AUU  1330
E   P   D   M   N   S   R   F   Q   L   L   S   T   T   T    303
GAA CCA GAC AUG AAU UCA AGG UUU CAG CUU UUG AGU ACC ACU ACC  1375
N   G   D   F   V   Q   G   K   D   L   A   M   V   K   V    318
AAU GGU GAU UUU GUU CAA GGA AAA GAU CUC GCA AUG GUU AAA GUU  1420
N   V   A   C   A   A   V   G   L   T   S   S   Y   T   P    333
AAU GUA GCA UGU GCU GCC GUU GGC UUG ACA UCA AGU UAU ACU CCA  1465
T   P   L   L   E   S   G   L   Q   K   D   R   G   L   I    348
ACU CCA UUG UUG GAA UCU GGU UUG CAA AAA GAC AGA GGG UUA AUU  1510
V   E   Y   F   G   R   M   S   Y   V   A   H   N   V   N    363
GUG GAG UAU UUU GGA AGG AUG UCU UAC GUU GCU CAU AAC GUU AAU  1555
Q   P   Q   E   K   D   L   L   E   G   N   F   S   F   D    378
CAG CCC CAA GAG AAA GAU UUG UUG GAG GGA AAU UUU UCC UUU GAU  1600
I   K   S   L   S   R   L   E   K   V   S   S   T   K   A    393
AUU AAA UCU CUC UCU AGA UUG GAA AAG GUU UCC UCU ACA AAA GCA  1945
Q   F   V   S   G   K   T   F   K   Y   D   I   I   G   A    408
CAA UUU GUU AGU GGA AAA ACC UUC AAA UAU GAU AUA AUU GGU GCU  1690
G   S   H   S   S   E   D   F   P   E   K   G   D   Q   E    423
GGU UCA CAU UCU UCA GAA GAU UUU CCU GAA AAG GGA GAU CAA GAA  1735
K   P   K   K   I   D   A   R   L   R   Q   R   I   D   P    438
AAA CCC AAA AAG AUU GAU GCC AGA UUG AGA CAA AGA AUA GAU CCU  1780
                              L-CP
Q   Y   N   E   V   Q   A   Q  ▼M   E   T   N   L   F   K    453
CAA UAC AAU GAG GUU CAG GCU CAG AUG GAA ACA AAU UUG UUU AAA  1825
L   S   L   D   D   V   E   T   P   K   G   S   M   L   D    468
UUG UCU CUU GAU GAU GUU GAA ACU CCU AAA GGU UCC AUG UUG GAU  1870
L   K   I   S   Q   S   K   I   A   L   P   K   N   T   V    483
CUU AAA AUU UCU CAA UCU AAA AUU GCA CUU CCC AAG AAC ACA GUU  1915
G   G   T   I   L   R   S   D   L   L   A   N   F   L   T    498
GGA GGA ACC AUU CUG CGU AGU GAU CUA UUG GCA AAU UUU UUG ACA  1960
E   G   N   F   R   A   S   V   D   L   Q   R   T   H   R    513
GAG GGC AAU UUU AGA GCA AGU GUU GAU UUG CAG CGC ACU CAU CGU  2005
I   K   G   M   I   K   M   V   A   T   V   G   I   P   E    528
AUU AAA GGA AUG AUU AAA AUG GUG GCC ACA GUU GGU AUU CCU GAG  2050
N   T   G   I   A   L   A   C   A   M   N   S   S   F   R    543
AAU ACA GGU AUA GCA UUG GCC UGU GCU AUG AAU AGU UCU UUU AGG  2095
G   R   A   S   S   D   I   Y   T   I   C   S   Q   D   C    558
GGG CGC GCC AGU UCU GAU AUU UAC ACC AUC UGC UCU CAA GAC UGU  2140
```

(Figure 2.4 continued)

```
E   L   W   N   P   A   C   T   K   A   M   T   M   S   F    573
GAA UUA UGG AAU CCU GCU UGC ACA AAA GCA AUG ACU AUG UCA UUU  2185
N   P   N   P   C   S   D   A   W   S   L   E   F   L   K    588
AAU CCA AAC CCG UGU UCU GAU GCA UGG AGU UUG GAA UUU CUG AAG  2230
R   T   G   F   H   C   D   I   I   C   V   T   G   W   T    603
CGU ACU GGA UUU CAU UGU GAU AUC AUU UGU GUC ACU GGA UGG ACU  2275
A   T   P   M   Q   D   V   Q   V   T   I   D   W   F   I    618
GCC ACC CCA AUG CAG GAU GUU CAG GUU ACA AUU GAU UGG UUU AUU  2320
S   S   Q   E   C   V   P   R   T   Y   C   V   L   N   P    633
UCC UCU CAG GAA UGU GUU CCC AGG ACC UAC UGU GUU UUA AAU CCA  2365
Q   N   P   F   V   L   N   R   W   M   G   K   L   T   F    648
CAA AAU CCU UUU GUG UUA AAU AGG UGG AUG GGA AAA CUG ACU UUC  2410
P   Q   G   T   S   R   S   V   K   R   M   P   L   S   I    663
CCC CAG GGC ACU UCC CGG AGU GUU AAG AGA AUG CCU CUU UCU AUA  2455
G   G   G   A   G   A   K   N   A   I   L   M   N   M   P    678
GGG GGA GGA GCU GGU GCA AAG AAU GCU AUU CUC AUG AAU AUG CCA  2500
N   A   V   L   S   M   W   R   Y   F   V   G   D   L   V    693
AAU GCU GUU CUU UCA AUG UGG AGA UAU UUU GUU GGA GAU CUC GUC  2545
F   E   V   S   K   M   T   S   P   Y   I   K   C   T   V    708
UUU GAA GUU UCU AAG AUG ACU UCU CCC UAC AUU AAA UGU ACA GUC  2590
S   F   F   I   A   F   G   N   L   A   D   D   T   I   N    723
UCU UUC UUC AUA GCA UUU GGA AAU UUG GCU GAU GAU ACC AUC AAU  2635
F   E   A   F   P   H   K   L   V   Q   F   G   E   I   Q    738
UUU GAG GCU UUU CCC CAC AAG CUG GUG CAG UUU GGA GAG AUU CAG  2680
E   K   V   V   L   K   F   S   Q   E   E   F   L   T   A    753
GAA AAA GUU GUA UUG AAA UUU UCA CAA GAG GAA UUU CUU ACA GCU  2725
W   S   T   Q   V   R   P   A   T   T   L   L   A   D   G    768
UGG UCA ACU CAG GUG CGU CCU GCA ACA ACU CUG UUG GCU GAU GGG  2770
C   P   Y   L   Y   A   M   V   H   D   S   S   V   S   T    783
UGU CCA UAU UUG UAU GCU AUG GUG CAU GAU AGU UCA GUG UCU ACA  2815
I   P   G   D   F   V   I   G   V   K   L   T   T   I   N    798
AUA CCA GGU GAU UUU GUC AUU GGU GUU AAG UUG ACA ACC AUA AAC  2860
N   M   C   A   Y   G   L   N   P   G   I   S   G   S   R    813
AAU AUG UGU GCA UAC GGG CUU AAU CCU GGU AUU UCA GGU UCU CGU  2905
                        S-CP
L   L   G   T   I   P   Q  ▼S   I   S   Q   Q   T   V   W    828
CUU UUG GGC ACC AUU CCU CAG UCC AUU UCA CAG CAA ACU GUU UGG  2950
N   Q   M   A   T   V   R   T   P   L   N   F   D   S   S    843
AAU CAA AUG GCA ACA GUG AGA ACA CCA UUG AAU UUU GAU UCU AGC  2995
K   Q   S   F   C   Q   F   S   I   D   L   L   G   G   G    858
AAG CAG AGC UUU UGU CAA UUU UCU AUU GAC CUU CUC GGU GGA GGA  3040
I   L   V   D   K   T   G   D   W   I   T   L   I   Q   N    873
AUU UUA GUG GAC AAA ACU GGA GAU UGG AUC ACA CUU AUA CAA AAU  3085
S   P   I   S   N   L   L   R   V   A   A   W   K   K   G    888
UCU CCA AUU AGU AAC UUG UUG AGA GUU GCU GCU UGG AAG AAA GGC  3130
C   L   M   V   K   I   V   M   S   G   N   A   A   V   K    903
UGU UUA AUG GUC AAG AUU GUG AUG UCU GGG AAU GCA GCA GUC AAA  3175
R   S   D   W   A   S   L   V   Q   V   F   L   T   N   S    918
AGG AGU GAU UGG GCC UCA UUG GUA CAA GUG UUU UUA ACA AAC AGC  3220
N   S   T   E   H   F   D   A   C   K   W   T   K   S   E    933
AAC AGU ACA GAG CAU UUU GAU GCA UGU AAG UGG ACA AAA UCA GAA  3265
```

(Figure 2.4 continued)

```
P    H    S    W    E    L    I    F    P    I    E    V    C    G    P     948
CCA CAU UCC UGG GAA UUG AUC UUC CCA AUA GAG GUA UGU GGU CCU  3310
N    N    G    F    E    M    W    S    S    E     W    A    N    Q    T    963
AAU AAU GGU UUU GAA AUG UGG AGU UCU GAG UGG GCA AAU CAA ACU  3355
S    W    H    L    S    F    L    I    D    N    P    K    Q    S    T    978
UCA UGG CAU UUG AGU UUC CUU AUU GAC AAU CCC AAA CAG UCU ACA  3400
V    F    D    I    L    L    G    I    S    Q    D    F    E    I    A    993
GUU UUU GAU AUU CUC UUG GGA AUC UCU CAA GAU UUU GAA AUU GCU  3445
G    N    T    L    M    P    A    F    S    V    P    Q    A    T    A    1008
GGU AAU ACU CUU AUG CCA GCU UUU UCU GUU CCA CAA GCU ACU GCC  3490
R    S    S    E    N    A    E    S    S    A    *                        1019
AGA UCU UCU GAA AAU GCG GAA UCC UCU GCA UGA UCUGGAAUUUGUGUU  3538

UCCUUUUGCUUGUUCGUUUGUUUAAUUUAAUAAAGGAAAUUAGGCAUGACCCUCUUGUUG 3598

AGUAUGCUCUGCCUAUUUGAAAAUUUCCACACCUCUUUUAAUUGUCGUAAUGAUGUGUGA 3658

AGUGUGUGUUAUUUU                                                3673
```

A

RNA1 / RNA2	K-G7	K-Ha1	K-Ho1
K-G7		85.5	98.1
K-Ha1	86.9		85.9
K-Ho1	87.2	98.8	

B

RNA1 / RNA2	K-G7	K-Ha1	K-Ho1
K-G7		95.9	98.0
K-Ha1	96.3		97.3
K-Ho1	96.6	99.6	

Figure 2.5. Percentage nucleotide and deduced amino acid identity of RNA1and RNA2 between BPMV strains. (A) Full-length RNA1, above diagonal, and full-length RNA2, below diagonal. (B) Deduced amino acid sequence identity of polyproteins encoded by RNA1 (above diagonal) and RNA2 (below diagonal). Values are the identity scores generated by the GAP program in the UWGCG package.

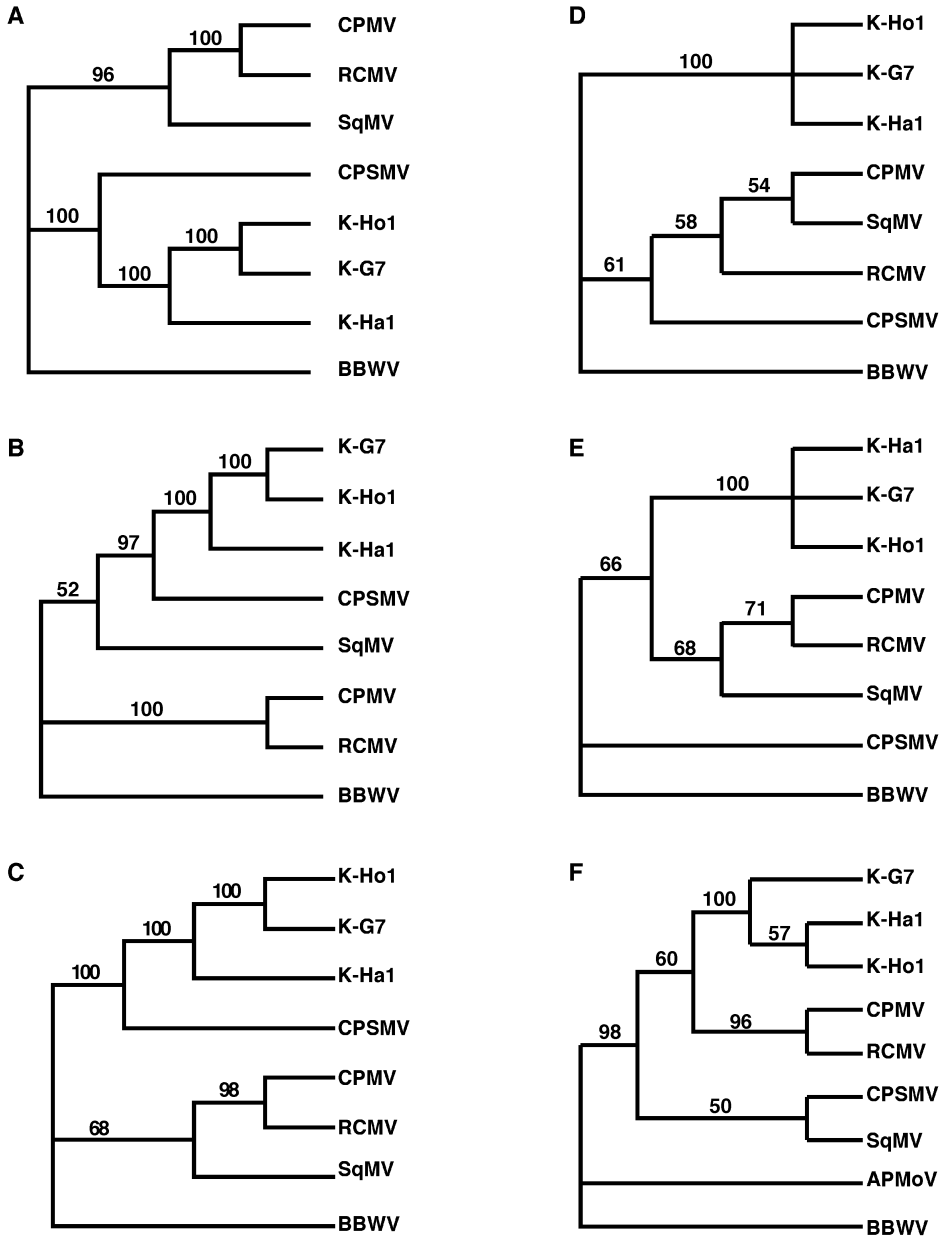

Figure 2.6. Phylogeny relationships of *comovirus* RNA1-encoded polyproteins and individual mature proteins. Amino acid sequences were aligned using the Clustal X Multiple Alignment Program and phylogenetic trees were constructed using the neighbor-joining method, as implemented in the PAUP program. The resulting consensus trees of 1000 bootstrap replicates are shown for: (A) RNA1 encoded precursor polyproteins, (B) Co-pro, (C) Hel, (D) VPg, (E) Pro, and (F) RdRp. The bootstrap values are indicated above each branch. *Broad bean wilt virus* (BBMV) RNA sequence was derived from Kobayashi et al. (1999).

Protease cofactor

	BPMV K-G7	BPMV K-Ha1	BPMV K-Ho1	CPMV	CPSMV	RCMV	SqMV
K-G7		96.1	99.7	29.4	38.9	30.6	28.8
K-Ha1	97.4		96.4	29.7	39.2	31.9	29.7
K-Ho1	99.7	97.7		29.4	38.9	30.9	28.8
CPMV	43.2	43.2	43.6		29.5	38.5	29.6
CPSMV	51.6	52.0	52.0	42.0		31.1	30.2
RCMV	42.1	42.4	42.4	51.5	45.3		23.3
SqMV	40.2	40.5	40.5	42.7	40.3	36.2	

Helicase

	BPMV K-G7	BPMV K-Ha1	BPMV K-Ho1	CPMV	CPSMV	RCMV	SqMV
K-G7		96.0	98.8	42.3	52.1	44.0	42.3
K-Ha1	97.8		96.5	42.7	52.6	44.4	41.4
K-Ho1	99.0	98.2		42.4	52.4	44.2	41.8
CPMV	52.1	52.2	52.1		39.6	50.7	41.7
CPSMV	62.4	62.9	62.6	49.7		43.6	41.4
RCMV	54.3	54.3	54.3	61.1	54.6		43.9
SqMV	52.8	51.7	81.7	51.5	50.8	52.4	

VPg

Figure 2.7. Percentage deduced amino acid sequence identity/similarity of RNA1-encoded proteins between BPMV strains and other comoviruses. Deduced amino acid identity scores are shown above diagonal, and similarity scores are shown below diagonal. The similarity and identity scores were generated by the GAP program in the UWGCG package.

(Figure 2.7 continued)

	BPMV K-G7	BPMV K-Ha1	BPMV K-Ho1	CPMV	CPSMV	RCMV	SqMV
K-G7		100.0	100.0	64.3	67.9	67.9	71.4
K-Ha1	100.0		100.0	64.3	67.9	67.9	71.4
K-Ho1	100.0	100.0		64.3	67.9	67.9	71.4
CPMV	78.6	78.6	78.6		67.9	75.0	85.7
CPSMV	89.3	89.3	89.3	82.1		64.3	78.6
RCMV	78.6	78.6	78.6	78.6	54.8		75.0
SqMV	85.7	85.7	85.7	85.7	92.9	82.1	

Protease

	BPMV K-G7	BPMV K-Ha1	BPMV K-Ho1	CPMV	CPSMV	RCMV	SqMV
K-G7		98.6	98.6	51.0	45.7	44.7	46.4
K-Ha1	98.6		99.0	51.4	46.6	45.2	46.4
K-Ho1	98.6	99.0		51.0	46.6	45.7	46.4
CPMV	62.2	63.5	63.0		45.2	54.8	50.7
CPSMV	52.4	53.4	53.4	58.2		43.8	44.9
RCMV	55.8	56.3	56.7	63.0	54.8		50.5
SqMV	56.0	56.5	56.5	60.9	54.1	60.2	

RdRp

	BPMV K-G7	BPMV K-Ha1	BPMV K-Ho1	APMoV	CPMV	CPSMV	RCMV	SqMV
K-G7		94.8	96.3	47.1	56.8	55.0	57.0	52.6
K-Ha1	95.9		97.8	49.6	57.3	57.6	57.5	53.4
K-Ho1	97.2	98.3		49.3	57.6	56.5	57.5	53.2
APMoV	58.3	60.6	60.4		49.8	49.3	50.2	50.5
CPMV	64.5	66.0	66.1	59.3		55.6	61.6	57.0
CPSMV	64.0	66.0	64.9	57.9	63.3		52.0	54.9
RCMV	65.6	66.0	66.0	59.8	69.2	60.2		55.0
SqMV	62.6	62.7	62.9	60.5	64.9	63.9	63.3	

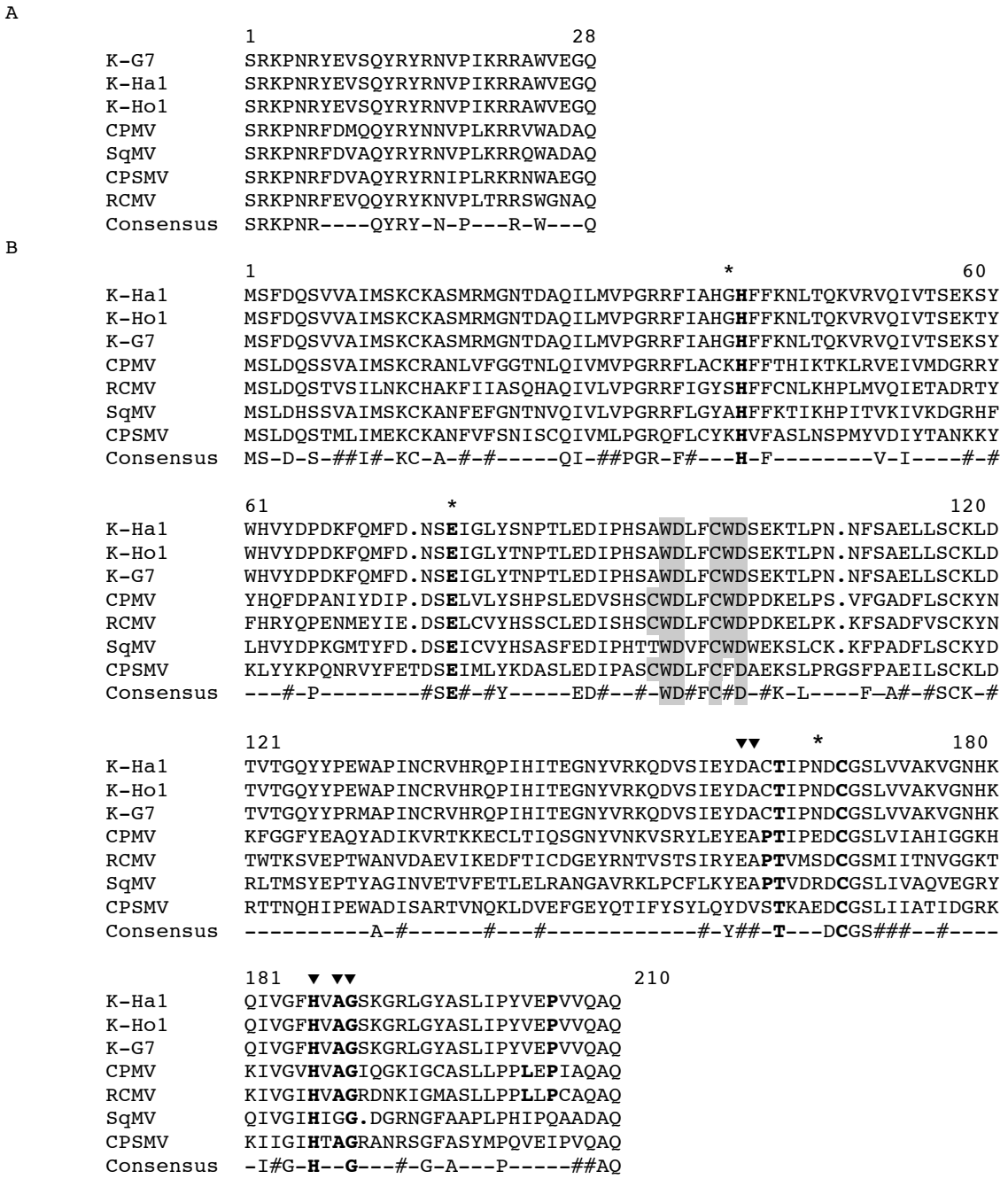

Figure 2.8. Sequence comparison of comovirus VPg and protease. (A) Multiple alignment of the deduced amino acid sequence of VPg from different BPMV strains and other members of the genus *Comovirus*. (B) Multiple alignment of the deduced amino acid sequence of protease of comoviruses. The conserved catalytic triad of comovirus proteases is indicated by bold letters and the positions are noted by asterisks (*). The residues of substrate binding pocket, according to Bazan and Fletteric model (1988), are indicated by arrowheads. The repeated "CWD" sequences are shaded. The multiple alignments were generated by the PRETTY program of the UWGCG package.

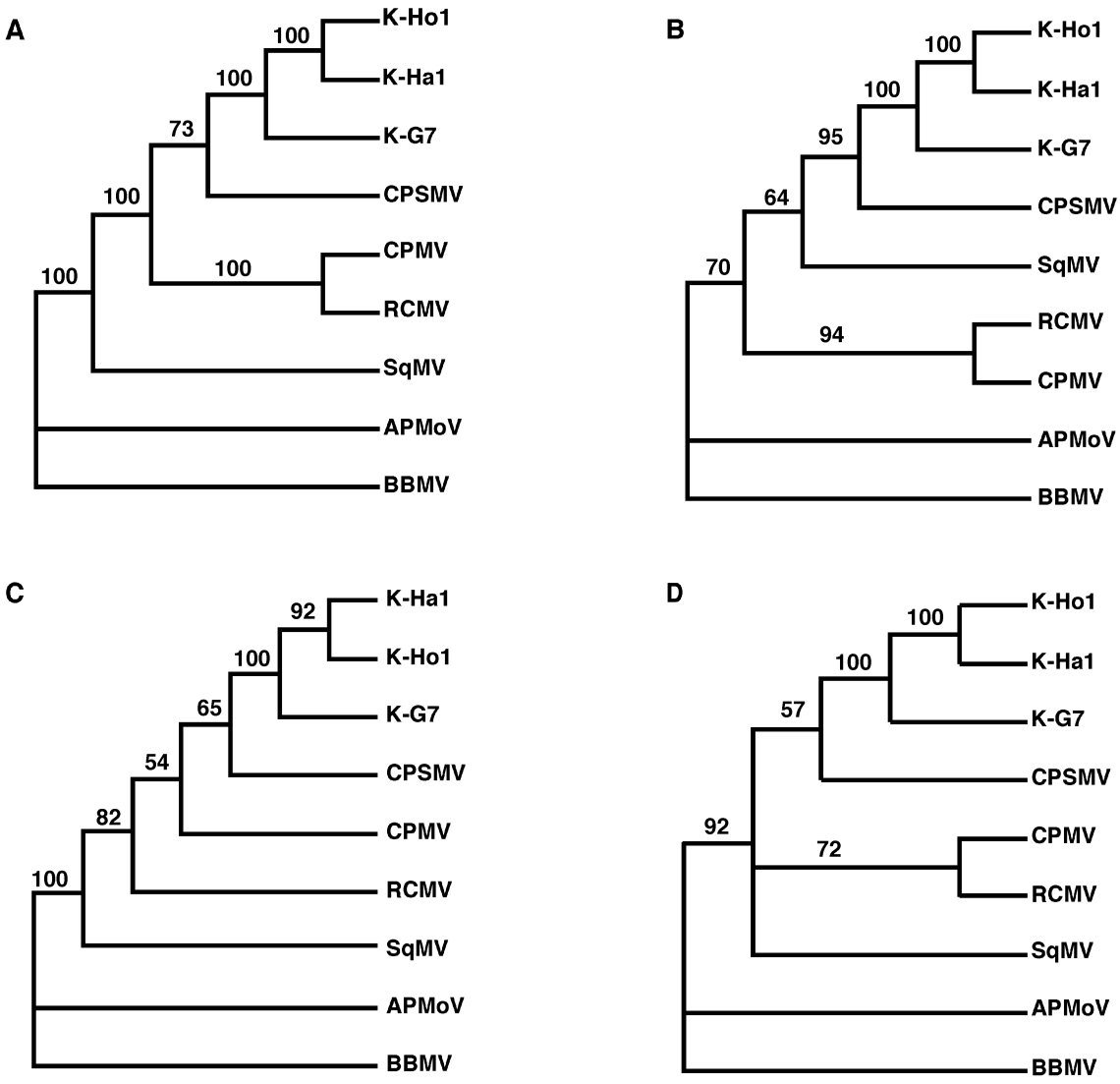

Figure 2.9. Phylogeny relationships of *comovirus* RNA2-encoded polyproteins and individual mature proteins. Amino acid sequences were aligned using the Clustal X Multiple Alignment Program and phylogenetic trees were constructed using the neighbor-joining method, as implemented in the PAUP program. The resulting consensus trees of 1000 bootstrap replicates are shown for: (A) RNA2 encode precursor polyproteins, (B) CR, (C) LCP and (D) SCP. The bootstrap values are indicated above each branch. *Broad bean wilt virus* (BBMV) RNA1 sequence was derived from Kobayashi et al. (2003).

RNA2 replication cofactor

	BPMV K-G7	BPMV K-Ha1	BPMV K-Ho1	APMoV	CPMV	CPSMV	RCMV	SqMV
K-G7		94.6	95.1	30.7	36.4	42.0	32.0	37.9
K-Ha1	96.4		99.3	30.4	36.8	42.3	32.2	37.7
K-Ho1	96.9	99.3		30.8	36.5	42.8	32.0	37.9
APMoV	42.2	40.9	41.9		28.6	30.7	30.7	32.7
CPMV	45.2	45.4	45.2	39.3		34.3	37.7	31.5
CPSMV	52.8	52.4	52.9	40.4	44.5		30.5	35.2
RCMV	43.5	43.5	43.5	41.6	47.6	42.8		32.9
SqMV	47.3	48.3	48.6	44.2	42.3	44.2	43.7	

Large Coat protein

	BPMV K-G7	BPMV K-Ha1	BPMV K-Ho1	APMoV	CPMV	CPSMV	RCMV	SqMV
K-G7		98.9	99.2	39.0	55.9	55.6	58.6	56.7
K-Ha1	98.9		99.7	38.7	55.9	55.6	58.6	56.4
K-Ho1	99.2	99.7		38.7	55.9	55.9	58.6	56.7
APMoV	47.4	47.1	47.1		36.2	30.3	39.7	40.2
CPMV	62.6	62.3	62.6	44.6		49.5	54.2	50.0
CPSMV	65.5	65.5	65.8	42.9	59.4		49.3	49.7
RCMV	64.5	64.5	64.5	46.4	62.7	56.8		55.4
SqMV	64.2	63.9	64.2	50.1	59.9	60.4	63.4	

Small coat protein

	BPMV K-G7	BPMV K-Ha1	BPMV K-Ho1	APMoV	CPMV	CPSMV	RCMV	SqMV
K-G7		95.0	95.0	34.6	37.1	41.0	46.1	43.4
K-Ha1	98.0		100.0	34.0	36.6	41.5	45.1	43.4
K-Ho1	98.0	100.0		34.0	36.6	41.54	45.1	43.4
APMoV	41.9	42.9	42.9		30.7	30.3	28.7	33.9
CPMV	44.3	44.3	44.3	38.6		33.5	49.0	39.6
CPSMV	48.2	49.2	49.2	36.2	41.2		42.5	35.6
RCMV	52.9	52.3	52.3	36.5	52.4	49.7		42.5
SqMV	52.8	53.9	53.9	42.1	44.0	45.0	48.1	

Figure 2.10. Percentage deduced amino acid sequence identity/similarity of RNA2-encoded proteins between BPMV strains and other comoviruses. Deduced amino acid identity scores are shown above diagonal, and similarity scores are shown below diagonal. The similarity and identity scores were generated by the GAP program in the UWGCG package.

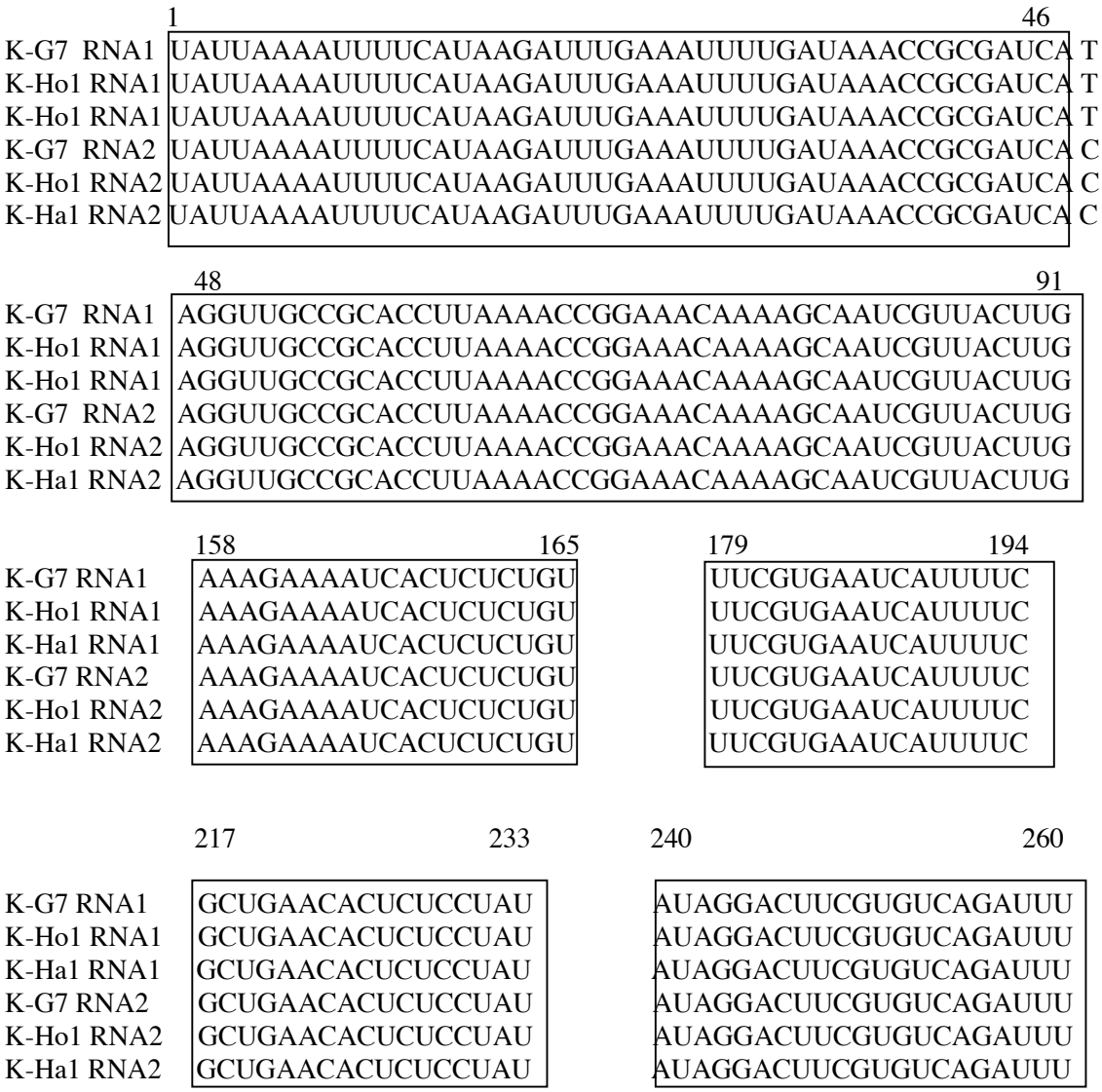

Figure 2.11. Multiple alignment of the RNA1 and RNA2 5'UTRs of different BPMV strains. The absolutely conserved sequences are boxed. The conserved regions are indicated above each box. The alignment was generated by PILEUP program of the UWGCG package.

Chapter III

Diversity among isolates of the comovirus *Bean pod mottle virus*

Introduction

Bean pod mottle virus (BPMV) is a member of the genus *Comovirus* in the family *Comoviridae* (Goldbach et al., 1995). Like other comoviruses, BPMV has a bipartite positive-strand RNA genome consisting of RNA-1 and RNA-2, which are separately encapsidated in isometric particles 28 nm in diameter. Purified BPMV preparations can be separated by density gradient centrifugation analysis into three viral sedimenting components designated top (T), middle (M) and bottom (B). The T component contains empty particles, whereas the M and B components contain single molecules of RNA-2 (approximately 3.6 kb) and RNA-1 (approximately 6.0 kb), respectively. The three components have identical protein composition, consisting of 60 copies each of a large (L) and small (S) coat protein of 41 kDa and 22 kDa, respectively. The S-coat protein occurs in two major size classes; the intact protein and a C-terminus-truncated version (Lomonossoff and Ghabrial, 2001).

BPMV genomic RNAs are polyadenylated and have a small basic protein, VPg, covalently linked to their 5' termini. The BPMV genome is expressed via the synthesis and subsequent cleavage of large polyprotein precursors (Goldbach et al., 1995; Lomonossoff and Ghabrial, 2001). The complete nucleotide sequences of the two genomic RNAs of BPMV strain KY-G7 have been reported (Di et al., 1999; MacFarlane et al., 1991). BPMV RNA-1 codes for five mature proteins required for replication (from 5' to 3', a protease cofactor [32K], a putative helicase [58K], a viral genome-linked protein [VPg], a protease [24K] and a putative RNA-dependent RNA polymerase, RdRp [87K]), whereas RNA-2 codes for a putative cell-to-cell movement protein and the two coat proteins (Di et al., 1999; MacFarlane et al., 1991).

BPMV is widespread in the major soybean-growing areas in many of the southern and southeastern United States (Ghabrial et al., 1990; Hartman et al., 1999; Ross and Butler, 1985; Skotland, 1958). A recent severe outbreak in BPMV incidence in the North

Central and Northern Great Plains states is currently the cause of serious concerns to soybean growers and to the soybean industry in this region (John Hill, personal communication; Ghabrial, unpublished). BPMV is efficiently transmitted in nature, within and between soybean fields, by several species of leaf-feeding beetles in the family Chrysomelidae (mainly the genera *Ceratoma* and *Diabrotica*). Beetle vector species also occur in the families Coccinellidae, Curulionidae, and Meloidae (Gergerich and Scott, 1996; Hartman et al., 1999; Ross, 1963). The deleterious effects of BPMV infection are not limited to seed yield but extend to seed quality since BPMV is known to predispose soybeans to *Phomopsis* spp. seed infection (Stuckey et al., 1982), a major cause of poor seed quality in soybean (Schmitthenner and Kmetz, 1980). Furthermore, BPMV interacts synergistically with the potyvirus *Soybean mosaic virus* (SMV) with drastic reduction of yield and seed quality (Anjos et al., 1992; Calvert and Ghabrial, 1983; Ross, 1968).

Disease management through genetic resistance is not possible at present because no soybean cultivars with resistance to BPMV are commercially available. A limited number of transgenic soybean lines expressing BPMV capsid polyprotein have been produced and were shown to confer protection against BPMV infection (Di et al, 1996; Reddy et al., 2001). Such transgenic resistance could be incorporated into commercial varieties. Because of the recent BPMV outbreaks, a concerted effort is currently underway to screen available soybean germplasm for resistance/tolerance to BPMV infection (Gu et al., 2002). No information, however, is available on strain diversity among BPMV isolates. Knowledge of the genetic diversity among BPMV isolates is necessary to ensure that selected or newly developed soybean germplasms may offer broad protection against the full range of BPMV strains found in nature. In this communication, we report the occurrence of at least two distinct subgroups of BPMV strains that can be clearly distinguished based on nucleic acid hybridization analysis. Furthermore, we present evidence for the occurrence in nature of reassortants between the two strain subgroups. Additionally, we developed an RT-PCR protocol based on the sequence of a highly conserved region in the capsid polyprotein coding sequence that provides efficient and highly sensitive detection of all BPMV isolates tested, regardless of their strain classification.

Materials and Methods

Virus isolates

A list of the BPMV isolates used is shown in Table 3.1. In most cases, the isolates are designated by the county and the state from which they were originally collected. The various BPMV isolates were propagated in the soybean cultivars York or Essex, and infected tissues were used for virion purification as previously described (Ghabrial et al., 1977). The complete nucleotide sequences of the genomic RNAs from isolates K-G7 (Di et al., 1999; MacFarlane et al., 1991), K-Ha1 and K-Ho1 (Chapter II) have been determined and the sequences have been deposited in the GenBank (see Table 3.1 for the GenBank accession numbers). Where appropriate, the well-characterized isolates, K-G7, K-Ha1 and K-Ho1, will be referred to as strains.

Nucleic acid hybridization analysis

Viral RNAs were extracted from purified virions by the SDS-phenol method of Peden and Symons (1973), and assayed by northern hybridization analysis using a cloned cDNA probe to RNA-1 from strain K-G7. As the ratio of RNA-1 to RNA-2 may vary among virus isolates and to ensure that equivalent amounts of RNA-1 are subjected to northern hybridization analysis, the amounts of total viral RNA loaded per lane were adjusted based on image analysis of scanned gels. For this purpose, images of ethidium bromide-stained gels were generated using an Alpha Innotech digital imaging system (Alpha Innotech Corporation, San Leandro, CA) and the resulting images were analyzed by the AlphaEase program (Alpha Innotech Corporation, San Leandro, CA) using the 1D-multi line densitometry tool. The integrated area under the RNA-1 peak was then determined for all the RNA preparations to be tested. The volumes of the RNA samples were adjusted accordingly so that equivalent amounts of RNA-1 were loaded for each BPMV isolate, as demonstrated in Figure 3.1. For northern hybridization analysis, viral RNA samples from the same preparations used for image analysis and in similar amounts to those shown in Figure 3.1 (approximately 1.0 mg) were used. The viral RNAs were separated by electrophoresis on 1.5% formaldehyde agarose gels (Sambrook and Russell,

2001), and transferred onto Hybond-N^+ membranes (Amersham, Piscataway, NJ), according to the manufacturer's instructions. The membranes were then prehybridized for 12 h at 42^0 C in a solution containing 5X SSC (1X SSC: 150 mM NaCl, 15 mM sodium citrate), 50% formamide, 0.5% SDS and 100 mg/ml salmon sperm DNA. Hybridization to the radiolabeled cloned cDNA probes was in the same buffer for 16 h at 42^0C. The membranes were then washed two times for 10 min each at room temperature in 2X SSC containing 0.1% SDS, followed by washing two times with 1X SSC/0.1% SDS at room temperature and two times in 0.1X SSC containing 0.1% SDS at 65^0 C for 30 min. The membranes were air-dried and exposed to Bio-Max film (Eastman Kodak, Rochester, NY) for 2-12 h. For slot-blot hybridization analysis, RNA was extracted from purified virions and, after concentration by ethanol precipitation, suspended in 2X SSC, and spotted onto Hybond-N^+ membranes using a "Minifold II" filtration apparatus (Schleicher & Schuell, Keene, NH). The RNA samples (200 ng, unless otherwise specified) were fixed onto membranes according to the manufacturer's instructions. Procedures and conditions for prehybridization, hybridization, and membrane washings were as described for northern hybridization analysis. The probes were prepared by nick-translation of full-length cDNA clones of RNA-1 and RNA-2 from isolates K-Ha1, K-Ho1 and K-G7 (except that a partial clone containing nucleotides 1511 to 3117 was used for K-G7 RNA-1). The membranes were then air-dried and exposed to a phosphorimager screen (Molecular Dynamics, Sunnyvale, CA) for 2-5 h. The images were visualized by a PhosphorImager 445 SI system and analyzed with the ImageQuant 4.1 program (Molecular Dynamics, Sunnyvale, CA).

Western blot analysis

Following SDS-PAGE on 12% polyacrylamide gels (Laemmli, 1970), the proteins were transferred to an Immobilon-P transfer membrane (Millipore), using a mini Trans-Blot electrophoretic transfer apparatus (Bio-Rad Laboratories, Hercules, CA) at 70 V for 1 h. The membrane was then incubated overnight at 4^o C in TBS (20 mM Tris-HCl buffer containing 150 mM NaCl, pH 7.4) containing 5% non-fat milk and incubated at room temperature for 2 h with an antiserum to BPMV virions (Ghabrial and Schultz, 1983) diluted 1: 2,000 in 1X TBS/5% non-fat milk. The membrane was washed twice in

water and three times in 1X TBS, then reacted with the secondary antibody, goat anti-rabbit IgG conjugated with alkaline phosphatase (Sigma-Aldrich, St. Louis, MO; diluted 1:2,000 in 1X TBS containing 5% non-fat milk) for 1 h at room temperature. Following washing as described before, the bound antibody was detected using 5-bromo-4-chloro-3-indolyl phosphate *p*-toluidine salt and *p*-nitro blue tetrazolium (Promega, Madison, WI) as substrates.

Field studies

A field plot experiment was carried out in 1999 to study the effects of infection with each of five different isolates of BPMV (K-G7, K-Ha1, K-Ho1, K-D1 and K-U1) on soybean yield. The selected BPMV isolates differ in the severity of the symptoms they induce in soybean from mild to severe (Table 3.2). Three Essex isolines (Essex, Essex-*Rsv*1, and Essex-Rsv4; provided by Glenn Buss, Virginia Polytechnic Institute & State University), which differ in the presence of resistance genes to SMV, were inoculated with the five BPMV isolates. Inoculum for each isolate was prepared by extracting infected leaf tissue with 0.05 M potassium phosphate buffer, pH 7.0 (1:10 wt/vol). Young leaves from Essex soybean seedlings infected with the individual BPMV isolates (2 weeks postinoculation) were used as sources of inoculum. Carborundum (600-mesh) was added to the inoculum prior to application.

The experiment was conducted as a completely randomized design. Each virus x isoline combination was replicated three times. Each isoline was included as a noninoculated control three times. The experiment was planted May 24, and all plants in a plot were rub-inoculated with sap from infected soybeans, as described before, on June 17 at growth stage V2/V3 (Fehr et al., 1971). Each plot consisted of six plants spaced 2 inches apart in a hill with hills spaced 30 inches apart. The plots were sprayed every other week with Malathion in an attempt to reduce the spread of BPMV by bean leaf beetles to noninoculated plants in control plots. Virus-like symptoms were first detected on noninoculated control plants on July 29. Because these symptoms appeared on all three isolines we judged them not to be due to be SMV. The experiment was irrigated five times (approximately every two weeks) from mid July to early September.

Yield (g/plot) was measured on each 6-plant hill plot. The yield data were analyzed by analysis of variance and linear contrasts using statistical analysis software (Windows version 6.12, SAS Institute, Cary, NC). Linear contrasts were used to compare inoculated versus noninoculated plots, strain subgroups I and II versus the reassortants, and strain subgroup I versus II.

Results and discussion

Symptomatology

Twenty-one different BPMV isolates were tested for symptom production on soybean cv. Essex (Table 3.2). The resulting symptoms were scored as follows: severe (severe stunting, yellow mottling, leaf distortion and extensive blistering), intermediate (yellow/green mottling with some stunting and blistering) and mild (mild green mottling and little or no stunting). All soybean cultivars tested, including Essex, Williams 82, Fayette, Jack, Davis, Clark, York, Dare and Hutchinson, were susceptible to BPMV infection regardless of the viral isolate used.

All BPMV isolates induced necrotic lesions on *Phaseolus vulgaris* (L.) cv. Pinto and chlorotic lesions with necrotic centers on *P. vulgaris* cv. Bountiful. The response of the bean cv. Black Valentine to BPMV varied with the source of seeds. Seedlings germinated from seeds obtained from three commercial sources (Vermont Bean Seed Co.,Vaucluse, SC; Sauk River Seed, Albany, MN; and Seeds For The South, Graniteville, SC), showed chlorotic lesions on inoculated leaves without systemic infection. However, Black Valentine bean seedlings generated from seeds supplied by Rose Gergerich (University of Arkansas) produced systemic symptoms similar in severity to those induced by the different isolates on soybean (Table 3.2). Skotland (1958) also reported that Black Valentine seed obtained from four different seed companies yielded plants that were only susceptible to local lesion infection. The original report of Zaumeyer and Thomas (1948) on BPMV, however, listed Black Valentine bean as a systemic host. All BPMV isolates tested, except for K-Ho1, induced symptomless infections in cowpea. The K-Ho1 isolate, the most severe on soybeans, produced mild mosaic symptoms on cowpea. Although all BPMV isolates from Kentucky produced chlorotic local lesions on

Chenopodium quinoa (L), the majority of those from other states induced ringspots and line patterns on inoculated leaves.

Identification of two distinct subgroups of BPMV strains

Northern hybridization analysis with a cloned cDNA probe to RNA-1 from strain K-G7 indicated that viral RNA-1 from representative BPMV isolates from four states can be separated into at least two distinct hybridization groups (Figure 3.2). As the purpose of the northern hybridization analysis was to ascertain whether RNA from a particular isolate hybridized with the probe, it was necessary in some cases to overexpose the membranes to the x-ray film. Consequently, smears were evident in some of the lanes that hybridized strongly with the probe (Figure 3.2). A total of six isolates out of 19 tested hybridized strongly with the K-G7 RNA-1 probe whereas no or very weak hybridization signals were observed with the remaining 13 isolates. Of the six isolates whose RNA-1 hybridized strongly with the K-G7 RNA-1 probe, five were collected from different locations in Kentucky. In contrast, RNA-1 from six of seven isolates tested from other states belonged to a different hybridization subgroup.

The results of northern hybridization analysis (Figure 3.2) were reproducible when the same BPMV isolates were subjected to slot-blot hybridization analysis using cloned cDNA probes to RNA-1 from either strain K-G7 (data not shown) or strain K-Ho1 (Figure 3.3A; same hybridization group as K-G7). Thus, RNA-1 from six isolates hybridized strongly with K-G-7 and K-Ho1 RNA-1 probes whereas the remaining 10 isolates showed no hybridization signals after a short exposure time (2 h; Figure 3.3A). With a longer exposure of the same membrane to the phosphorimager screen (4 h), comparable weak hybridization signals were detected for all 10 isolates, thus verifying that equivalent amounts of RNA samples were applied to the membrane (Figure 3B). The high stringency conditions used in our northern and slot blot hybridization experiments require that sequences have at least 90% identity to the probe in order to generate strong hybridization signals (Memelink et al., 1994). The production of weak or no hybridization signals indicates that the percent identity between RNA-1 from each of the 10 isolates and the RNA-1 probes are less than 90%. This conclusion is supported by

the nucleotide sequencing data of representative isolates of the two hybridization groups (see below).

The biological, molecular and structural properties of strain K-G7 have been well characterized and it is regarded as the type strain of BPMV (Chen et al., 1989; Di et al., 1999; Ghabrial et al., 1977; Ghabrial and Schultz, 1983; MacFarlane et al., 1991). As the type strain with known sequence and to facilitate the classification of our isolates based on nucleic acid hybridization analysis, we designated BPMV K-G7 as the prototype of hybridization subgroup I. We selected strain K-Ha1, which did not hybridize with the K-G7 RNA-1 probe, as a representative of a second hybridization subgroup (designated subgroup II). The complete nucleotide sequences of both genomic RNAs of isolate K-Ha1 were recently determined (Chapter II) and their sequences deposited in the GenBank (see Table 3.1 for GenBank accession numbers). The values for percentage nucleotide sequence identity between strains K-G7 and K-Ha1, as determined by sequence comparison using the GAP program, are 85.5% and 86.9%, respectively, for RNA-1 and RNA-2.

Surprisingly, the classification of BPMV isolates into two distinct subgroups based on hybridization with the RNA-1 probes (Figures. 3.2 and 3.3A) was not so apparent when a cDNA clone to K-G7 RNA-2 was used as a probe. Only two isolates (V-S98-1 and K-G7) of the six tentatively classified as belonging to subgroup I (based on hybridization with the RNA-1 probes) hybridized strongly with the K-G7 RNA-2 probe (Figure 3.4). Isolates K-C2, K-Ho1, K-U1, and K-Fu1 showed very weak or no hybridization signals and were comparable to signals from the subgroup II isolates.

The results of the reciprocal slot-blot hybridization assays using cloned cDNA probes to RNA-1 and RNA-2 from the subgroup II strain K-Ha1 are shown in Figure 3.5. Ten of the isolates tested hybridized strongly to the RNA-1 probe from the subgroup II strain, whereas the remaining six isolates, previously classified as belonging to subgroup I, showed little or no hybridization. In contrast, the 14 BPMV isolates that did not react with the subgroup I RNA-2 probe hybridized strongly with subgroup II probe (compare Figure 3.5B with Figure 3.4). The results thus indicate that the genomes of only two isolates out of the 16 tested hybridized solely to subgroup I probes, whereas the genomes of 10 isolates hybridized solely with subgroup II probes. The remaining four isolates (K-

C2, K-Ho1, K-Fu1 and K-U1) displayed hybridization patterns indicative of reassortants between the two subgroups with RNA-1 derived from a subgroup I strain and RNA-2 derived from a subgroup II strain. The reassortant status of strain K-Ho1 was confirmed by nucleotide sequencing of full-length cDNA clones of its genomic RNAs (sequences deposited in the GenBank, Table 3.1; Chapter II).

In a study on the diversity of the comovirus *Squash mosaic virus* (SqMV), Haudenshield and Palukaitis (1998), using northern hybridization analysis with a cDNA probe to RNA-2, classified the SqMV isolates into two hybridization groups. Like BPMV, the SqMV isolates were disproportionately distributed between the two groups. Of the five SqMV isolates tested, one belonged to hybridization group 1 whereas the remaining four were placed in hybridization group 2 (Haudenshield and Palukaitis, 1998). Because probes to SqMV RNA-1 probes were not used in that study, it cannot be ruled out that some of the isolates are reassortants. It is clear from the present BPMV study that comovirus reassortments occur in nature, and that the use of cloned probes to both genomic RNAs would be required to unravel the extent of diversity.

It is of interest that the recent increase in BPMV incidence has been concomitant with a large increase in the beetle vector populations as well as with an augmentation in disease symptom severity. Molecular characterization of such severe BPMV isolates revealed that they are reassortants between two distinct subgroups of strains (this study). The BPMV isolates collected in earlier virus surveys in the 1970s and 1980s (Ghabrial et al., 1977; Ghabrial et al., 1990) induced only mild or moderate symptoms on soybeans (Table 3.2; Ghabrial, unpublished). The relatively mild BPMV strain K-G7, the prototype of subgroup I, was isolated in 1975 (Ghabrial et al., 1990), and the mild Isolates included in the present study were shown to belong to subgroup II. The severe reassortant strains thus appear to be of recent occurrence and to be associated with the recent large increases in the beetle vector populations and in BPMV incidence. Although these severe natural reassortants appear to be stable in the short term, their long survival in nature remains to be seen.

Field plot experiments

No significant differences (F = 0.8) in the yield of the three Essex isolines were detected (Essex, 104 g/hill; Essex Rsv_1, 108 g/hill; Essex Rsv_4, 116 g/hill). The

interaction of isoline x virus isolate was also nonsignificant (F = 0.64). Thus, the SMV resistance alleles Rsv_1 and Rsv_4 did not provide resistance to BPMV. Although all inoculated plants showed virus symptoms, the various BPMV isolates differed in symptom severity and the magnitude of yield loss they caused (Table 3.3). The greatest yield loss was due to infection with the reassortant strain K-Ho1, which induced the most severe symptoms. The plants in non-inoculated plots (133 g/plot) yielded significantly more than the plants in inoculated plots (103 g/plot; F = 9.05, $P \leq 0.01$). Yield reductions caused by the different strain subgroups (I, II and the reassortants I/II) were not significantly different (strains vs. reassortants; F = 3.10, strain subgroup I vs. II; F = 2.59). The number of isolates tested, however, was limited and only one isolate from strain subgroup I was included in the study. Large differences in yield reductions were observed within a strain subgroup; K-D1 reduced yield significantly less than strain K-Ha1, both in strain subgroup II. This is of interest because K-D1 was considered a mild isolate under greenhouse conditions, but induced relatively more severe symptoms under field conditions. All other isolates induced symptoms similar to those described under greenhouse conditions.

In summary, the data presented in this study represents the first report on strain diversity among BPMV isolates. Nucleic acid hybridization analysis and nucleotide sequencing data clearly revealed the occurrence in nature of two distinct subgroups of BPMV strains as well as reassortants between the two subgroups. The nucleotide and deduced amino acid sequence analyses of the entire genomes of three BPMV strains indicated that the coding sequence of the capsid polyprotein is more highly conserved among BPMV isolates than the remainder of the genome. This finding was supported by the results from our RT-PCR and western blot analyses. In a field plot experiment, different BPMV isolates induced symptoms on soybeans that varied in severity and in the incurred yield losses. There was, however, no relationship between the magnitude of yield loss and strain classification. Knowledge of strain diversity among BPMV isolates is necessary to ensure that selected or newly developed soybean germplasms may offer broad protection against the full range of BPMV strains found in nature. This is of special importance in view of the recent increase of BPMV incidence in the major soybean-growing regions in the United States.

Table 3.1. Sources of BPMV isolates used and year of collection

Virus isolate (GenBank accession numbers)	Location	Year collected
A-S98-50M	Kaiser, AR	1998
I-JH1	Ames, IA	2000
K-C1	Caldwell Co., KY	1997
K-C2	Caldwell Co., KY	1997
K-D1	Daviess Co., KY	1998
K-Fa1	Fayette Co., KY	1996
K-Fu1	Fulton Co., KY	1987
K-G7 (M62738, U70866)	Graves Co., KY	1975
K-Ha1 (AF394606, AF394607)	Hancock Co., KY	1998
K-He1	Henderson Co., KY	1998
K-He2	Henderson Co., KY	1998
K-He3	Henderson, Co., KY	1998
K-Ho1 (AF394608, AF3946080	Hopkins Co., KY	1998
K-U1	Union Co., KY	1998
M-D3	Delta Station, MS	1997
M-G1	Greenville, MS	1997
V-W1	Warsaw, VA	1997
V-W2	Warsaw, VA	1997
V-S98-1	Mount Holly, VA	1998
V-S98-15	Warsaw, VA	1998

Table 3.2. Host reactions to representative BPMV isolates collected from four states.

Isolate	Strain subgroup[a]	Soybean 'Essex'	Cowpea 'Blackeye'	*Chenopodim. quinoa*
A-S98-50M	II	M[b]	SL (+)	CL (+)
K-C1	II	I	SL (+)	CL
K-C2	I/II	I	SL (+)	CL
K-D1	II	M	SL (+)	CL
K-Fa1	II	M	SL (+)	CL
K-Fu1	I/II	S	SL (+)	CL
K-G7	I	I	SL (+)	CL
K-Ha1	II	M	SL (+)	CL
K-He1	II	M	SL (+)	CL
K-He2	II	M	SL (+)	CL
K-He3	II	I	SL (+)	CL
K-Ho1	I/II	S	Mo	CL
K-U1	I/II	I	SL (+)	CL
M-D3	II	M	SL (+)	RS, LP
M-G1	II	M	SL (+)	RS, LP
V-S98-1	I	M	SL (+)	CL (+)
V-S98-15	II	M	SL (+)	RS, LP
V-W1	II	M	SL (+)	RS
V-W2	II	M	SL	RS, LP

[a] Strain subgroup affiliation (I, II or an assortment I/II) was determined by nucleic acid hybridization analysis (see Figs.2-5).

[b] Abbreviations: M=mild (green mottling with little or no stunting/blistering); I=intermediate (green/yellow mottling, some stunting, leaf distortion and blistering); S=severe (yellow mottling, severe stunting and blistering and necrosis on inoculated leaves); SL=symptomless; (+)= non-inoculated upper leaves are ELISA positive; Mo= mosaic; CL=chlorotic lesions; RS= ringspots; LP= line patterns

Table 3.3. Comparative yield losses in soybean (Essex and two Essex SMV-resistant isolines) inoculated with five different isolates of BPMV.

BPMV isolate	Strain subgroup[a]	Seed weight (g/plot)[b]	% of noninoculated control
Noninoculated		133 a	
K-G7	I	96 b	72%
K-Ha1	II	134 a	100%
K-D1	II	98 b	74%
K-Ho1	I/II	80 b	60%
K-U1	I/II	105 b	79%

[a] Strain subgroup affiliation (I, II or an assortment I/II) was determined by nucleic acid hybridization analysis (see Figures. 3.2-3.5).

[b] Means followed by the same letter are not significantly different at P=0.05.

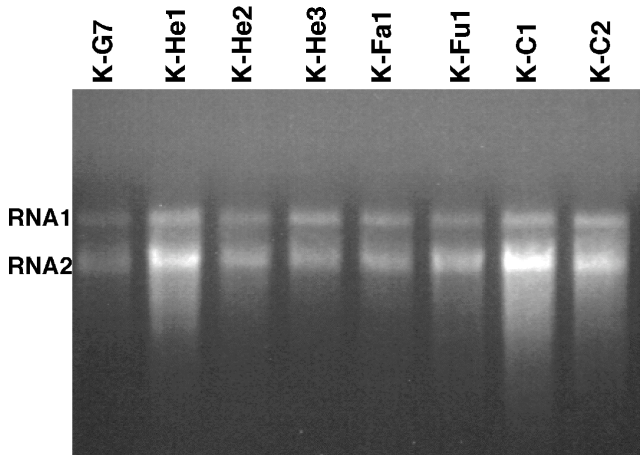

Figure 3.1. Agarose gel electrophoresis of purified *Bean pod mottle virus* (BPMV) RNA. RNA samples from representative BPMV isolates were electrphoresed on 1% agarose gels and stained with ethidium bromide. The amounts of total viral RNA loaded per lane were adjusted based on prior image analysis of scanned gels of the same RNA preparations so that all lanes contain equivalent amounts of RNA-1 (see Materials and Methods for details). The designations of the BPMV isolates from which the RNAs were isolated are marked at the top of each lane.

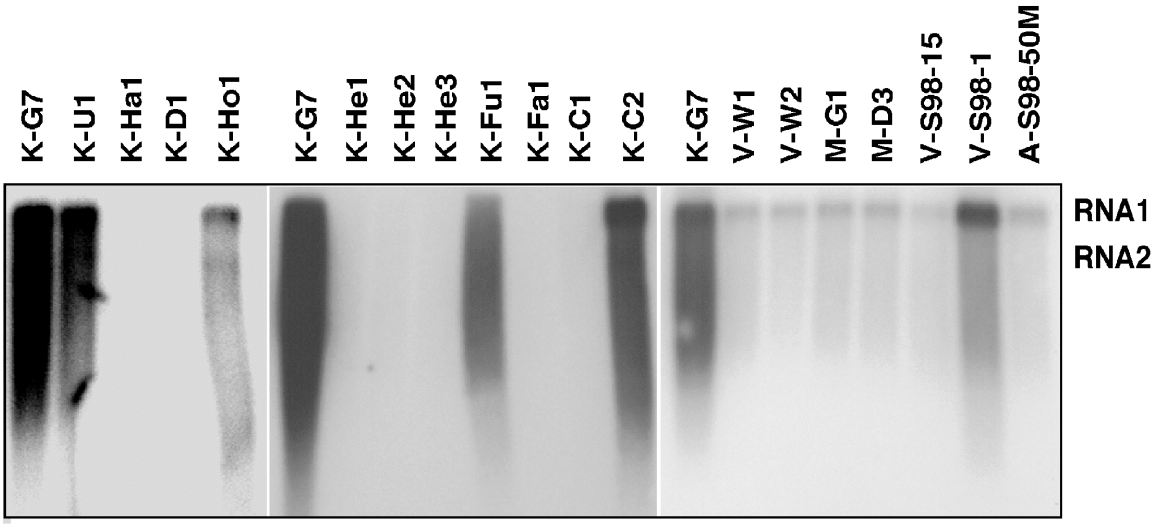

Figure 3.2. Northern hybridization analysis of RNA isolated from purified virions from 19 *Bean pod mottle virus* (BPMV) isolates. Viral RNA samples from the same preparations and in similar amounts to those shown in Fig. 1 (approximately 1.0 mg) were separated by electrophoresis on 1.5% formaldehyde agarose gels, and transferred onto Hybond-N^+ membranes (Amersham, Piscataway, NJ). The RNA blots were hybridized with a cloned RNA-1 cDNA probe to K-G7.

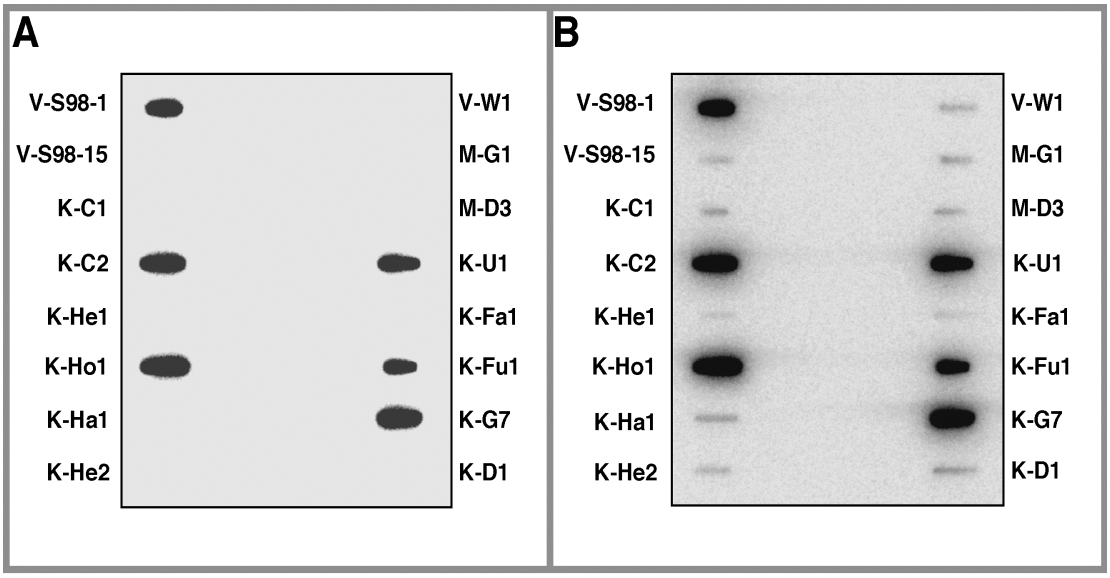

Figure 3.3. Slot blot hybridization analysis of RNA isolated from purified virions of 16 *Bean pod mottle virus* (BPMV) isolates. RNA samples (200 ng) were blotted onto a Hybond-N^+ membranes (Amersham, Piscataway, NJ) and hybridized with a cloned cDNA probe to RNA-1 from isolate K-Ho1. A, Membrane was exposed for 2h. B, Same membrane was exposed for 4 h demonstrating that all samples were applied to the membrane.

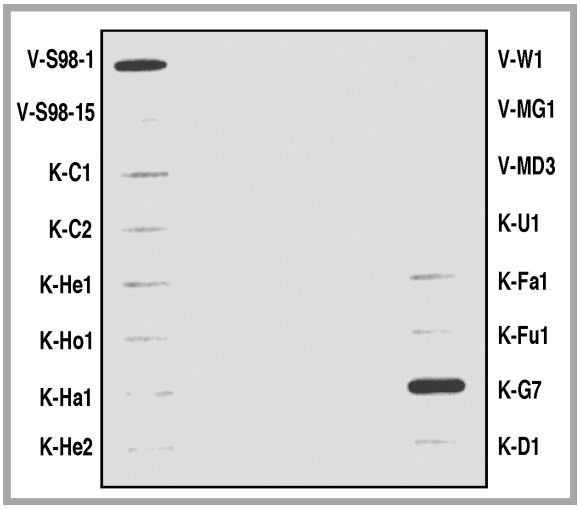

Figure 3.4. Slot blot hybridization analysis of RNA isolated from purified virions of 16 *Bean pod mottle virus* (BPMV) isolates. RNA samples were blotted (200 ng) onto a Hybond-N^+ membrane (Amersham, Piscataway, NJ) membrane and hybridized with a cloned cDNA probe to RNA-2 from isolate K-G7.

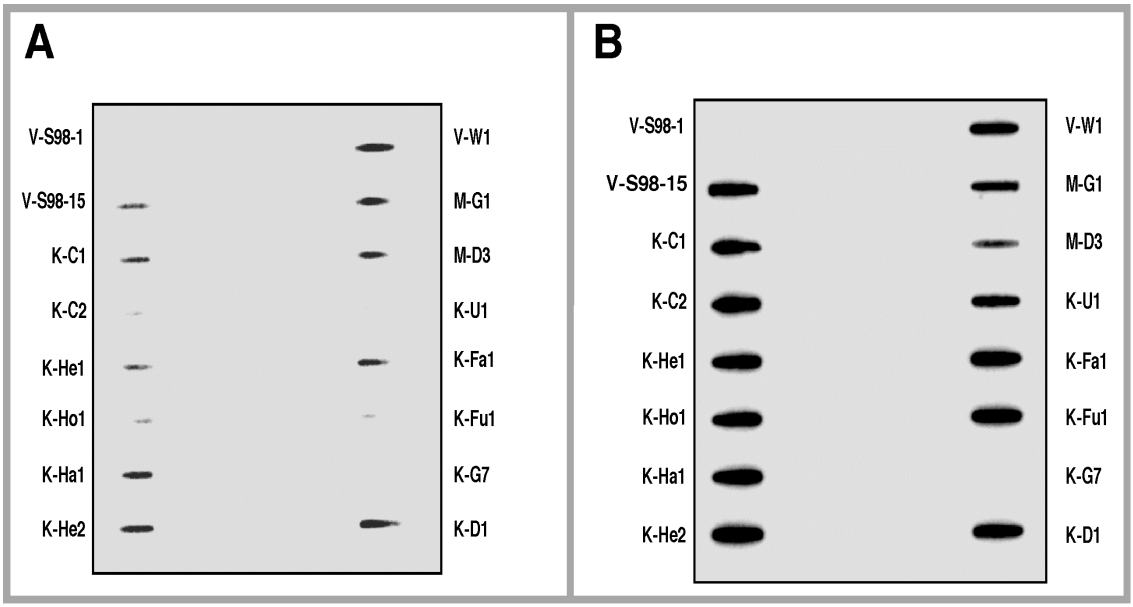

Figure 3.5. Slot blot hybridization analysis of virion RNA isolated from 16 *Bean pod mottle virus* (BPMV) isolates. RNA samples (50 and 200 ng, respectively, for A and B) were blotted onto a Hybond-N^+ membrane (Amersham, Piscataway, NJ) and hybridized with cloned cDNA probes to: (A) RNA-1, and (B) RNA-2 from BPMV strain K-Ha1, respectively.

Chapter IV

Cross-protection among isolates of BPMV

Introduction

Cross-protection refers to the phenomenon in which systemic infection by one virus protects plants from subsequent infection by another related virus. It was first described in 1929 by McKinney who demonstrated that prior inoculation of tobacco plants with a light green mosaic strain of *Tobacco mosaic virus* (TMV) protected against infection with a yellow mosaic strain. The phenomenon of cross-protection was subsequently reported among different strains of several other viruses (Salaman, 1933; Posnette and McA, 1955; Simmonds, J. H., 1959; Yeh *et al*., 1987; Wen *et al*., 1991; Fraser, 1998). Cross-protection was successfully utilized to control several viral diseases under field conditions, such as *Citrus tristeza virus* (CTV; Adams 1988), *Papaya ringspot virus* (PRSV; Wang et al., 1987), *Tomato mosaic virus* (ToMV; Ahoonmanesh and Shalla, 1981) and *Zucchini yellow mosaic virus* (ZYMV; Walkey, 1992). In addition, cross-protection was exploited as an indictor of the relatedness among viral strains (Matthews, 1991).

Numerous mechanisms have been proposed in order to explain this phenomenon. One mechanism suggests that the multiplication sites of one virus are limited and specific for the this virus. Infection by one virus isolate may occupy all the sites and therefore the challenging viral isolate cannot multiply. On the other hand, multiplication of an unrelated virus is not affected (Bawden and Kassanis, 1945). Another mechanism is based on shortage in metabolites as infection by the protecting strain may deplete some essential metabolites required by the challenging strain, thereby blocking its multiplication (Matthews, 1991). The RNA-RNA interaction mechanism proposes that the minus strand copies of the protecting RNA virus anneal with the RNA from the challenging virus and prevents its replication (Zaitlin, 1976; Palukaitis and Zaitlin, 1984). However, protection afforded by coat protein-mediated transgenic resistance to some viruses is inconsistent with this mechanism (Powell-Abel *et al*., 1986; Tumer *et al*.,

1987). No convincing evidence is presently available in support of any of these proposed mechanisms.

Based on cross-protection among isolates of TMV, the reencapsidation of the challenging viral RNA mechanism was proposed (de Zoeten and Fulton, 1975; Sherwood and Fulton, 1982; Lu *et al.*, 1998). This mechanism is based on the prediction that the coat protein from the protecting strain will recoat the 5' terminal end of challenging virus RNA, which is exposed upon uncoating. Several lines of evidences support this mechanism. For instance, plants infected by one TMV isolate could not protect it from infection by viral RNA of another isolate (Sherwood and Fulton, 1982; Culver, 1996). Sherwood (1987) showed that protection by a CP deficient mutant was overcome by another TMV isolate. Lu et al. (1998) demonstrated that the ability of the protecting virus CP to bind the challenging virus RNA was necessary for cross-protection. Additionally, the CP subunits should be able to form virion-like helical aggregates in order to prevent infection from a challenging viral strain. However, results from several independent studies were inconsistent with this mechanism. For example, results from some studies indicated that CP-less strains were able to protect against other TMV strains in tobacco plants (Gerber and Sarkar, 1989). Plants inoculated with a PVX vector expressing segments of the TMV POL domain conferred high protection levels against infection by TMV (Goregaoker *et al.*, 2000). Cross-protection among strains of *Potato virus A* (PVA) was efficient against challenge-inoculation with viral RNA (Valkonen *et al.*, 2002). In addition, the helper component proteinase (HCpro) coding region was also critical to cross-protection.

Posttranscriptional gene silencing (PTGS) is a general host defense phenomenon observed in plant and animal kingdoms (Cogoni and Macino, 2000). PTGS can be triggered by virus infection (Ratcliff *et al.*, 1997; Voinnet, 2001). Ratcliff et al.(1999) demonstrated that PTGS is involved in cross-protection using transient gene expression experiments, in which *N. benthamiana* plant was previously inoculated with *Tobacco rattle virus* (TRV) vector expressing green fluorescent protein (GFP) and challenge-inoculated with Potato virus X (PVX) vector expressing truncated GFP.

Bean pod mottle virus (BPMV) is an important pathogen of soybean in the United States. Yield loss of 10-40% has been reported (Giesler et al., 2002). BPMV can be

efficiently transmitted by several leaf-feeding beetles. In the major soybean-growing regions, the bean leaf beetle (BLB; *Cerotoma trifurcata* Förster) is considered the major vector of BPMV (Giesler et al., 2002). Symptoms induced by different isolates vary from mild, moderate to very severe (Chapter III).

In this chapter, cross-protection among isolates of BPMV was investigated. Results indicated that cross-protection was evident regardless of whether virions or BPMV-RNA were used as inocula. Prior infection of soybean plants with mild isolates of BPMV conferred efficient protection against infection with a severe isolate, regardless of the method of virus inoculation, rub-inoculation or by BLB. Protection was complete and durable. The mechanism of cross-protection among BPMV isolates is discussed.

Materials and Methods

Virus strains

Four isolates (K-Ha1, K-Ho1, K-G7 and K-D1) were used in this part of the study. The symptoms induced by these four isolates were described in Table 1, Chapter III. The complete nucleotide and deduced amino acid sequences of the genomic RNAs of strains K-Ha1 and K-Ho1 were determined and shown in Figures 2.1- 2.4. The sequences of the genomic RNAs of strain K-G7 have been published previously (MacFarlane *et al*., 1991; Di *et al*., 1999). The soybean cultivars Essex, York, Clark and Calhoun were maintained in greenhouse. Soybean seedlings with fully expended primary leaves were used.

Cross-protection assays

Primary leaves of soybean seedlings (7-10 days after planting) were sap inoculated with the protecting isolates. The trifoliolate leaves were challenge-inoculated with the challenging virus using sap from infected plants. Mock-inoculated plants were employed as a control. In some experiments, inoculation with the challenging virus was made by the beetle vector. BLBs were collected from soybean fields and maintained on healthy soybean plants at room temperature for several days to one week. After a 48 hr starvation period, the beetles were allowed a 72 hr acquisition feeding on Ho1-infectred plants. Beetle regurgitant was tested by ELISA to confirm that they were free of virus

before acquisition and that they contained virus after maintaining in BPMV K-Ho1 infected plants. A group of 5 beetles were used to challenge inoculate each protected or mock-inoculated plant. Beetle feeding leaves were harvested from protected and unprotected plants once severe symptoms appeared on mock-inoculated plants.

Slot blot hybridization analysis

Viral RNAs were extracted from purified virions as described previously (Peden and Symons, 1973). Total RNA from BPMV-infected plants was isolated from either inoculated or systemically infected leaves (Naidu *et al.*, 1995). The RNA concentration was assessed spectrophotometricaly using UltrospecIII (Amersham). Procedures for preparation of cloned cDNA probes and conditions for pre-hybridization and hybridization were as described previously (Gu *et al.*, 2002). The blots were exposed to a phosphorimager screen and the images were visualized by a PhosphorImager 445 SI system and analyzed with the ImageQuant 4.1 program (Molecular Dynamics, Sunnyvale, CA).

Plant growth conditions and symptom documentation

Soybean plants were kept in a greenhouse maintained at 22°C with 16hr/8hr light/dark conditions. Disease symptoms were photographed with a digital camera (Olympus C-2500L).

Results

Cross-protection between BPMV isolates

The BPMV strain K-Da1 (subgroup II) induces mild symptoms on soybean plants (Gu *et al.*, 2002), whereas strain K-Ho1, a diploid reassortant, induces very severe symptoms (Chapter V; Gu *et al.*, 2002). To determine whether cross-protection occurs between BPMV strains, plants previously infected with K-D1 were challenge inoculated with K-Ho1. The results indicated that prior infection with the mild isolate K-D1 conferred protection against infection with the severe strain K-Ho1. The protected plants showed symptoms typical of isolate K-D1 and similar to those exhibited by control plants

in which the primary leaves were inoculated with K-D1 and the first trifoliolate leaves were rub inoculated with the inoculation buffer alone (compare panels A and B, figure 4.1). In contrast, the unprotected plants, which were rub inoculated onto their primary leaves with buffer and their first trifoliolate leaves with strain K-Ho1, showed distinct necrotic lesions on the inoculated leaves (data not shown). Furthermore, the unprotected plants developed severe systemic mottling and blistering symptoms (Figure 4.1C). Total RNA from 1st trifoliolate (challenge-inoculated leaf) and 3rd trifoliolate leaves of cross-protected and unprotected plants were isolated and subjected to slot blot hybridization analysis using a cloned cDNA probe to K-Ho1 RNA1 (subgroup I). As controls, viral RNA extracted from purified K-Ho1 and K-D1 virions as well as total RNA from K-Ho1 and K-D1-infected plants were used. Although purified viral RNA as little as 10 ng per slot was readily detectable, no K-Ho1 RNA (challenge virus) was detected in either inoculated or systemic leaves of any of three protected plants (Figure 4.2A). As predicted, no hybridization signals were evident neither with purified K-D1 virion RNA nor with RNA samples from K-D1-infected plants. Strong hybridization signals, on the other hand were observed with purified K-Ho1 virion RNA and with RNA samples from K-Ho1-infected plants. These results suggest that inoculation of soybean plants with a mild strain conferred complete protection against infection with the severe strain.

In a recent study, Valkonen et al. (2002) reported that potato and tobacco strains of *Potato virus A* (PVA) differ in their abilities to confer effective cross-protection. Since BPMV strains K-G7 and K-Ha1 were previously classified into two distinct strain subgroups (subgroup I and II, respectively; Gu *et al.*, 2002), it was of interest to compare subgroup I and subgroup II strains of BPMV for their cross-protection capabilities. Plants cross-protected by prior inoculation with either a subgroup I or a subgroup II strain were challenge inoculated with a virus in a different strain subgroup. Total RNA extracted from challenge-inoculated (the first trifoliolate leaves) and systemically-infected leaves were analyzed by slot blot hybridization using cloned cDNA probes specific to either subgroup I RNA2 or subgroup II RNA2. The results showed that cross protection was effective in all cases and that it is independent of the virus strain used as the protecting virus. No K-Ha1 specific RNA (subgroup II) was detected in either the inoculated or systemic leaves in plants protected by prior infection with strain K-G7.

Likewise, no K-G7 RNA was detected in plants previously inoculated with strain K-Ha1 (Figure 4.3 B and C). These results suggest that cross-protection between BPMV isolates blocks infection by the challenging virus at an early stage of infection.

Cross-protection is evident in plants challenge-inoculated with viral RNA

It was previously reported that cross-protection between TMV strains can be overcome if the protected plants were inoculated with TMV RNA (Sherwood and Fulton, 1982). To determine whether a similar mechanism is in operation in BPMV cross-protected plants, RNA isolated from purified BPMV virions of the severe strain K-Ho1 was used to challenge protected plants previously infected with the mild strain K-Ha1. Total RNA from challenge-inoculated and systemic leaves was subjected to slot blot hybridization analysis using a cloned cDNA probe to K-Ho1 RNA1. The results indicated that, following challenge inoculation with viral RNA (a total of 3.75 μg RNA were applied per plant), all unprotected plants (5 plants were tested) showed very severe symptoms typical of those induced by strain K-Ho1 (data not shown). Strong hybridization signals were detected when the RNA blot was hybridized with the K-Ho1 RNA1 probe. All of the 10 protected plants challenge-inoculated with either 3.75μg or 7.5μg virus RNA did not show severe symptoms. Three randomly selected plants were subjected to slot blot hybridization analysis. The results indicated the strain K-Ha1 confer protection against infection with the strain K-Ho1 (Figure 4.3). The data implied that cross-protection among strains of BPMV was not due to reencapsidation of challenging virus RNA.

Cross-protection is effective regardless of the means of inoculation

BPMV is mainly transmitted in nature by BLB and transmission efficiency of 70-80% has been observed (Giesler *et al.*, 2002). It is thus important to determine whether cross-protection is effective against challenge inoculation by the beetle vector. In this study, nonviruliferous BLBs, which were maintained on healthy soybean plants for a week, were confirmed to be BPMV-free by ELISA testing of regurgitants. Following a short starvation period, the beetles were allowed an acquisition feed period of 72 hr on soybean plants infected with the severe strain, K-Ho1. After a 48 hr starvation period,

beetles were allowed an inoculation feed period of 72 hr on either soybean plants previously infected with the mild strain K-Ha1 (protected plants) or mock-inoculated control plants (unprotected plants). The results indicated that all the unprotected plants exhibited the severe symptoms typical of strain K-Ho1, and the presence of K-Ho1 viral RNA was verified by slot blot hybridization analysis. None of the protected plants, on the other hand, showed severe symptoms and K-Ho1 viral RNA was not detected (Figure 4.4). These results were reproducible in a second independent experiment. Thus, cross-protection is effective against challenge inoculation by beetles.

Cross-protection is independent of soybean cultivar and timing of inoculation with the challenge virus

To determine whether cross-protection is influenced by the soybean cultivar used, the following four soybean cultivars were compared in cross-protection assays: Clark, Calhoun, Essex and York. The results indicated that cross-protection is equally effective with all four soybean cultivars tested (data not shown).

It was previously reported that BPMV-infected plants exhibit two flushes of symptoms separated by a recovery period and that symptom severity and BPMV concentration varied with leaf position (Calvert and Ghabrial, 1983). It was thus of interest to determine whether cross-protection can be overcome by challenge inoculation of leaves at different positions. Sap from soybean plants infected with the severe strain K-Ho1 was used to inoculate the first, second, third or fourth trifoliolate leaves of plants previously inoculated onto their primary leaves with the mild strain K-Ha1. Total RNA was isolated from both inoculated and systemic leaves and subjected to slot blot hybridization analysis using a cloned cDNA probe to K-Ho1 RNA1 (subgroup I). The results showed that cross-protection was effective in all cases regardless of the leaf position and timing of challenge inoculation (Figure 4.5).

Discussion

The results presented in this chapter indicated that cross-protection among strains of BPMV is effective and durable regardless of the viral strain used to confer protection. Inoculation with one strain, whether belonging to strain subgroup I or II, protects plants

completely from infection with another strain. Furthermore, cross-protection is independent of the soybean cultivars used and it cannot be overcome if the protected plants are challenge-inoculated with viral RNA. This finding rules out reencapsidation as the underlying mechanism for cross-protection among BPMV strains (de Zoeten and Fulton, 1975; Sherwood and Fulton, 1982, Lu *et al.*, 1998).

In the chapter V, I demonstrated that naturally occurring diploid reassortants or pseudorecombinants that contain two RNA1 segments belonging to two different strain subgroups induced very severe symptoms in soybean plants and that symptom severity correlated well with viral RNA accumulation. Since infected cells could support very well the replication of two distinct RNA1s, shortage in metabolites and/or limitation of replication sites are not likely to account for the observed cross-protection among BPMV strains (Bawden and Kassanis, 1945; Matthews, 1991).

In all the cross-protection experiments described here, the challenge viral RNA was not detected in either inoculated or systemically infected leaves. These results suggest that prior infection of soybean with one strain of BPMV may trigger PTGS that leads to degradation of the challenging viral RNA. It was proposed that the detection of short interfering RNA (siRNA), 21-26 nt in length, provides a reliable diagnostic feature of PTGS (Hamilton and Baulcombe, 1999; Hamilton *et al.*, 2002). If this diagnostic feature is also true for the BPMV system, BPMV sequence specific short (21–22 nt) and long (24–26 nt) size classes of siRNAs should be detected (Hamilton et *al.*, 2002). It will be of interest to determine whether inoculation of plants with BPMV transcripts from a recombinant BPMV vector expressing a strong suppressor of PTGS would confer cross-protection against challenge inoculation with a second strain.

Although cross-protection has been successfully utilized in control of several viral diseases, it is not recommended as a management strategy unless the viral diseases have reached epidemic level and no other methods can be applied (Fraser, 1998; Hull, 2002). As an annual crop, inoculation of soybean plants in the field may not be cost effective. Furthermore, BPMV interacts synergistically with SMV resulting in a devastating disease (Calvert and Ghabrial, 1983). Therefore, SMV-resistant soybean cultivars must be used if the cross-protection approach is to be implemented. It is important, however, to gain an understanding of the mechanism underlying cross-

protection among strains of BPMV. Such an understanding may lead to the development of novel control measures.

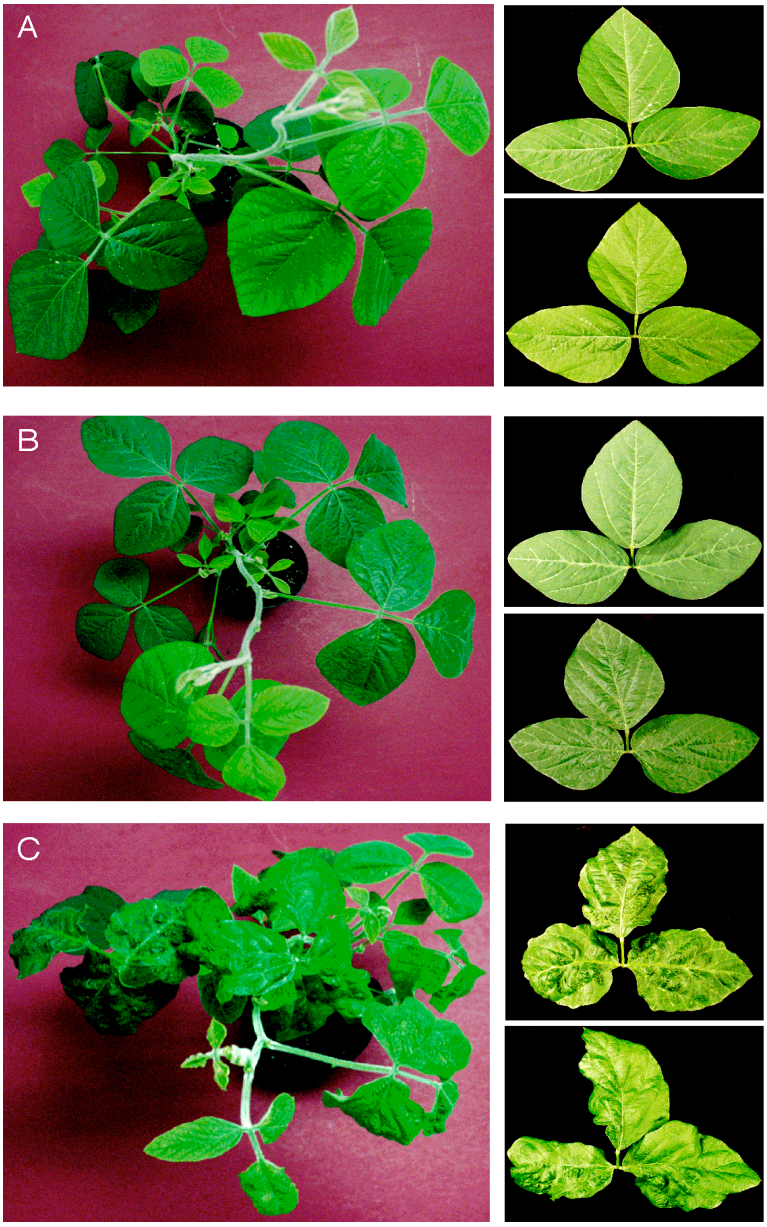

Figure 4.1 Symptoms elicited in cross-protected and unprotected soybean plants (c.v. Essex). A and B. Plants were inoculated with the mild strain K-Da1 on primary leaves and challenge-inoculated on the first trifoliolate leaves with either buffer (A) or BPMV K-Ho1 (B). C. The plant was mock inoculated on primary leaves with buffer and challenge-inoculated with the severe strain K-Ho1 on the first trifoliolate leaves. The two close up photographs to the right of panels A–C show symptoms exhibited by the 4th and 5th trifolilate leaves of each plant.

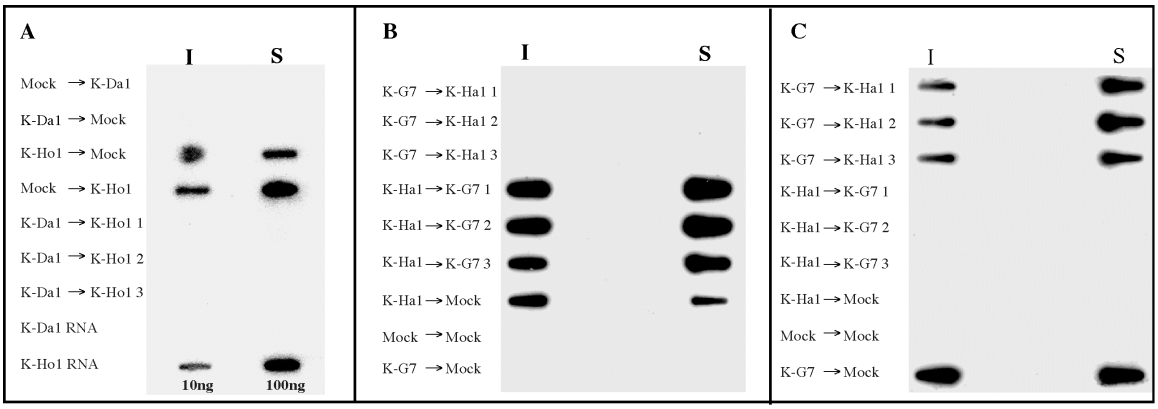

Figure 4.2 Slot blot hybridization analysis of total RNA from cross-protected and unprotected plants. A. Plants infected with the mild strain, BPMV K-Da1 (subgroup II), were challenge-inoculated with the severe strain K-Ho1. Three cross-protected plants (1-3) were selected for analysis by the slot blot assay. Controls included plants that were rub-inoculated with inoculation buffer on the primary leaves and challenge-inoculated on the first trifoliolate leaves with either K-D1 or K-Ho1. Additional controls included plants that were rub-inoculated with either K-D1 or K-Ho1 on the primary leaves and rub-inoculated on the first trifoloiolate leaves with inoculation buffer (mock). Total RNA from the challenge-inoculated (I) and systemically infected (S) leaves was subjected to slot blot hybridization analysis using a cloned cDNA probe to K-Ho1 RNA1 (subgroup I). The sequence of inoculations is indicated to the left of the hybridized membrane. Samples of purified viral RNA (10 and 100 ng/slot) were also included as controls. B and C. Plants previously inoculated on the primary leaves with strain K-Ha1 (subgroup II) or K-G7 (subgroup I) were challenge-inoculated with a BPMV strain of a different subgroup. Samples of total RNA from the challenge-inoculated (I) and systemically infected (S) leaves were subjected to slot blot hybridization analysis using a cloned cDNA probe to K-Ha1 RNA1 (subgroup II; panel B) or a cloned cDNA probe to K-G7 RNA1 (subgroup I; panel C). The sequence of inoculations is indicated to the left of the hybridized membrane.

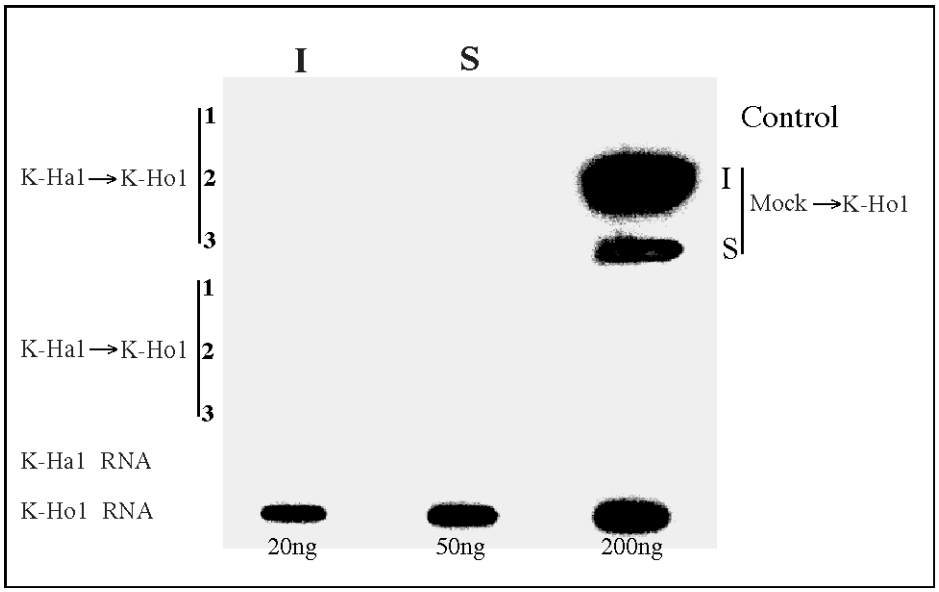

Figure 4.3 Cross-protection conferred by BPMV is not overcome by inoculation with viral RNA. Plants previously inoculated with the mild strain K-Ha1 (subgroup II) on the primary leaves were challenge-inoculated with viral RNA, isolated from purified virions of the severe strain K-Ho1 (3.75 or 7.5 mg RNA per trifoliolate leaves), on the first trifoliolate leaves. Control plants were rub-inoculated with inoculation buffer on the primary leaves and with BPMV RNA on the first trifoliolate leaves. Total RNA from the challenge-inoculated (I) and systemically infected (S) leaves was subjected to slot blot hybridization analysis using a cloned cDNA probe to K-Ho1 RNA1 (subgroup I). The sequence of inoculation is indicated to the left of the hybridized membrane (control treatment is indicated to the right of the blot). Purified BPMV K-Ha1 and K-Ho1 viral RNAs (20, 50 and 200 ng per slot) were included as negative and positive controls, respectively.

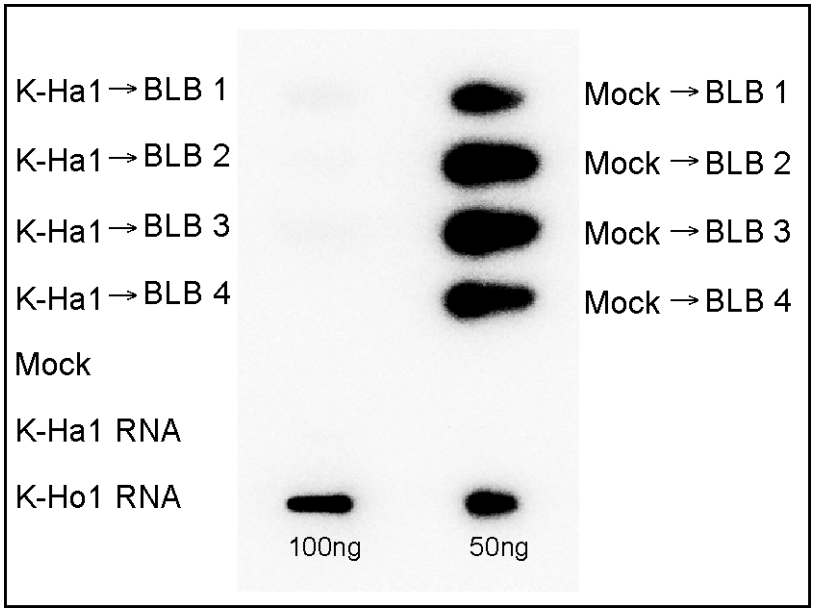

Figure 4.4 Cross-protection is evident regardless of the means of virus inoculation. Bean leaf beetles (BLBs) were allowed an acquisition feed for 72 hr on plants previously infected with the severe strain K-Ho1. Following a 48 hr starvation period, the BLBs were allowed an inoculation feed period of 72 hr on soybean plants previously inoculated on the primary leaves with the mild strain K-Ha1 (subgroup II) or on mock-inoculated control plants. Total RNA from systemically infected leaves was subjected to slot blot hybridization analysis using a cloned cDNA probe to K-Ho1 RNA1 (subgroup I). The sequence of inoculation is indicated to the left and right of the hybridized membrane. Purified BPMV K-Ha1 and K-Ho1 viral RNAs (50 and 100 ng per slot) were included as negative and positive controls, respectively.

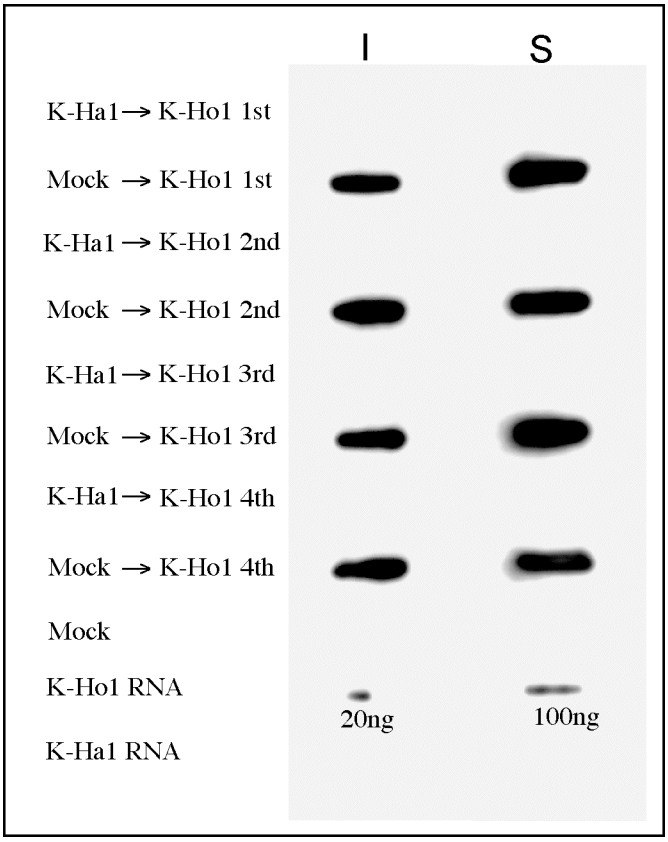

Figure 4.5 Cross-protection conferred by BPMV is effective regardless of timing of challenge inoculation. Soybean plants previously inoculated on the primary leaves with the mild strain K-Ha1 were challenge-inoculated with the severe strain K-Ho1 at different times on trifoliolate leaves at different leaf positions. The sequence of inoculation is listed to the left of the hybridized membrane. Total RNA from the challenge-inoculated (I) and systemically infected (S) leaves was subjected to slot blot hybridization analysis using a cloned cDNA probe to K-Ho1 RNA1 (subgroup I). Purified BPMV K-Ha1 and K-Ho1 RNAs (20 and 100 ng per slot) were used as negative and positive controls, respectively

Chapter V

Characterization of diploid reassortants of BPMV

Introduction

Viruses with mutltipartite RNA genomes are able to reassort their genome segments either in their hosts or insect vectors as a consequence of mixed infections. The frequency of dual infection is high both in the vectors and cultured cells infected with viruses in the family *Bunyaviridae* (Hendenson et al., 1995, Pringle, 1996; Rodriguez et al., 1998). Genetic reassortment has been observed in viruses belonging to several families of mutltipartite RNA viruses including the families *Bunyaviridae*, *Bromoviridae*, *Arenaviridae*, *Potyviridae* (genus *Bymovirus*), *Reoviridae* and *Orthomyxoviridae* (Ushijima et al., 1981; Barry et al., 1985; Murphy and Webster, 1990; Hendenson et al., 1995; Kashiwazaki Hibino, 1996; Fraile et al., 1997). It has been demonstrated that genetic reassortment plays an important role in the evolution, pathogenesis and epidemiology of some of these mutltipartite viruses.

Partial diploid reassortment, which arises as a consequence of dual infection, refers to the genotype of the virus progeny that is diploid for one or more genome segments and haploid for the other genome segments. This phenomenon has been observed with both plant and animal viruses in the laboratory under experimental conditions (Kashiwazaki and Hibino, 1996; Fraile at al, 1997; Rodriguez et al, 1998). It was further demonstrated that viruses with partial diploid genomes are transient and evolve into either the parental genotypes or genome reassortants (Kashiwazaki and Hibino, 1996; Rodriguez et al, 1998). It is not clear if genetic reassortment in natural populations also results in the emergence of partial diploid reassortants.

One of the most common ways to study reassortment is through phylogenetic analysis of the individual viral segments or the individual viral genes/proteins (Kool et al., 1992; Hendenson et al.; 1995; Miranda et al., 2002; Lin et al., 2004). As the sequences of an increasing numbers of viruses have recently become available, phylogenetic analysis of inter- and intraspecific reassortments is becoming more feasible. Other methods used to analyze reassortments include RNase protection assay (Fraile et

al., 1997), RT-PCR (Kashiwazaki and Hibino, 1996; Rodriguez et al., 1998), and Northern blot hybridization analysis (White et al., 1995).

Bean pod mottle virus (BPMV), a member of the genus *Comovirus* in the family *Comoviridae*, has a bipartite positive-strand RNA genome consisting of RNA-1 and RNA-2, which are separately encapsidated in isometric particles 28 nm in diameter. Purified BPMV preparations can be separated by density gradient centrifugation analysis into three viral sedimenting components designated top (T), middle (M) and bottom (B). The T component contains empty particles, whereas the M and B components contain single molecules of RNA-2 (approximately 3.6 kb) and RNA-1 (approximately 6.0 kb), respectively.

In this chapter, novel, naturally occurring, diploid reassortant strains of BPMV, which induce strikingly severe symptoms, were isolated and characterized. The possible mechanism(s) underlying the generation of diploid reassortant strains is discussed.

Material and Methods

RNA extraction and nucleic acid hybridization analysis

Procedures for total RNA extraction from infected plants, slot blot hybridization analysis and for preparation of radiolabeled probes were previously described (Gu, et al, 2002). Following hybridization, the slot blots were exposed to a phosphoimager screen and the images were visualized with a PhosphorImager 445 SI system and analyzed with the ImageQuant 4.1 program (Molecular Dynamics, Sunnyvale, CA).

RT-PCR amplification

Specific primers for K-Ho1 and K-G7 RNA1 were designed based on examination of sequence alignments. The 5' terminal primers specific for K-G7 and K-Ho1, which corresponded to nucleotide positions 254 to 281 were designated G7-spec-F (5'-CTCTTACCTCTCTTTCTC-3') and Ho1-Spec-F (5'CTCCTATCTTTCTTTCTC-3'), respectively. The 3' terminal primer specific for K-G7, which is complementary to nucleotide positions 3518 to 3547, was designated G7-spec-R (5'ATTCTGGGATAATA-

TTGT-3'). The 3' terminal primer specific for K-Ho1, which is complementary to nucleotide positions 3519 to 3547, was designated Ho1-Spec-R (5'-CATTCTGGGTAA-ACTGTC-3'). RT-PCR was performed using the SuperScript one step RT-PCR system with platinum *Taq* DNA Polymerase (Invitrogen). Samples of total RNA from systemically infected leaves were used as templates for RT-PCR.

Plant growth conditions and symptom documentation

Soybean plants (c.v. Essex) were kept in a greenhouse maintained at 22°C with 16h/ 8h. light/dark condition. Disease symptoms were photographed with a digital camera (Olympus C-2500L).

Results

The naturally occurring strain K-Ho1 is a diploid reassortant

Previous results showed that inoculation of soybean with transcripts derived from full-length cDNA clones of K-Ho1 RNA1 and RNA2 induced severe symptoms, but not as severe as those induced by the field isolate K-Ho1. Slot blot hybridization analysis of purified virion RNA of field isolate K-Ho1 showed a weak hybridization signal to a subgroup II RNA1-cDNA probe (Gu et al., 2002). This suggested that isolate K-Ho1 might have two species of RNA1. To test this hypothesis, slot blots of total RNA from soybean plants previously inoculated with either the field isolate K-Ho1 or with transcripts of K-Ho1 infectious cDNA clones ($RNA1_I$+$RNA2_{II}$) were probed with cloned cDNA probes specific for K-Ho1 RNA1 ($RNA1_I$) and K-Ha1 RNA1 ($RNA1_{II}$). The results revealed quite clearly that total RNA derived from plants infected with the field isolate K-Ho1 hybridized strongly with both probes, whereas total RNA derived from K-Ho1 transcript-infected plants only reacted only with the K-Ho1 $RNA1_I$ probe (Figure 5.1). Thus, strain K-Ho1 contains two distinct species of RNA1, i.e., diploid for RNA1, and only one species of RNA2. To determine whether diploid reassoratants are of common occurrence among BPMV field isolates, we examined the genotypes of several field isolates collected from soybean fields in six states and Canada (Table 5.1). Five out of 51 field isolates examined were found to be diploid reassortants, all of which caused

very severe symptoms on soybean plants under field and greenhouse conditions. All 5 reassortants are diploid for RNA1 and haploid for RNA2, i.e., they are partial diploids. In 4 reassortants, the RNA2 was derived from a subgroup II strain whereas the fifth reassortant had its RNA2 derived from a subgroup I strain (Table 5.1).

Enhancement of symptom severity induced by co-infection by two distinct RNA1s

It was previously established that symptom severity maps to RNA1 and that it is strain specific (Chapter IV). It was thus of interest to determine whether the strikingly severe symptoms induced by the diploid reassortants require the presence of specific RNA1s, or the mere presence of two distinct RNA1s is sufficient. Partial diploid pseudorecombinants ($RNA1_I$ + $RNA1_{II}$ +$RNA2_I$ or $RNA2_{II}$) were constructed from infectious transcripts derived from strains K-Ho1 (severe), K-G7 (moderate) and K-Ha1 (mild). The source of $RNA1_I$ transcripts was RNA1 cDNA from either strain K-Ho1 or K-G7, whereas strain K-Ha1 RNA1 cDNA provided the source of $RNA1_{II}$ transcripts. Sap from transcript-infected plants was used to inoculate healthy soybean plants and total RNA was extracted from systemically infected leaves 10 dpi and subjected to slot blot hybridization analysis. The results showed that both RNA1s ($RNA1_I$ and $RNA1_{II}$) were able to replicate in all plants inoculated with progeny from the diploid pseudorecombinants (Figure 5.2 A and C). RNA-2 from one subgroup (either I or II) was also detected (Figure 5.2 B and D). Apparently, encapsidation of viral RNA and virus movement were as efficient as the natural isolates. The presence of two genetically distinct types of RNA-1 in naturally infected soybean plants is associated with enhanced symptom severity. This phenomenon was also observed with comparable diploid pseudorecombinants constructed with the pertinent infectious transcripts (Figure. 5.3). Although diploid pseudorecombinants that contain one type of RNA1 and two types of RNA2 were shown to replicate in the infected plants, there was no enhancement in symptom severity. Furthermore, the RNA2 diploids were unstable since only a single RNA2 type was detected after three passages.

Co-infection with RNA1s derived from the same strain subgroup does not enhance symptom severity

To determine whether infection with two RNA1s derived from same subgroup may also enhance symptom severity, transcripts from K-Ho1 and K-G7 RNA1-cDNA clones (both belong to subgroup I strains) were combined with K-Ho1 RNA2 transcript. Five plants were inoculated with the $RNA1_{G7}$+$RNA1_{Ho1}$+$RNA2_{Ho1}$ transcript set. Total RNA, isolated from systemically infected leaves, was used as template for RT-PCR using K-G7 and K-Ho1 RNA1 specific primers. Samples of total RNA from the wild type K-Ho1 and K-G7 strains were used as controls. The results indicated that two out of five plants contained both K-Ho1 and K-G7 RNA1s (Figure 5.4, lanes 1 and 3). K-G7 RNA1 was the only RNA1 detected in two other plants (Figure 5.4, lanes 4 and 5), and K-Ho1 RNA1 alone was present in the remaining plant (Figure 5.4, lane 2). Similar results were obtained in two other independent experiments. Interestingly, one or the other RNA1 was lost in the subsequent passages of progeny from plants containing $RNA1_{G7}$+$RNA1_{Ho1}$ (data not shown). Plants containing RNA1s from the same subgroup did not show enhancement of symptom severity and induced symptoms similar to those produced by progeny derived from K-Ho1 infectious clones ($RNA1_{Ho1}$+$RNA2_{Ho1}$; Figure 5.4 B, compare panels 1 and 2 with panel 6).

Discussion

Emergence of new virus variants as a consequence of genomic reassortment has been observed in natural populations of animal and plant multipartite viruses (Fulton, 1980; Henderson et al., 1995; Lin et al., 2004). Contribution of reassortment to virus evolution may vary. High frequencies of reassortment have been reported for members of the family *Bunyaviridae* (Beaty et al., 1985; Pringle, 1996; Rodriguez et al, 1998). Fraile et al. (1997), on the other hand, suggest that reassortment is not important to the evolution of CMV in nature.

It is of interest to know the course of events leading to the development of reassortment. When the frequency and characteristics of reassortment of hantaviruses, which include rodent-borne viruses that cause pulmonary syndrome in infected patients, were studied, Rodriguez and co-workers (1998) observed that 30% of 294 progeny plaques contained more than one S or M segments derived from both parental strains after dual infection. They suggested that the viruses with partial diploid genomes were in

a transient state of genetic reassortment because most of the progenies of these diploid viruses evolved into either the parental or reassortant genotypes. No confirmed partial diploid reassortants were detected among 217 field isolates of CMV, even though 16% of the field isolates showed evidence of mixed infections by type 1 and type 2 isolates in subgroup I (Fraile et al., 1997). One out of 64 infected plants was indeed infected by a diploid reassortant of CMV (RNA1 and RNA3 from type 1 and RNA2 from both type 1 and type 2) following inoculation by aphids, which were allowed acquisition feeding on plants co-infected with CMV types 1 and 2 isolates. Diploid reassortment was also demonstrated in plants inoculated with a mixture of two different isolates of *Barley mild mosaic virus* (BaMMV, Kashiwazaki and Hibino, 1996). The stability of such reassortants was not tested.

This study represents the first report for the presence of stable diploid reassortants among field isolates of BPMV. The BPMV diploid reassortants are stable at least under greenhouse conditions since the strain K-Ho1 has been maintained in soybean since 1998 by successive passages at monthly intervals by sap inoculation. Both subgroup I and II RNA1s were present in the infected greenhouse plants after at least 60 passages (Figure 5.2 A and C). Whether the diploid reassortants are stable under field conditions has not been experimentally determined. It will be of interest to determine whether haploid reassortants can be generated from the diploid reassortant by beetle transmission. A high frequency (5/51) of diploid reassortants was found among field isolates suggesting that diploid reassortment may play an important role in the evolution of BPMV. It is possible that the diploid reassortants may serve as a reservoir for the emergence of new strains of multipartite viruses.

Both the virus host and the vector could be the site for the generation of diploid reassortment. Beaty and co-workers (1985) reported that a new virus genotype of bunyavirus was derived from the dually infected mosquitoes, which is the vector of bunyaviruses. However, they did not provide direct evidence that the vertebrate host could serve as a site of RNA reassortment. Aphid transmission experiments indicated that a diploid reassortment of CMV was detected in one out of 64 infected tomato plants. However, it was not clear how this diploid reassortment of CMV emerged (Fraile et al., 1997). As for BPMV, the diploid reassortants were first isolated during the recent

outbreak in BPMV incidence in the major soybean growing regions, which was apparently linked to the warm winters of the past few years that allowed the bean leaf beetle, vector of BPMV, to overwinter and emerge in the spring in unprecedented numbers. With large beetle populations and concomitant high incidence of BPMV, individual beetles may have been able to accumulate more than one strain of BPMV and introduce the mixture into healthy plants. My recent studies on cross-protection between BPMV strains (Chapter III) indicated that it was not possible to introduce different strains sequentially into the same plant via viruliferous beetles, as infection with the first strain cross-protects against infection with other strains. The finding that one of the RNA2 segments was lost in subsequent passages after plants were initially manually inoculated with a diploid pseoudrecombinant (diploid for RNA2 but haploid for RNA1) is of interest in view of the fact that all the field isolates of diploid reassortants contain only one type of RNA2 (Table 5.1). Diploid reassortment may originate from mixed infections with two distinct BPMV strains since these strains may accumulate in the vector prior to feeding on healthy plants. Alternatively, mixed infections may arise as a result of simultaneous feeding by beetles that have acquired two different BPMV strains.

Naturally occurring diploid reassortants as well as diploid pseudorecombinants with similar genotypes (i.e., containing $RNA1_I$ + $RNA1_{II}$) induced very severe symptoms compared to haploid pseudorecombinants, suggesting that enhancement of symptom severity is due to the presence of two distinct RNA1s. The mechanism underlying this apparent synergism is not clear. The results presented in Chapter VI showed that symptom severity maps to the Co-pro and Hel coding regions and that it correlates well with viral RNA accumulation. Recent evidence indicates CPMV Co-pro (32K protein) is involved in targeting the replication complex to the ER and that Co-pro and 60K(NTBM + VPg) induce massive proliferation of ER. The enhancement of symptoms severity in infections involving two distinct RNA1s may be related to an enhacement in ER proliferation and vesicle formation resulting in an increase in the total surface area accessible for virus replication. This, in turn, may account for the increase in viral RNA accumulation detected in infections with diploid reassortants. Such an enhancement, apparently does not occur if the two RNA1s are closely related (in the same subgroup), as their gene products (Co-pro and 60K) may compete for the same sites for membrane

binding. On the other hand, different Co-pro and 60K proteins may bind to different membrane sites.

Symptom severity determinants have been reported to function as suppressor of RNA silencing (Anandalakshmi et al., 1998; Brigneti et al., 1998; Qiu et al., 2002). Attempts to demonstrate RNA silencing suppressor by either Co-pro or Hel were unsuccessful. Liu at al (2004) showed the SCP of CPMV is a weak suppressor of RNA silencing. Sequence analysis demonstrated that the deduced amino acid sequences of SCP from strains K-Ho1 and K-Ha1 are identical (Figure 2.10). However, the diploid reassortant K-Ho1 induced very severe symptom and K-Ha1 elicited mild symptoms in infected plants. Thus, suppression of RNA silencing does not appear to play an important role in determining symptom severity induced by BPMV strains.

In summary, evidence was presented for the occurrence of BPMV diploid reassortants in nature. The frequency of diploid reassortment emergence may be enhanced by the beetle vectors since individual beetles may accumulate more than one strain and introduce the mixture into healthy plants. The diploid reassortants may also represent an intermediate form that leads to the emergence of genomic reassortants. If this is true, one can predict that diploid reassortants may also be present in natural populations of other comoviruses.

Table 5.1 Genotypes of field isolates of BPMV [a]

Location	No. of Isolates tested	Genotype[c] RNA1	Genotype[c] RNA2	Symptoms [b]	Reference
Arkansas	1	I	I	M	Gu, et al., 2002.
Illinois	5	II	II	M	This study.
	2	I, II	II	S	This study.
	1	I, II	I	S	This study.
Indiana	4	II	II	M	This study.
	1	I, II	II	S	This study.
Kentucky	1	I	I	Mo	Gu et al, 2002.
	7	II	II	M	Gu et al, 2002.
	1	I, II	II	S	This study.
	3	I	II	S, Mo	Gu et al, 2002.
Mississippi	2	II	II	M	Gu et al., 2002.
Virginia	1	I	I	M	Gu et al., 2002.
	5	II	II	M	Gu et al., 2002, This study.
Canada	17	II	II	M	This study.

[a] Strain subgroup affiliation (I, II or an assortment I/II) was determined by nucleic acid hybridization analysis.

[b] Abbreviations: M=mild (green mottling with little or no stunting/blistering); Mo=intermediate (green/yellow mottling, some stunting, leaf distortion and blistering); S=severe (yellow mottling, severe stunting and blistering and necrosis on inoculated leaves).

[c]Genotype classification: haploid; I or II, diploid I and II.

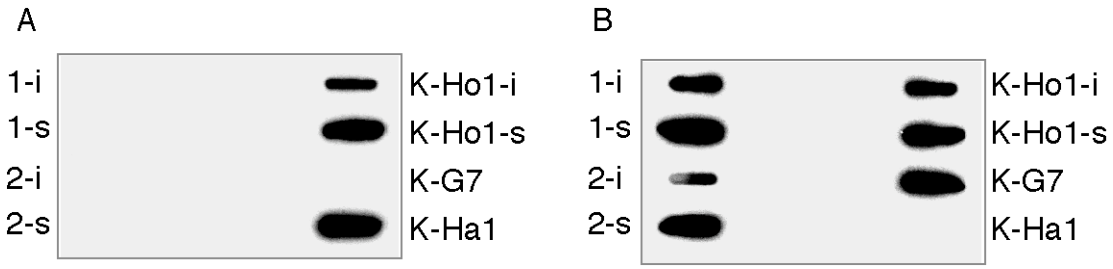

Figure 5.1. Slot blot hybridization analysis of BPMV RNA. Total RNA from soybean plants previously inoculated with either infectious transcripts derived from strain K-Ho1 cloned cDNAs of genomic RNAs (left column in panels A and B; extracts from two plants, 1 and 2, with "i" indicating inoculated leaf and "s" indicating systemic leaf), or extracts from plants infected with a subculture of strain K-Ho1 (top two slots of the right columns in panels A and B). Control RNAs from purified virions of strains K-G7 (subgroup I) and K-Ha1 (subgroup II) were blotted onto the lower slots of the right columns. The blots were probed with cloned cDNA probes specific for subgroup II (panel A) or subgroup I (panel B).

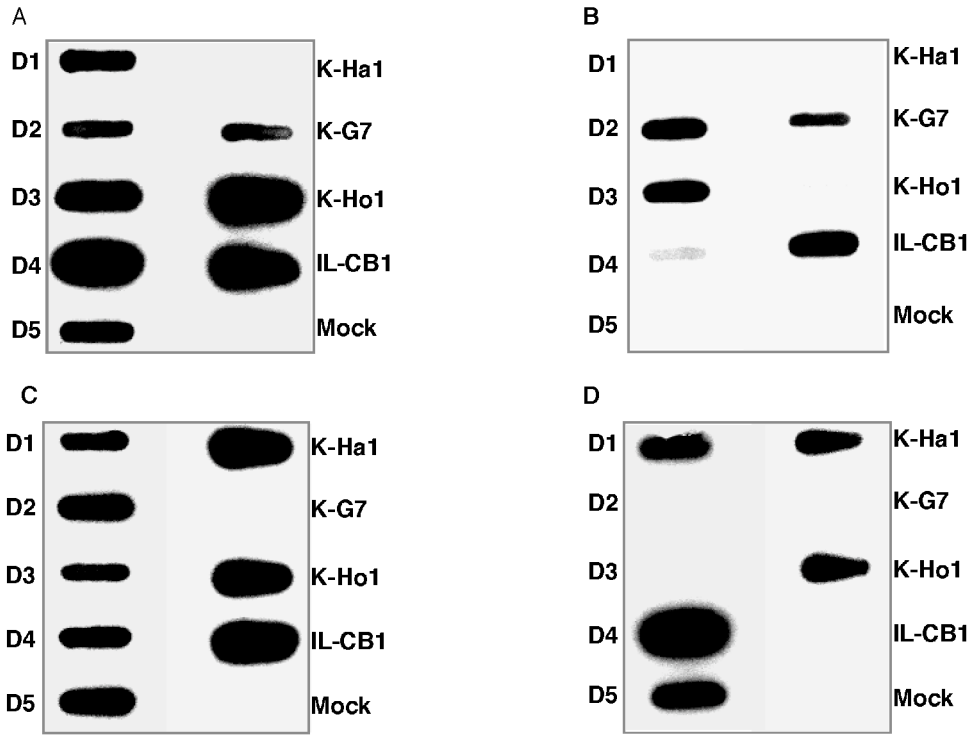

Figure 5.2. Slot blot hybridization analysis of total RNA from soybean plants previously inoculated with different diploid pseudorecombinants (D1 to D5). Right column in panels A-D: D1, $RNA1_{Ha1}$+ $RNA1_{Ho1}$ +$RNA2_{Ho1}$; D2, $RNA1_{Ha1}$+ $RNA1_{Ho1}$+ $RNA2_{G7}$; D3, $RNA1_{Ha1}$+ $RNA1_{G7}$ + $RNA2_{G7}$; M4, $RNA1_{Ha1}$+ $RNA1_{Ho1}$+ $RNA2_{Ha1}$; D5, $RNA1_{Ha1}$+ $RNA1_{G7}$ + $RNA2_{Ha1}$. The field isolates used to inoculate the soybean plants are indicated to the right. The blots were probed with cloned cDNA probes specific for subgroup I RNA1 (panel A), RNA2 (panel B) or subgroup II RNA1 (Panel C), RNA2 (Panel D).

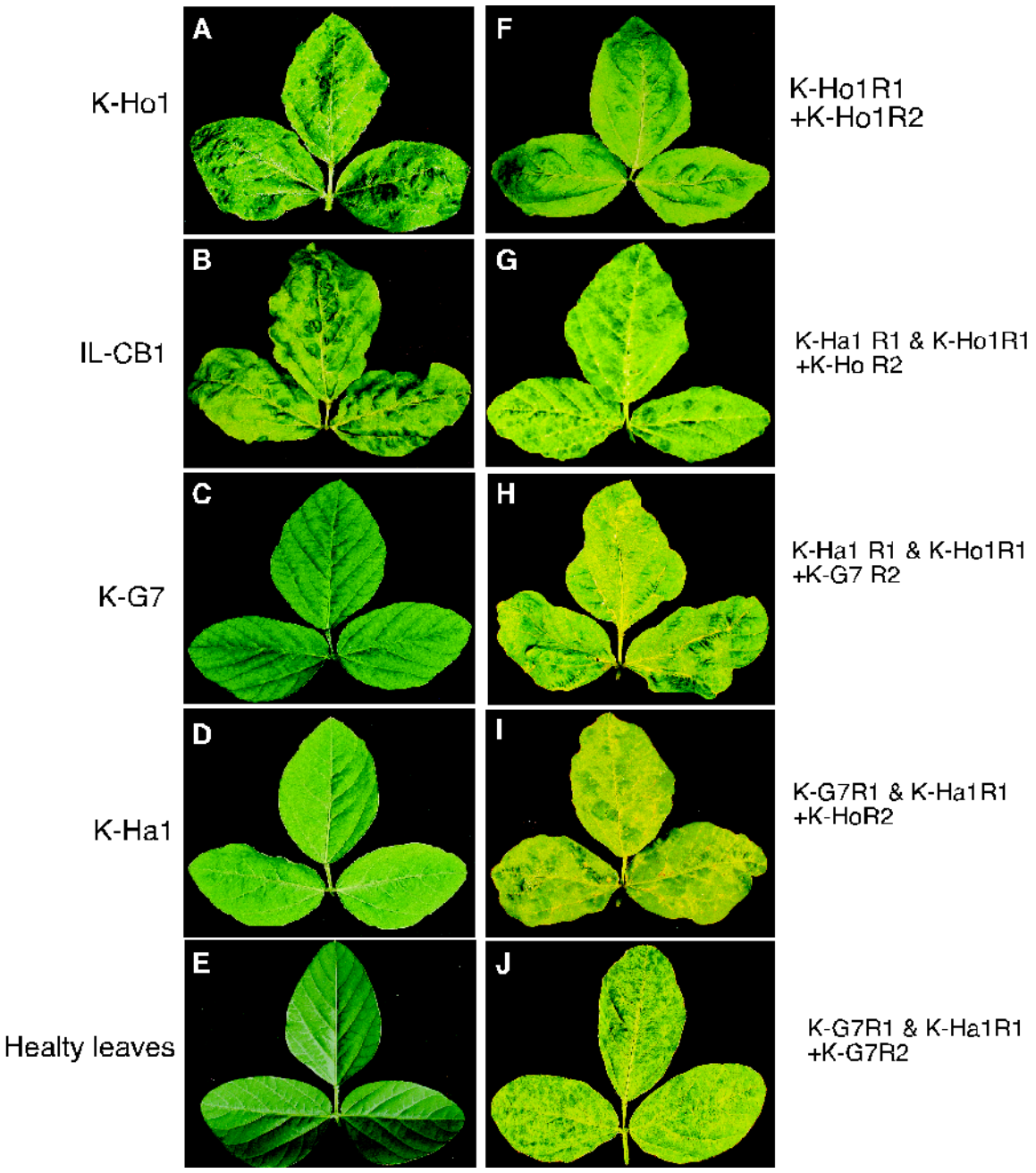

Figure 5.3 Systemic symptoms exhibited by soybean plants (cv Essex) previously inoculated with various field isolates of BPMV (left panels, A-D) or with different haploid/diploid pseudorecombinants (right panels, F-J). A trifoliolate leaf from a healthy control plant is shown in panel E. In all cases, the fourth trifoliolate leaves were photographed 21 dpi.

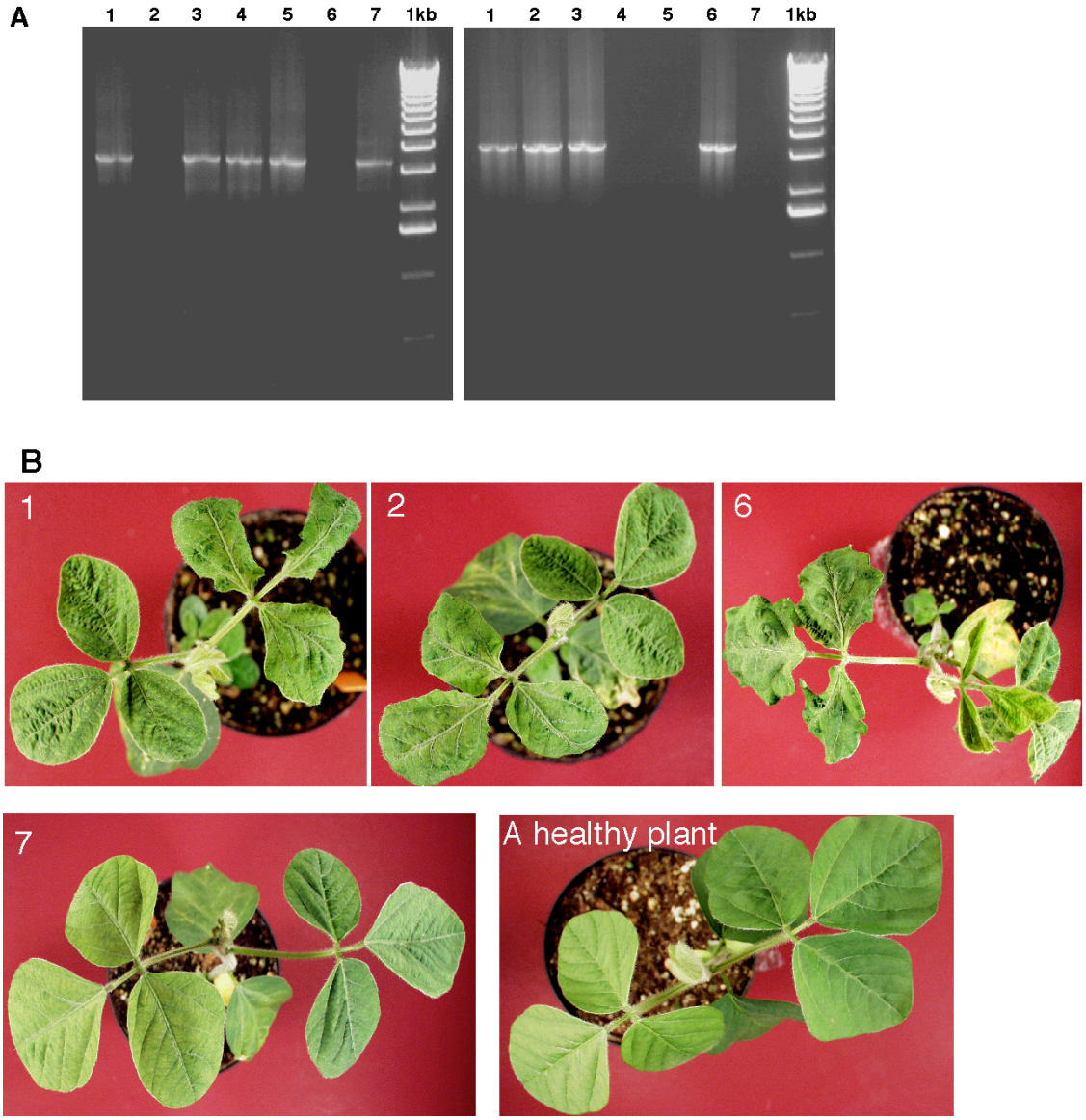

Figure 5.4. RT-PCR analysis and symptom development on soybean plants (cv Essex) previously inoculated with an RNA1 diploid pseudorecombinant (a transcript set consisting of $RNA1_{G7}$+$RNA1_{Ho1}$+ $RNA2_{Ho1}$). (A) RT-PCR analysis of total RNA extracted from systemically infected leaves 10 dpi using K-Ho1 RNA1-specific primers (left panel) or K-G7 RNA1 specific primers (right panel). Lanes 1-5, total RNA from five individual plants (designated 1-5) inoculated with the transcript set, $RNA1_{G7}$+$RNA1_{Ho1}$+ $RNA2_{Ho1}$; Lanes 6 and 7, total RNA from plants previously inoculated with the field isolates K-Ho1 and K-G7, respectively. (B) Symptoms induced on plants 1 and 2 (panels 1 and 2) and plants inoculated with the field isolates (panels 6 and 7). A healthy plant is shown in the lower panel to the right. Photographs were taken 10 dpi.

Chapter VI

The *Bean pod mottle virus* proteinase cofactor and putative helicase are symptom severity determinants

Introduction

Bean pod mottle virus (BPMV) is a member of the genus *Comovirus* in the family *Comoviridae* (Lomonossoff and Ghabrial, 2001). Like other comoviruses, BPMV has a bipartite positive-strand RNA genome consisting of RNA1 and RNA2, which are separately encapsidated in isometric particles 28 nm in diameter. Both genomic RNAs are polyadenylated and have a small basic protein, VPg, covalently linked to their 5' termini. The BPMV genome is expressed via the synthesis and subsequent cleavage of large polyprotein precursors. The complete nucleotide (nt) sequences of the two genomic RNAs of BPMV strain KY-G7 have been reported (Di et al.,1999; MacFarlane et al., 1991). BPMV RNA1 codes for five mature proteins required for replication (from 5' to 3', a protease cofactor (Co-pro), a putative helicase (Hel), a viral genome-linked protein (VPg), a protease (Pro) and a putative RNA-dependent RNA polymerase (RdRp), whereas RNA2 codes for a putative cell-to-cell movement protein (MP) and the two coat proteins (L-CP and S-CP).

BPMV is widespread in the major soybean-growing areas in many of the southern and southeastern states. A recent severe outbreak in BPMV incidence in the north central and northern Great Plains states is currently the cause of serious concerns to the soybean industry in this region (Giesler et al., 2002). BPMV is efficiently transmitted in nature, within and between soybean fields, by several species of leaf-feeding beetles in the family Chrysomelidae (mainly the bean leaf beetle). The outbreak was apparently linked to the warm winters of the past few years that allowed the bean leaf beetle, vector of BPMV, to overwinter and emerge in the spring in unprecedented numbers. Concomitant with the increased incidence of BPMV has been an augmentation in disease symptom

severity and the emergence of apparently new and unusual severe strains. Molecular characterization of such severe BPMV isolates revealed that they are reassortants/recombinants between two distinct subgroups of strains (Gu *et al.*, 2002; and unpublished). In order to develop efficient virus control strategies, it is essential to unravel the extent of diversity among BPMV isolates and to gain an understanding of the molecular basis of symptom severity associated with infection with such BPMV reassortants and sequence variants.

The two known subgroups of naturally occurring BPMV strains (subgroups I and II), as well as reassortants between the two subgroups, can be clearly distinguished based on nucleic acid hybridization and nucleotide sequence analyses (Gu et al., 2002). The availability of full-length cDNA clones of genetically distinct strains that differ in symptom severity allowed us to generate the appropriate chimeric constructs needed for mapping the determinants of symptom severity. In this study, we report that the Co-pro and Hel proteins specify symptom severity in BPMV-infected soybean plants. Furthermore, expression of the individual BPMV genes from a PVX vector showed that separate expression of either the Co-pro or Hel genes, unlike CP genes, induced necrotic lesion formation in *N. benthamiana*. Neither the Co-pro nor Hel protein, however, could be demonstrated as suppressors of RNA silencing using the agrobacterium infiltration assay.

Materials and methods

Virus strains

A list of the BPMV strains used is shown in Table 1. The strains are designated by the name of the Kentucky county where they were originally collected. The BPMV strains were propagated in the soybean cultivar 'Essex', and infected tissues were used for virion purification as previously described (Ghabrial et al., 1977). The complete nucleotide sequences of the genomic RNAs from strain K-G7 (Di et al., 1999; MacFarlane et al., 1991), K-Ha1 and K-Ho1 (Gu et al., 2002) have been determined and the sequences have been deposited in the GenBank (see Table 3.1).

Production of full-length cDNA clones

cDNA synthesis, cloning and sequencing of the genomic RNAs of strains K-Ha1 (subgroup II) and K-Ho1 (a reassortant between subgroups I and II) were previously described (Chapter II). Near full-length cDNA clones of K-Ho1 RNAs 1 and 2, which were missing the 5'-terminal 21 and 14 nts, respectively, were used as templates for generating full-length cDNA clones by PCR. An RT-PCR approach, however, was used to construct full-length cDNA clones of the genomic RNAs from strains K-G7 (subgroup I) and K-Ha1 (subgroup II). Because the 5' terminal 46 nts of the two genomic RNAs of all 3 strains are identical, the same forward (sense) primer (F1) was used for cDNA amplifications in all cases (Chapter II, figure 2.11). The F1 primer (Table 3. 2) contains a modified T7 promoter sequence (underlined), an extra G (bold) and the 5' terminal 39 nts of the genomic RNAs. A reverse (antisense) primer (R1) containing 18 (dT) residues (complementary to the poly(A) tail of viral RNAs), an engineered *Sal*I site (italicized) and 21 extra nucleotides (bold) for efficient restriction enzyme digestion, was used for first strand cDNA synthesis and for PCR amplification. For RT-PCR, first strand cDNA synthesis was made using viral RNAs from purified virions of K-Ha1 and K-G7 and a Superscript II reverse transcriptase kit (BRL), according to the manufacturer's instructions. The near full-length cDNAs of K-Ho1 genomic RNAs or first strand cDNA of strains K-G7 and K-Ha1 were amplified by 35 cycles of PCR including 3 cycles of 30 seconds at 94°C, 30 seconds at 47 °C and 6.5 minutes at 68°C and 32 cycles of 30 seconds at 94°C, 30 seconds at 62 °C and 6.5 minutes at 68°C. The High Fidelity Platinum® *Taq* DNA polymerase (Invitrogen) was used in all cases. PCR products were purified with the QIAquick PCR purification kit (QIAGEN, Valenchia, CA) according to the manufacturer's instructions. The purified PCR products (RNA1 cDNA of strain K-Ho1 and RNA2 cDNAs from all three strains) were then cloned into the pGem-T easy vector (Promega, Madison, WI). The PCR products (RNA1 cDNAs of strains K-G7 and K-Ha1), on the other hand, were cloned into the pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA). The plasmids containing full-length cDNA clones of the genomic RNAs from which infectious transcripts can be generated were designated as follows: pCR**G7**R1 and pG**G7**R2, pCRHaR1 and pGHaR2, and pGHoR1 and pGHoR2 (Fig. 3.1).

Chimeric constructs between K-G7 and K-Ho1 RNA1 cDNAs

Chimeric RNA1 constructs involving strains K-Ho1 and K-G7 (both belonging to subgroup I) were generated using plasmids pGHoR1 and pCRG7R1 (Figure 6.1). Eleven chimeric constructs, designated pHoG-1 through pHoG7-11 were produced. To generate construct pHoG-1, the *Nde*I fragment (nt positions 527-3764) of pCRG7R1 was replaced by the corresponding fragment from plasmid pGHoR1 (Figure 6.3A). Construct pHoG-2 was generated by replacing the *Apa*I-*Xho*I fragment of pGHoR1, with the corresponding fragment from plasmid pCRG7R1, which contains part of the multiple cloning sites of the pCR TOPO vector and the N-terminal 3068 nts of K-G7 RNA1 (Figure 6.1 and 6.3A). To generate construct pHoG-3, pCR**G7**R1 was digested with *EcoR*I and the *Eco*R1 fragment was cloned into pUC119, previously digested with *EcoR*I, to give rise to plasmid pUCG7R1 (Figure 6.1). The latter was then digested with *Sal*I and religated to generate plasmid pUCG7R1-1 (Figure 6.1). The C-terminal (nt positions 3069-6004) of plasmid pGHoR1 was replaced by the corresponding fragment from plasmid pUC**G7**R1 through double digestion with restriction enzymes *Apa*I and *Xho*I followed by religation of related portions of the two constructs. The intermediate plasmid pUCHoR1 was generated from pGHoR1 using the same strategy as described for pUCG7R1 (Figure 6.1). Constructs pHoG-4 and pHoG-5 were generated by exchanging the *Xba*I fragment between plasmids pUCHoR1 and pUCG7R1. The *Xba*I site located in the multiple cloning sites of construct pUCHoR1 and pUCG7R1 were removed by *Sal*I digestion followed by religation to produce plasmids pUCHoR1-1 and pUCG7R1-1, respectively (Figure 6.1). Constructs pHoG-6 and pHoG-7 were obtained by exchanging the *Xba*I-*Xho*I fragment between plasmids pUCHoR1-1 and pUCG7R1-1 (Figure 6.3A). To generate construct pHoG-8, plasmid pUCG7R1 was subcloned as an *Eco*R1-*Sal*1 fragment into the pGem-T easy vector (to avoid the redundant *Ava*I sites in vector pUC119) to generate plasmid pGG7R1 (Figure 6.1). Construct pHoG-8 was then generated by replacing the *Ava*I fragment (nt positions 1797-3069) of plasmid pGG7R1 with the corresponding fragment from plasmid pGHoR1. Similarly, construct pHoG-9 was created by replacing the *Ava*I fragment of plasmid pGHoR1 with the corresponding fragment from plasmid pGG7R1 (Fig. 6.4A). Constructs pHoG-10 and pHoG-11 were obtained by exchanging the *BsmB*I-

*Xho*I fragment (nt positions 2097 to 3068) between plasmids pGHoR1 and pGG7R1 (Figure 6.4A)

Chimeric constructs between K-Ha1 and K-Ho1 RNA1 cDNAs

Chimeric RNA1 constructs involving strains K-Ho1 and K-Ha1 (belonging to subgroups and I and II, respectively) were generated using plasmids pGHoR1 and pCRHaR1 (Figure 6.1). Seven chimeric constructs, designated pHoHa-1 through pHoHa-7 were produced. For the generation of construct pHoHa-1, a restriction enzyme *Bbv*CI site was introduced at the 5' terminal sequence of a fragment derived from plasmid pGHoR1 (at nt positions 1973 to 3764) by PCR using forward primer F2 and reverse primer R2, the latter contains a restriction enzyme *Nde*I site and is complementary to nt positions 3746 to 3764 of pGHoR1 (Table 6.2). The PCR product was cloned into pGem-T easy vector, and following digestion of the recombinant plasmid, the *Bbv*CI-*Nde*I fragment was exchanged for the corresponding fragment in pCRHaR1 to create construct pHoHa1 (Figure 6.5A). A similar strategy was used to generate construct pHoHa1-2 (Figure 6.5A) by PCR using primer F3 and R3 (Table 2), except that restriction enzyme sites *Sbf*I and *Bst*Z17I were introduced at the 5' and 3' termini of a cDNA fragment (nt positions 997-2542) derived from pGHoR1. Construct pHoHa-3 was produced by replacing the *Sbf1-Ppum*I fragment (nt positions 2100 to 3450) of pHoHa-2 with the corresponding fragment of pGHoR1 (Figure 6.5A). The *Apa*I-*Ppu*MI fragment of pGHoR1 construct (both pHoHa-3 and pGHoR1 contain one *Apa*I site upstream of the 5' terminus of the cDNA inserts), was exchanged with the corresponding one of pHoHa-3 to generate construct pHoHa-4 (Figure 6.5A). Construct pHoHa1-5 was generated by a similar strategy to that used to produce construct pHoHa1. Briefly, an *Xho*I site was created at the 5' terminal sequence of a cDNA fragment (nt positions 3070 to 6007) derived from pCRHaR1 (Figure 6.1) by PCR using primers F4 and R1 (Table 6.2). The PCR product was digested by restriction enzymes *Xho*I and *Sal*I and subsequently used to replace the corresponding fragment of pGHoR1 resulting in construct pHoHa-5 (Figure 6.5A). To generate construct pHoHa-6, the *Bbv*CI restriction site was created at the 3' terminal sequence of a fragment (nt positions 1 to 1979) derived from pGHoR1 by PCR using the M13 universal forward primer and primer R5. The

cloned c-DNA fragment (nt positions 1 to 1979) was excised from pGem-T easy vector by restriction enzymes *Apa*I and *Bbv*CI and used to replace the corresponding fragment of pCRHaR1 to produce construct pHoHa-6 (Figure 6.5A). To obtain construct pHoHa-7, the *Nhe*I and *Bsm*BI restriction sites were created at the 5' and 3' terminal sequences of a cDNA fragment (nt positions 1020 to 2104) derived from pCRHaR1 by PCR using primers F7 and R7 (Table 6.2). The *Nhe*I-*Bsm*BI fragment was then used to replace the corresponding fragment in construct pHoHa-5 giving rise to construct pHoHa-7.

Construction of recombinant PVX vector

The potato virus X (PVX) vector pP2C2S (Chapman et al., 1992) was used to express the coding regions of K-Ho1 Co-pro, Hel, Hel/VPg, S-CP, LCP/SCp polyprotein (LS-CP), as well as K-Ha1 Co-pro, Hel/VPg. The cDNA fragments containing the desired coding regions were amplified by PCR using plasmids pGHoR1 and pCRHaR1 as templates and designated specific primers (Table 6.1). A *Cla*I restriction and a start cordon were introduced at the 5' terminal sequence of each coding regions using the designated specific forward primer except for the Co-pro and LCP/SCP coding regions, which contain start codons at their 5' terminal sequences. A stop codon followed by *Sal*I site were added to each primer with exception of LS-CP and LCP/SCP coding regions, which already contain termination codons (in the latter cases, only the *Sal*I sites were engineered using specific primers; Table 6.2). The PCR products were individually cloned downstream a duplicated PVX CP promoter into the PVX vector, which was predigested with *Cla*I and *Sal*I. Additional constructs, which contained mutated versions of the K-Ho1 Hel and Co-pro coding regions, were generated by PCR mutagenesis. The mutants were generated by the introduction of two stop codons, which rendered them in effect untranslatable. All the recombinant PVX constructs were verified by restriction sites analysis. The BPMV inserts were subjected to sequencing analysis.

Construction of binary vectors

The binary vectors used in this study were derived from the pGD vector (Goodin et al., 2002). To generate constructs pGD-Hel, pGD-Co-pro and pGD-SCP, restriction enzyme sites for *Bgl*II and *Sal*I were individually introduced at the 5' and 3' termini,

respectively, of each of the coding regions of K-Ho1 cistron, Co-pro, Hel, SCP, by PCR. Start codon (AUG) or stop codon or both were created in each cistron, in case they were not present. The PCR products were cloned into pGem-T easy vector prior to digestion and religation with pGD vector. The PCR primers were listed in Table 6.2. PCR products were ligated into the pGD vector after digestion with restriction enzymes *Bgl*II and *Apa*I except PCR products of CR and MP genes, in which *Xho*I and *Apa*I sites were introduced by PCR at the 5' and 3' termini, respectively, and the product were subsequently digested. The above constructs were designated pGD-Pro, pGD-RdRp, pGD-CR, pGD-MP, pGD-LCP, pGD-R1, pGD-R2, pGD-Pro+Hel and pGD-LSCP, respectively.

In vitro transcription and _in vitro_ translation

Constructs pHoG4, 5, 6, 7 were linearized with *Sal*I. All other inserted BPMV cDNAs were released using double digestion with *Not*I and *Sal*I. PVX-derived construct were linearized with *Spe*I. *In vitro* transcription was performed according to Hu *et al* (1998) except that 100μl reaction mixtures were used. Samples (5 μl) of transcription reaction mixture were analyzed by electrophoresis on 1% agrose gel to assess yield and quality. RNA transcripts were used to inoculate fully expanded leaves of soybean or *Nicotiana benthamiana* by rub inoculation. Plasmid DNAs from infectious clones linearized by restriction enzyme *Sal*I were used for *in vitro* translation. BPMV genomic RNAs were translated using a TNT® quick coupled transcription/translation systems (Promega) in the presence of ^{35}S methionine at 30°C for 90 min. Translation products were analyzed in 7.5% SDS-PAGE. The gels were dried and virtualized by PhosphorImager 445I system (Molecular dynamics).

Plant growth condition and symptom documentation

Soybean and *N. benthamiana* plants were maintained kept in a growth chamber maintained at 22°C with 16hr/8hr light/dark conditions. Disease symptoms were photographed with a digital camera (Olympus C-2500L).

RNA extraction and nucleic acid hybridization analysis

Viral RNAs used for slot blot hybridization were purified according Peden and Symons (1973). Total RNAs extraction from virus or transcripts inoculated plants and procedures of slot blot hybridization were described before (Gu, et al, 2002). For northern hybridization, total RNAs were extracted using hot phenol method (Verwoerd et al., 1989). RNA samples (10μg) were denatured in presence of glyoxal and dimethyl sulfoxide (DMSO) and separated on 1% agrose (W/V) gel, which is dissolved in 10mM sodium phosphate (pH 6.3). RNAs were transferred onto Hybond-N^+ membranes (Amersham, Piscataway, NJ), according to the manufacturer's instructions. Probe synthesis and hybridization conditions were described previously (Gu, et al., 2002). The images were visualized by a PhosphorImager 445 SI system and analyzed with the ImageQuant 4.1 program (Molecular Dynamics, Sunnyvale, CA).

RT-PCR amplification

K-Ho1 and K-G7 RNA1 specific primers were designed based on examination of pair-wise sequence alignments of their cDNAs. The 5' terminal primers specific for K-G7 and K-Ho1 (G7-spec-F and Ho1-Spec-F) correspond to nucleotide positions 254to 281. The 3' terminal primers, G7-spec-R and Ho1-Spec-R, are complementary to nucleotide positions 3518 to 3547 or 3519 to 3547, respectively in the corresponding cDNAs. RT-PCR was performed using SuperScript one step RT-PCR system with platinum *Taq* DNA Polymerase (Invitrogen) according to the manufacturer's instructions. Samples of total RNA from systemically infected leaves were used as templates.

Agrobacterium growth condition and infiltration

The constructs expressing K-Ho1 genes were transformed into *Agrobacterium tumefaciens* strain C58C1 (An et al., 1998). The growth conditions of the transformed *A. tumefaciens* were as previously described (Goodin et al., 2002). The plasmid pGD, pZP-TCVCP and pZP-GFP were similarly transformed into *A. tumefaciens* strain C58C1. The transformed *A. tumefaciens* containing pGD serial vectors were maintained on LB media containing rifampicin (100 μg/ml), spectinomycin (100 μg/ml) and tetracycline (5 μg/ml). The *Agrobacterium* infiltration experiments were performed as described by Qu et al. (2003). Transgenic *Nicotiana benthamiana* plants expressing GFP were obtained

from Dr. David C. Baulcombe. Infiltrated plants showing fluorescence were photographed with a N90-S AF digital camera (Nikon, Tokyo).

Results

Symptom severity determinants map to RNA1

The availability of infectious RNA transcripts derived from full-length cDNA clones of the genomic RNAs from three strains of BPMV (Fig. 6.2 A), which differ in symptom severity and belong to two distinct strain subgroups (Table 6.1), allowed the use of 9 possible combinations of RNA1 and RNA2 transcripts as inocula. Thus, in addition to the three homologous pairs of genomic RNAs, 6 different pseudorecombinants were constructed and tested for their reaction on soybean. The results showed that all combinations that contained K-Ho1 RNA1 and any one of three RNA2s, regardless of origin, induced necrotic primary lesions on inoculated leaves (Figure 6.1B, panels d-f; arrows) and systemic mottling and blistering (Figure 6.1B, panels d-f). Although the symptoms induced by all transcript pairs that contain K-Ho1 RNA1 induced severe symptoms, they were not as severe as those produced by the natural K-Ho1 isolate (Figure 6.1B, panel b; see below for an explanation). Combinations of K-Ha1 RNA1 and any RNA2 regardless of origin induced mild symptoms (Figure 6.1B, panels g-i) similar to those produced by the naturally occurring isolate K-Ha1 (Figure 6.1B, panel c). Likewise, combinations containing K-G7 RNA1 and any of the three RNA2s induced moderate symptoms (data not shown) similar to those produced by the naturally occurring isolate K-G7 (Figure 6.1B, panel a). The reasons why the symptoms induced by pairs of transcripts containing K-Ho1 RNA1 did not induce as severe symptoms as those produced by the natural isolate were investigated in a separate study (Gu, H., Zhang, C. and Ghabrial, S.A., unpublished observations). The results of this study showed that the natural isolate contained two distinct species of RNA1 and only one species of RNA2. In addition to the RNA1 species belonging to subgoup I, used in the present study, another RNA1 belonging to subgroup II was isolated and sequenced, i.e., strain K-Hop1 is diploid for RNA1. The enhancement of symptom severity that mimics the natural isolate K-Hop1 was demonstrated when the pertinent combinations of infectious RNA1transcripts were used (Chapter V).

Symptom severity maps to the C-terminal region of the putative helicase encoded by K-Ho1 RNA1

The results described above indicated that symptom severity maps to RNA1 and that it is strain-specific, but not subgroup-specific. This was demonstrated by the finding that RNA1 derived from strain K-Ho1, but not K-G7, is associated with severe symptoms even though they both belong to subgroup I. The fact that these two RNA1s share 98% nt sequence identity provided an excellent opportunity to delineate the genetic determinant(s) of symptom severity. For this purpose, chimeric RNA1 constructs were generated by exchanging full or partial coding regions of the five RNA1-encoded mature proteins between the full-length RNA1 cDNA clones, as shown schematically in Figure 6.3A. Transcripts derived from the chimeric RNA1 cDNA constructs plus transcripts from K-G7 RNA2 cDNA or K-Ho1 RNA2 cDNA constructs were inoculated onto soybean seedlings and the resultant symptoms were recorded. As shown in Figure 6.3B, severe symptoms were observed only when the coding region of Hel was derived from K-Ho1 RNA1 (pGHoR1, pHoG-1, pHoG-2, pHoG-5 and pHoG-6). In construct pHoG-6, only the Hel coding region was derived from K-Hop1, but the remainder of the RNA1 molecule including the 5' and 3' UTRs were derived from K-G7 RNA1. Interestingly, symptom severity correlated well with accumulation of viral RNA, as demonstrated by northern hybridization analysis (Figsure 6.3 C and D).

To determine whether the entire Hel coding region is required for induction of severe symptoms (necrosis on inoculated leaves and systemic severe mottling and blistering), chimeric constructs in which the C-terminal and N-terminal regions of Hel were of different origins were generated (Figure 6.4A). Plants inoculated with transcripts derived from these constructs were monitored for symptom development. The results showed that symptom severity determinants map to the C-terminal half of the Hel protein. Thus the amino acid residues that specify symptom severity are contained in the C-terminal fragment of 324 amino acid residues (pHoG-10; Figure 6.4A). Amino acid sequence comparison of Hel between K-G7 and K-Ho1 strain, and also considering Hel sequences encoded by RNA1 from two additional mild strains belonging to subgroup II and one

severe strain in subgroup I) revealed two critical amino acid changes, Ser to Asn (aa position 359) and Leu to Phe (aa position 408; Table 6.3).

Both the protease cofactor and helicase coding regions are required for symptom severity

The finding that the K-Hop1 Hel C-terminal region was necessary and sufficient for induction of severe symptoms when the remainder of the RNA1 molecule was derived from a subgroup I RNA1 (e.g., strain K-G7), raised the question of whether an RNA1 from a subgroup II strain (e.g., strain K-Ha1) known to induce mild symptoms, would produce severe symptoms if it includes the Hel C-terminal region from K-Hop1 RNA1. To address this question, several chimeric constructs were generated in which coding regions were exchanged between K-Ho1 and K-Ha1 RNA1 cDNAs (Figure 6.5A). Chimeric K-Ha1 RNA1 constructs that contained the K-Hop1 Hel C-terminal region as well as the upstream K-Ho1 coding regions for VPg and Pro (construct pHoHa-6; Figure 6.5A) or the coding regions for VPg, Pro and RdRp (construct pHoHa-4; Figure 6.5A) induced mild symptoms comparable to those induced by the wild type K-Ha1 isolate. On the other hand, chimeric K-Ha1 RNA1 constructs that contained the K-Ho1 Co-pro and Hel coding regions (construct pHoHa-5) or the K-Ho1 Co-pro and the Hel C-terminal region (construct pHoHa-7) produced severe symptoms comparable with those induced with transcripts from wild type K-Ho1 RNA1. Interestingly, transcripts from construct pHoHa-6, which is similar to construct pHoHa-7 except that the Hel C-terminal region was derived from K-Ha1 RNA1 cDNA, induced mild symptoms. Taken together, both the coding regions of Co-pro and the C-terminal part of Hel are determinants of symptom severity. Northern hybridization analysis of total RNAs from plants inoculated with the various chimeric transcripts showed that symptom severity correlated well with higher levels of viral RNA accumulation (Figure 6.5B).

Neither the Co-pro nor Hel protein functions as a suppressor of RNA silencing

To determine whether Co-pro or Hel may function as a suppresser of RNA silencing an *Agrobacterium*-mediated transient expression system was used. Co-

expression of GFP and Co-pro or Hel proteins in transgenic *N. benthamiana* (line 16c) did not suppress RNA silencing, since no fluorescence was evident. On the other hand, co-expression of GFP and TCV-CP, a known suppressor of RNA silencing, induced intense fluorescence (Figure 6.6, panel A).

Expression of the individual Hel and Co-pro coding regions from a PVX vector induces necrosis on inoculated N. benthamiana leaves

The coding regions of Co-pro, Hel, S-CP and L-S CP proteins were cloned between the ClaI and SalI sites of the PVX vector pPVX2C2S to generate pPVX-Co-pro, pPVX-Hel, PVX-SCP and pPVX-SLCP, respectively. Infectious transcripts, generated from these plasmids, were used to inoculate N. benthamiana seedlings. The production of progeny RNA of the predicted size in the recombinant PVX-infected plants was verified by Northern hybridization analysis (Figure 6.7A). Whereas N. benthamiana plants inoculated with transcripts of pPVX2C2S, pPVX-SCP or pPVX-SLCP produced mosaic on inoculated leaves, those inoculated with transcripts from pPVX-Co-pro or pPVX-Hel induced necrotic lesions in addition to mosaic (Figure 6.7B).

Discussion

In this chapter, convincing evidence was presented in support of the conclusion that symptom severity induced by infection with BPMV maps to RNA1, and more specifically to the coding regions of the Co-pro and Hel. Although symptom severity correlated well with accumulation of viral RNA, neither the Co-pro nor Hel protein could be demonstrated as a suppressor of RNA silencing. This suggests that pathogenicity determinants need not be suppressors of RNA silencing-mediated host defense.

The replication of CPMV, the type species of the genus *Comovirus*, has been extensively studied at the molecular and cellular levels and an understanding of the functions of its encoded proteins is beginning to emerge (Pouwels et al., 2002). Therefore, analogy to CPMV will be made to explain the observations made in this study.

The major question to address is how Co-pro and Hel modulate symptom severity and viral RNA replication (accumulation) in infected cells. In CPMV-infected cells, Co-

pro (CPMV 32K), complexed to the 170K polyprotein, is specifically targeted to ER membranes. As a consequence of the interaction of 32K with ER membranes, the 170K polyprotein undergoes further proteolytic processing releasing the 60K (Hel/VPg) protein, which is inserted in the ER membranes and thus affixing the replication proteins to the ER membranes. The interaction of 32K and 60K proteins with ER membranes is known to trigger the proliferation and vesicle formation of ER and eventually to an increase of the total membrane surface available for viral RNA replication. With this in mind, I will attempt to explain the differences in symptom severity between strains K-G7 and K-Ho1, which differ in only two significant amino acids in the critical C-terminal region of Hel. As shown in Table 6.3, one significant change is from Ser in the mild strains to Asn in the severe strains. This change entails the potential availability of an additional N-glycosylation site in the Hel protein of severe strains. It is not known, however, whether the Asn at position 359 is indeed glycosylated. Glycosylation is known to play important roles in the folding and stability of glycosylated membrane proteins (Helenius and Aebi, 2001). It is possible that changes in conformation and stability of Hel, as mediated by the additional glycosylation site, may enhance membrane proliferation and vesicle formation and, in turn, increase accessible membrane surface and replication levels. This proposal could be tested by mutational analysis of Asn and flanking sequences and monitoring the resultant effects on symptom severity and RNA accumulation.

Table 6.1. A list of virus strains used.

Strain	Strain subgroup	% nucleotide identity[a]	GenBank accession #	Symptoms
KY-Graves 7 (K-G7)	I	100 100	NC_003496 NC_003495	Moderate; moderate mosaic, little or no stunting and blistering
KY-Hancock 1 (K-Ha1)	II	85.5 86.9	AF394606 AF394607	Mild; mild mosaic, no stunting or systemic blistering
KY-Hopkins 1 (K-Ho1)	I/II RNA-1 (I) RNA-2 (II)	 98.1 87.2	 AF394608 AF394609	Severe; necrosis on inoculated leaf and systemic stunting and blistering

[a] Percentage nucleotide sequence identity between the strain indicated and strain K-G7, as dertermined by the pairwise alignment using the GAP program

Table 6.2. Primers used in generation of BPMV chimera and PVX derived constructs.

Primer	Sequence (5'-3')
F1	*TAATACGACTCACTATA*GTATTAAAATTTTCATAAGATTTGAAATTTT GATAAAC[a]
R1	TTCCGCGGCCGCTATGGCCGACGTCGACTTTTTTTTTTTTTTTTTT[b]
F2	CcTGCAgGGATGTAAAGAGCTGC[b]
R2	gtaTaCTGCACTCGCCTTCCT[b]
F4	CCGCTcGAGGCCCAAAGTAGAAAGCCCAAC[b]
R5	TGAGCcTCAGCAGATGGCTC[b]
F7	gcTAGCTTGGGGAATGTGTA[b]
R7	GCTTCAGGAGAcGGTTCCAA[b]
Ho1/G7-F	TTTTGTTTGGAAATGGGC
Ho1/G7-R	TTGTCCAAATATTGTTAACC
PVX-Ho1-Hel-F	ATCGAT**ATG**GCCCTTTCAGTTCTAGAA
PVX-Ho1-Hel-R	GTCGAC**CTA**TTGCGCCTCGAGCACACT
PVX-Ho1-60K-R	GTCGAC**CTA**TTGGCCCTCAACCCACTC
PVX-Ho1-Co-pro-F	ATCGAT**ATG**AAGTTCTATCCTGGT
PVX-Ho1-Co-pro-R	GTCGAC**CTA**TTGGAACTGTCCCTTAAT
PVX-Ho1-Co-pro-F'[d]	ATCGAT**ATG**AAGTTCTATCCTGGT**TAG**AATATTTCT**TAA**ATTG
PVX-Ho1-Hel-F'	ATCGAT**ATG**GCCCTTTCAGTTCTA**TAA**GGGTTGGCAACA**TAG**CTC
PVX-Ho1-LCCP-F	CCATCGAT**ATG**GAAACAAATTTGTTTAAATTG
PVX-Ho1-SCP-F	CCATCGAT**ATG**GAATCCATTTCACAGCAAACTGTT
PVX-Ho1-SCP-R	TTCCGCGGCCGCTATGGCCGACGTCGAC**TCA**TGCAGAAGATTCCGC
PVX-Ha1-Co-pro-F	ATCGAT**ATG**AAGTTTTATCCAGGACAAAATGTCTCTGAAATTG
PVX-Ha1-Co-pro-R	GTCGAC**CTA**CTGGAATTGCCCTTTAATTTCTTCAGAAGG
PVX-Ha1-60K-F	ATCGAT**ATG**GCTCTTTCTGTTTTAGAGGGATTGGCAACACAACTC
PVX-Ha1-60K-R	GTCGAC**CTA**CTGGCCTTCAACCCATGCTCTGCGCTT
PGDHel-F	AGATCT**ATG**GCCCTTTCAGTTCTAGAA
PGDHel-R	GTCGAC**CTA**TTGCGCCTCGAGCACACT
PGDCo-pro-F	AGATCT**ATG**AAGTTCTATCCTGGT
PGDCo-pro-R	GTCGAC**CTA**TTGGAACTGTCCCTTAAT
PGDSCP-F	AGATCT**ATG**TCCATTTCACAGCAAACTG
PGDSCP-R	GTCGAC**TCA**TGCAGAAGATTCCGC

[a] T7 promoter is italic.

[b] Introduced restriction sites are underlined, and modified nucleotides are indicated by lowercases.

[c] Start and stop codons are bold.

[d] Frameshift mutant primer.

Table 6.3. Amino acid sequences comparison of the C-terminal region of Hel.

Amino acid position	Virus strain				
	K-G7 Mo (I)	K-Ho1 S (I)	IL-Cb1 S (I)	K-Ha1 M (II)	IL-Cb1 IL (II)
318	Tyr	Asn	Asn	Asn	Asn
335	His	His	His	Tyr	Tyr
359	**Ser**	**Asn**	**Asn**	**Ser**	**Ser**
365	Tyr	Asn	Asn	Asn	Asn
380	Tyr	Ser	Ser	Ser	Ser
382	Asp	Asp	Asp	Asn	Asn
386	Asn	Asn	Asn	Glu	Glu
391	Glu	Glu	Glu	Asp	Asp
392	Arg	Arg	Arg	Lys	Lys
393	Asn	Asn	Asn	Ser	Ser
408	**Leu**	**Phe**	**Phe**	**Val**	**Val**
419	Thr	Thre	Thr	Ser	Ser
420	Val	Val	Val	Ile	Ile
426	Asp	Asp	Asp	Glu	Glu
432	Ala	Ala	Ala	Leu	Leu
522	Thr	Thr	Thr	Ser	Ser

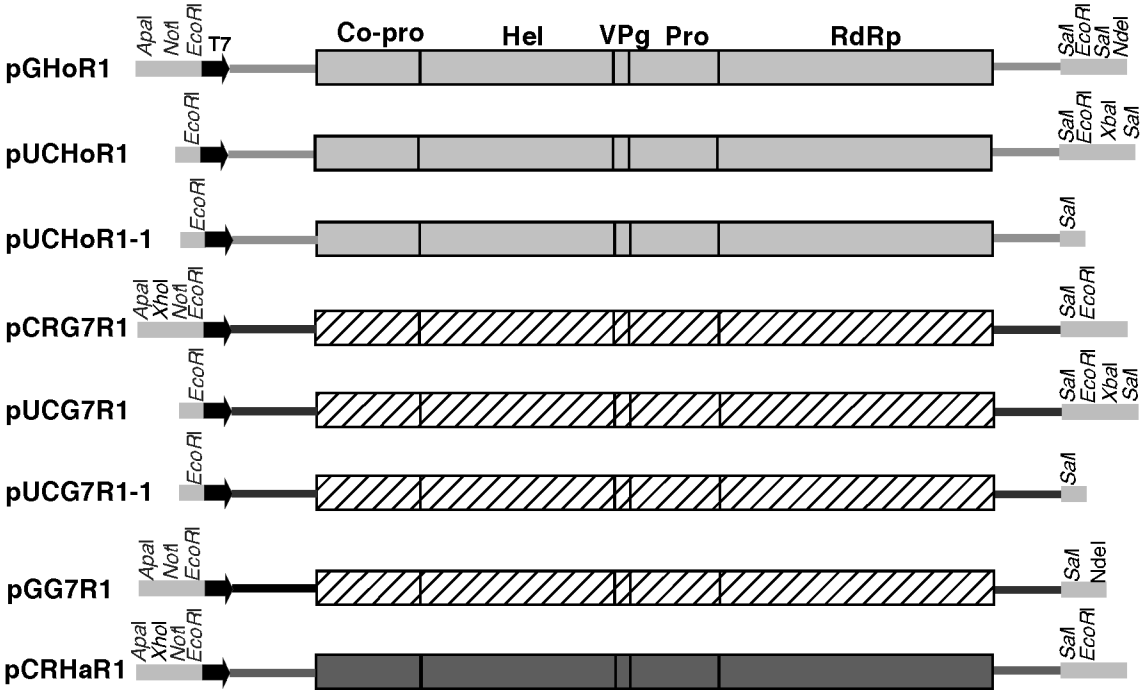

Figure 6.1. Schematic representation of BPMV RNA1 cDNA constructs used for generation of infectious transcripts.

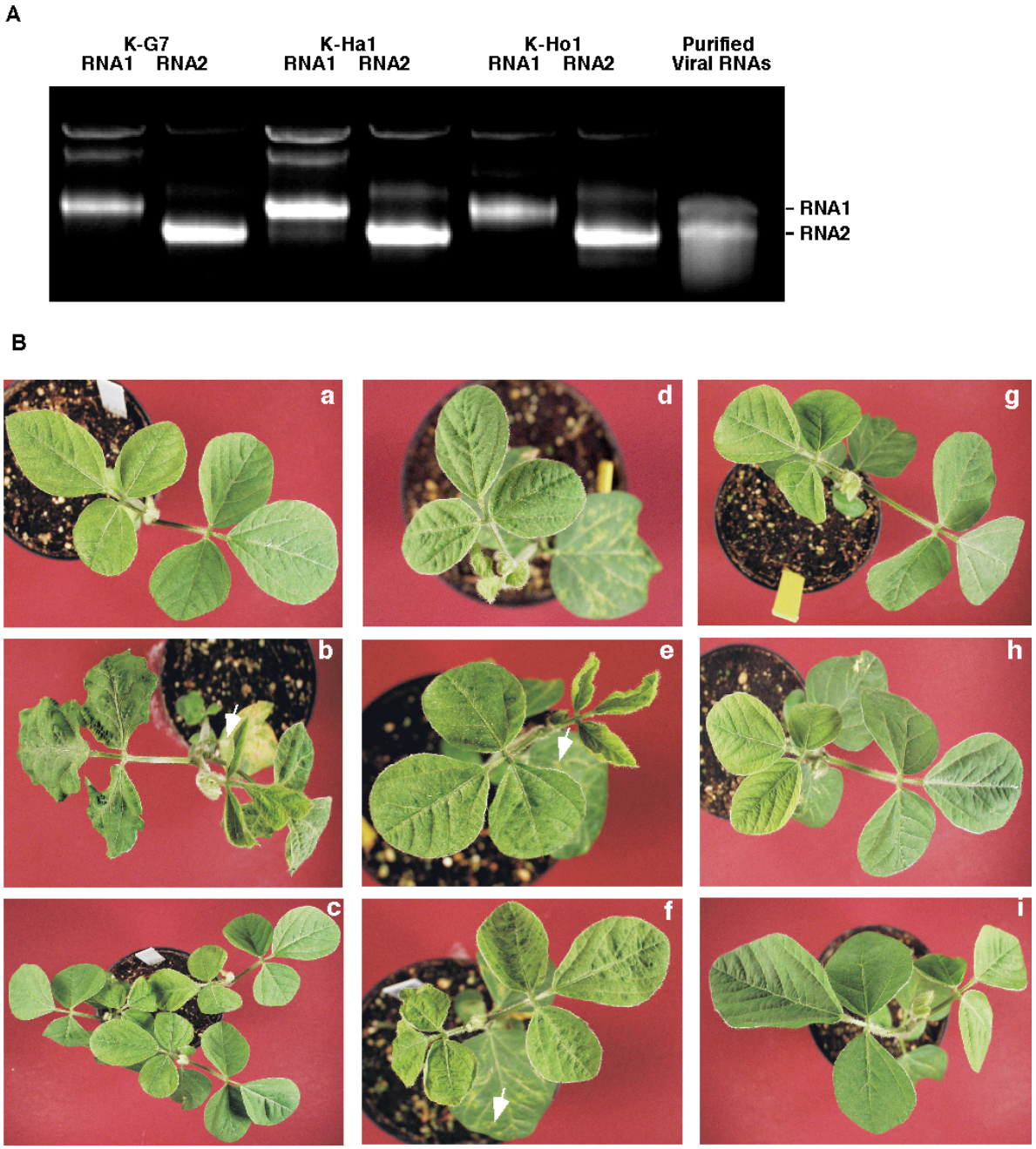

Figure 6.2. Symptom severity is mapped to BPMV RNA1. (A) Agarose gel electrophoresis of transcripts generated from full-length cDNA clones of genomic RNAs from three strains of BPMV. The gel was stained with ethidium bromide. The positions of RNA1 and RNA2 are indicated to the right. (B) Symptoms induced on soybean plants previously inoculated with field isolates and pseudorecombinants. Panels a-c. Symptoms induced by the field isolates K-G7, K-Ho1, K-Ha1, respectively; panels d-f, symptoms induced by inoculation with three different transcript sets that differ in the origin of RNA2: $RNA1_{Ho1}$+$RNA2_{G7}$, $RNA1_{Ho1}$+$RNA2_{Ho1}$ and $RNA1_{Ho1}$+$RNA2_{Ha1}$, respectively; panels g-I, symptoms induced by inoculation with additional three transcript sets that differ in the origin of RNA2: $RNA1_{Ha1}$+$RNA2_{G7}$, $RNA1_{Ha1}$+$RNA2_{Ho1}$ and $RNA1_{Ha1}$+ $RNA2_{Ha1}$. Necrotic lesions on inoculated leaves are indicated by arrows.

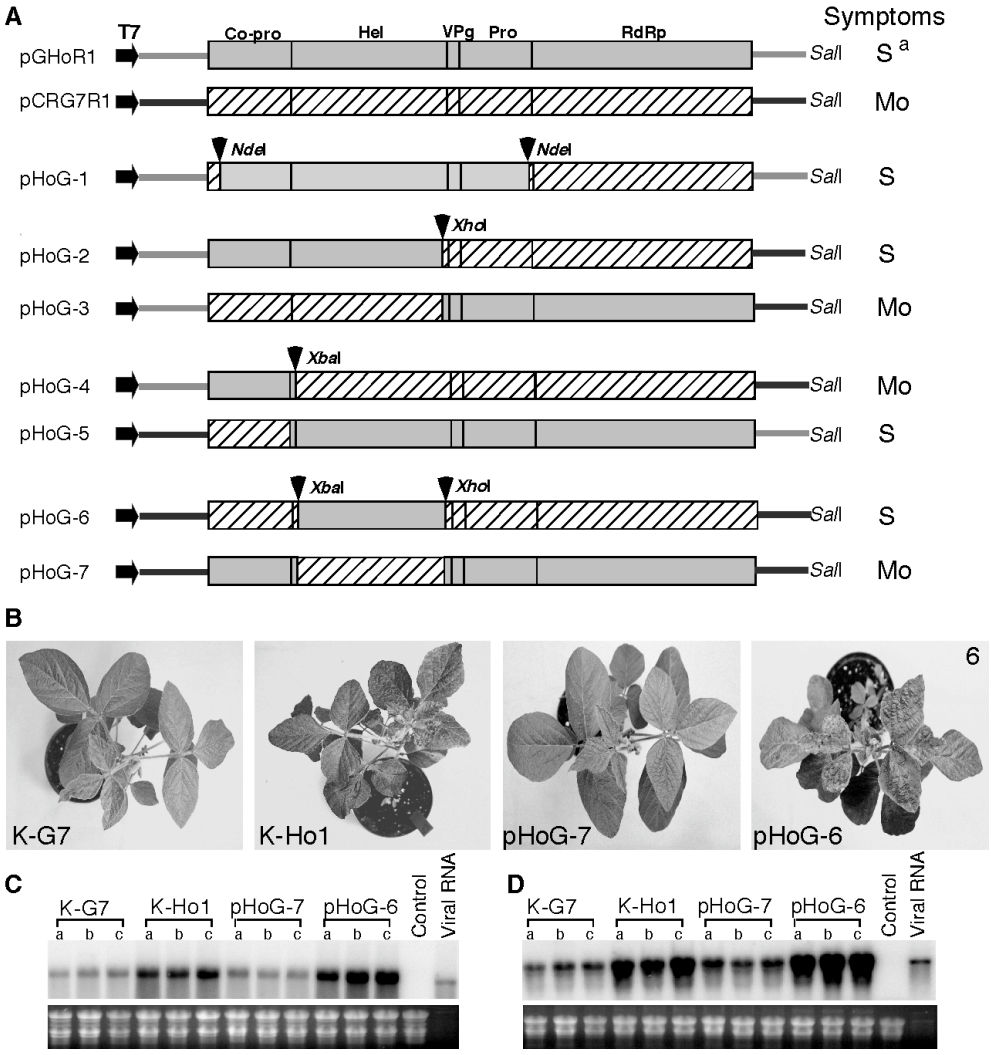

Figure 6.3. Mapping the symptom determinants in subgroup I of BPMV strains. (A) Schematic representation of BPMV RNA1 and chimeric constructs. The coding regions in K-Ho1 RNA1 are shown in gray boxes, and those in K-G7 RNA1) are indicated with hatched boxes. Restriction enzyme sites used to generate chimeric constructs are indicated with arrowheads. S, severe symptoms; Mo, moderate symptoms. (B) Symptoms induced on soybean plants (cv. Clark) inoculated with infectious transcripts. Numbers at the corners of each panel refer to the number of the RNA1 cDNA construct used to generate the transcripts. RNA2 transcripts used in all experiments was derived from the K-Ho1 RNA2 cDNA construct. (C) and (D). Northern blot hybridization analysis to assess viral RNA1 (panel C) and RNA2 (panel D) accumulation in the infected soybean plants 10 dpi. In panel C, an RNA1 cDNA fragment with sequences absolutely conserved between K-Ho1 and K-G7 RNA1s was used as a template template to synthesize the probe by PCR amplification using Primers Ho1/G7-F, Ho1/G7-R (Table 2). Three plants (a, b, c) were test for each treatment. RNAs from mock inoculated plants (M) and purified virions (V) are used as controls. Levels of RNA loading are assessed by ethidium bromide staining of ribosomal RNA (bottom panels).

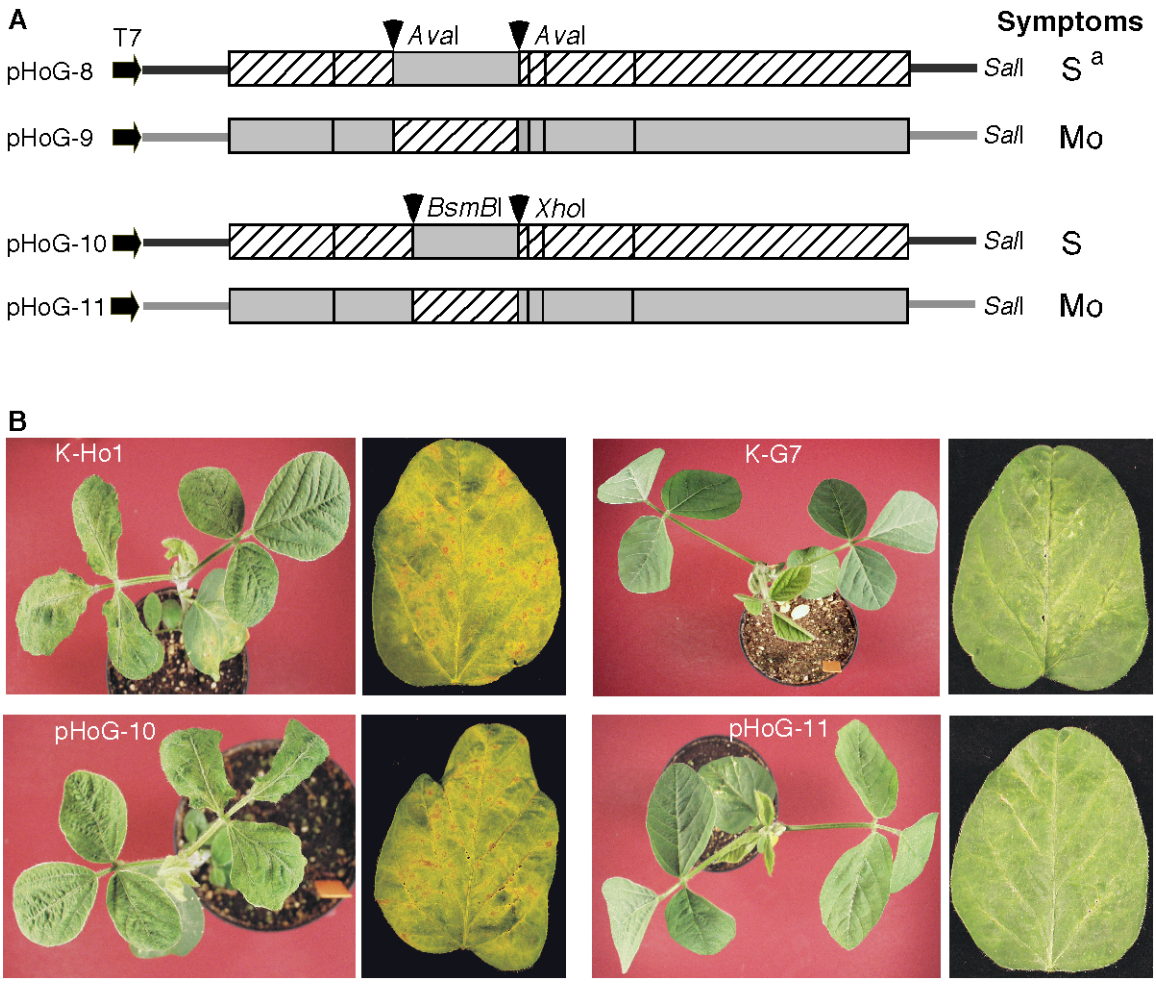

Figure 6.4. The helicase C-terminal region is critical for symptom severity. (A) Schematic representation of BPMV chimeric constructs generated by exchanging cDNA fragments between K-Ho1 and K-G7 RNA1-cDNA constructs. The letters S and Mo to the right refer to the symptoms (severe and moderate, respectively) induced by the indicated tranbscripts. (B) Symptoms induced by transcripts from chimeric clones (pHoG-10 and pHoG-11) and the parental constructs (pGHoR1 and pCRG7R1). The inoculated leaves are shown in the panels at right side of each of the photographed plants. Photographs were taken 10 dpi.

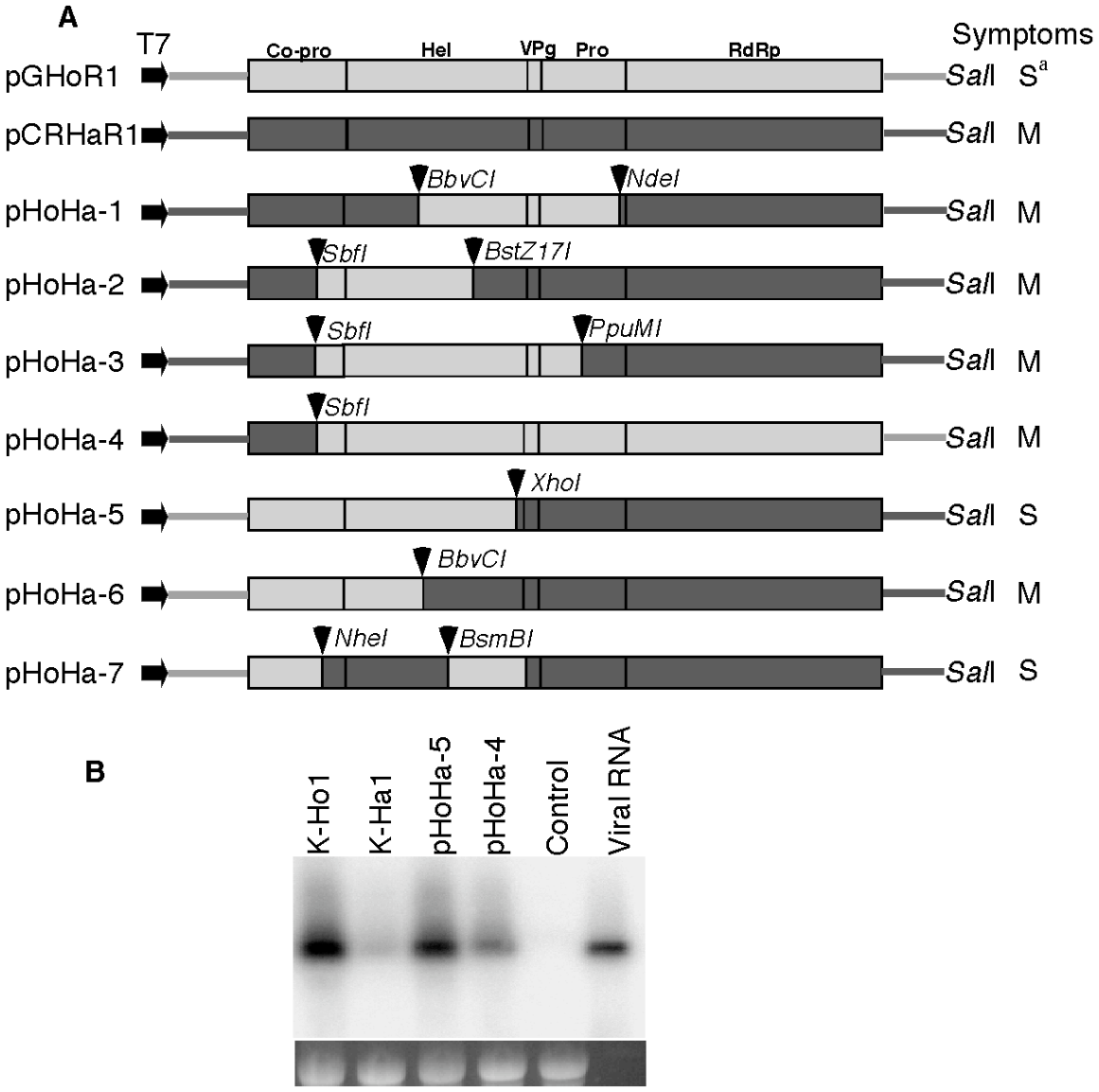

Figure 6.5. Mapping the symptom determinants in BPMV subgroup II strains. (A) Schematic diagrams of BPMV chimeric constructs generated by exchanging fragments between K-Ho1 RNA1 cDNA (subgroup I; gray) and K-Ha1 RNA1 cDNA (subgroup II; black). Symptom severity induced by inoculation with tranbscripts derived from the indicated constructs was indicated to the right by the letters, S (severe) and M (mild). (B) Northern blot hybridization analysis for assessment of BPMV RNA2 accumulation. Numbers above the individual lanes refer to the RNA1 constructs used for inoculation. Total RNA was extracted from infected soybean plants 10 dpi. Total RNA RNA from mock inoculated plants (lane M) and purified virions (lane V) were used as controls. K-Ho1 RNA2 cDNA was used as templates to synthesis the probe. The lower panels show the ethidium bromide-stained RNA gels before transfer as a loading control

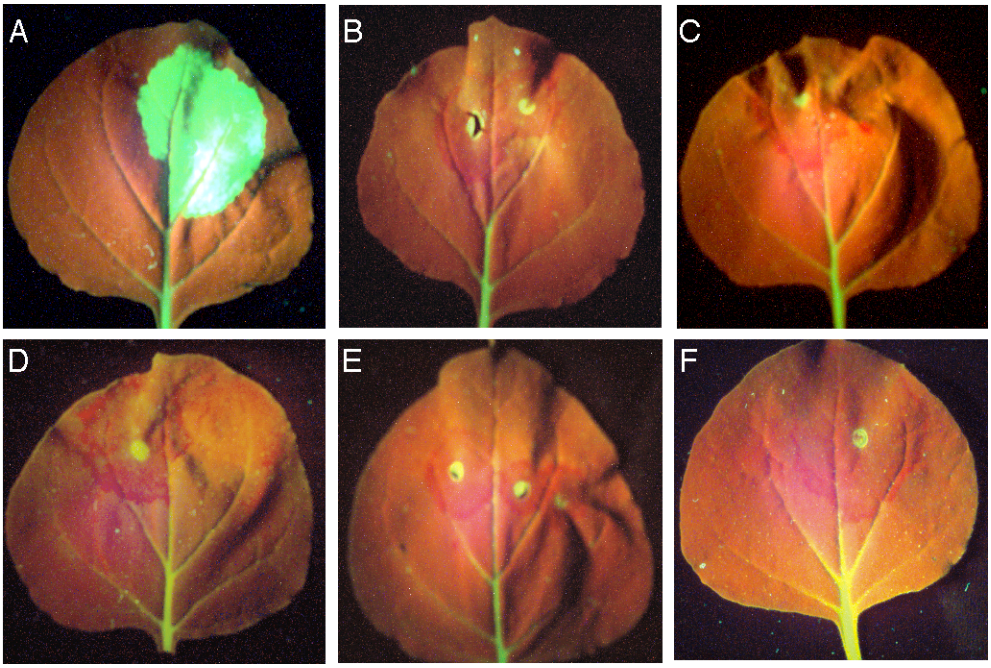

Figure 6.6. Identification of the BPMV suppressor of RNA silencing. Transgenic *Nicotiana* plants were agroinoculated with pGD-GFP and (A) pZP-TCVP, (B) pGD, (C) pGD-Hel, (D) pGD-Co-pro, (E) pGD-SCP. (F) Transgenic *Nicotiana* plant leaf was agroinfiltrated with pGD-GFP alone as negative control.

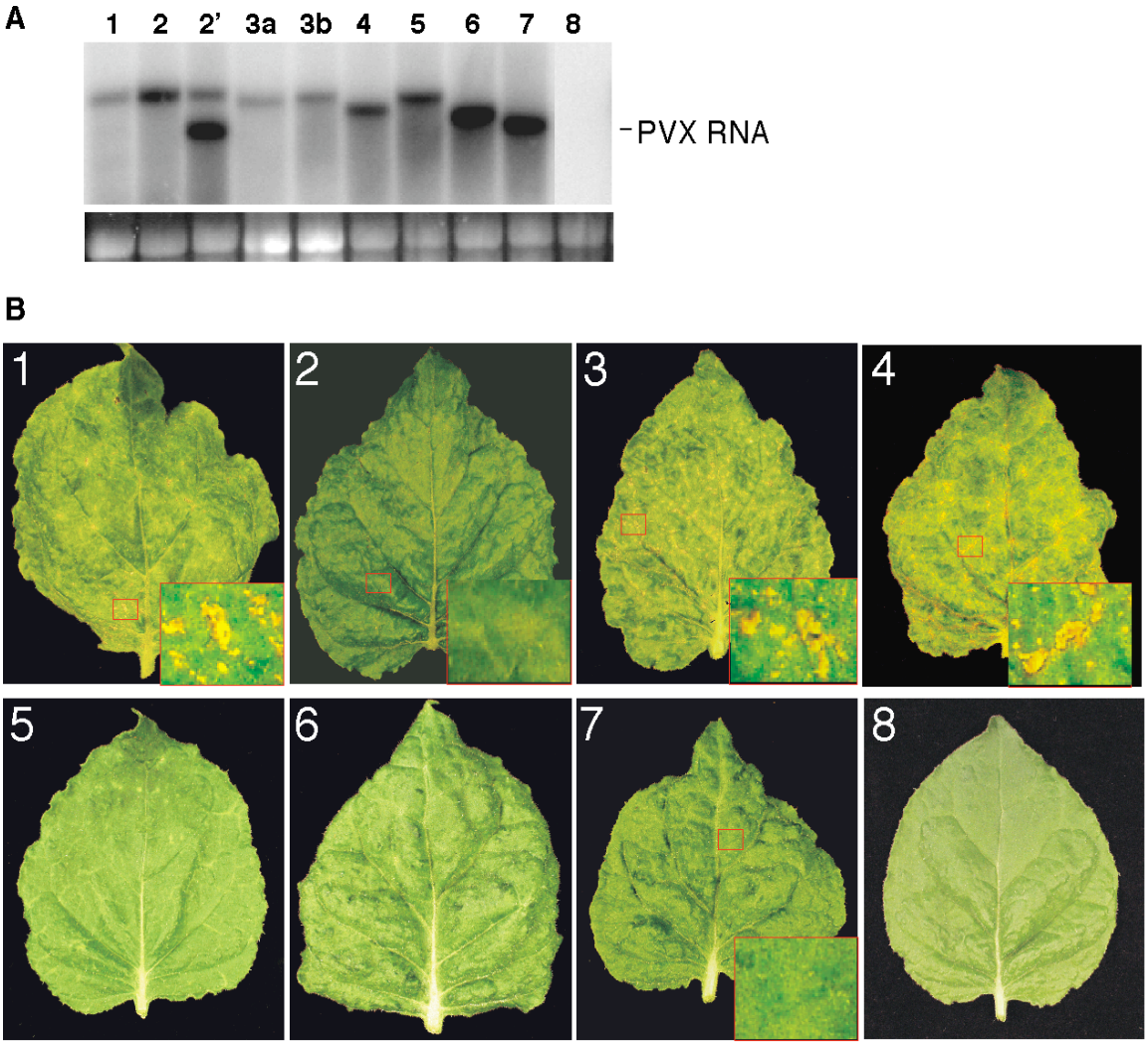

Figure 6.7. Expression of individual BPMV protein from a Potato virus X (PVX) vector. (A) Northern blot hybridization analysis to assess the size and level of progeny RNA accumulation in plants inoculated with transcripts from recombinant PVX constructs. Total RNA was extracted from the a fully expanded systemically infected leaf 14 d. p. i and probed with a PVX RNA specific probe. PVX RNA position is indicated to the right. The relative loading of the samples is indicated by ethidium bromide staining of ribosomal RNA (bottom panels). (B) Symptoms induced on *N. benthamiana* leaves inoculated with transcripts derived from recombinant PVX constructs. PVX and distinct PVX-derived constructs are indicated by different numbers, 1, PVX-Ha1-Hel/VPg; 2, PVX-Ho1-Hel/VPg-fm; 3a, PVX-Ho1-Hel; 3b, PVX-Ho1-Hel/VPg; 4, PVX-Ho1-Co-pro; 5, PVX-Ho1-LSCP, 6, PVX-Ho1-SCP; 7, PVX; 8, mock inoculation. Closeup images of photographs 1, 2, 3, 4 and 7 are shown in insets.

Chapter VII

Concluding Remarks

There are at least two distinct subgroups of *Bean pod mottle virus* (BPMV) strains that can be clearly distinguished based on nucleic acid hybridization and nucleotide sequencing analyses. Reassortants and partial diploid reassortants have also been identified among BPMV field isolates. The diversity of BPMV is more complex than I first predicated even though all the isolates I tested so far were collected from either the United States or Canada. It will important to decipher the genetic diversity of BPMV in broader geographical areas of the world, for instance, Brazil and China where soybean is an important economic crop. World trade not only benefits the soybean growers and consumers, but also facilitates the spread of seed-borne viruses. Furthermore, knowledge of the genetic diversity BPMV is critical for the development of soybean resistant germplasm with broad resistance to all known strains of BPMV.

To the best of my knowledge, the finding that stable diploid reassortants of BPMV can be isolated from natural populations of BPMV constitutes the first report in this regard. The phenomenon of diploid reassortment, however, has been experimentally documented for animal and plant viruses by using mixed inoculation (Kashiwazaki and Hibino, 1996; Fraile at al, 1997; Rodriguez et al, 1998). The discovery of diploid reassortants makes it prudent to analyze several single lesion isolates of potential new virus strains prior to making definitive identification. The stability of the partial diploid reassortants under field conditions should be investigated and their roles in virus epidemiology critically evaluated. A stable diploid reassortant may provide an excellent reservoir for generating new strains, provided that the beetle vectors efficiently transmit such strains.

Cross-protection among isolates of BPMV is demonstrated in this study. However, the mechanism is not clear. The results presented suggest that PTGS is the likely mechanism underlying cross-protection between BPMV strains. Direct evidence, however, is lacking. Characterization of the small RNAs (21-26 nt, indicators of PTGS) that might accumulate in the cross-protected plants would provide direct evidence for this

hypothesis. This could be further tested by challenge inoculation of protected tissue with transcript from a chimeric BPMV vector expressing a strong suppressor of PTGS (e.g., HC-Pro).

Symptom severity determinants map to the Co-pro and Hel. Furthermore, symptom severity is correlated with viral RNA accumulation level. By analogy to CPMV, the Hel and Co-pro induce extensive proliferation of ER membranes and vesicle formation thus enhancing the surface area utilized for virus replication. It is possible that the Hel and Co-pro encoded by the severe strains are more efficient in eliciting more membrane proliferation than those of the mild strains. The mechanism that specifies the extent of membrane proliferation, however, is not known.

References

Adams, S. 1988. Orlando lab puts fresh squeeze on Florida citrus problems. *Agric. Re.USA* 36, 6-10.

Ahoonmanesh, A. and Shalla, T. A. 1981. Feasibility of crossprotection for control of tomato mosaic virus in fresh market field-grown tomatoes. *Plant Dis.* 65,56-57.

Anandalakshmi, R., Pruss, G. J., Ge, X., Marathe, R., Smith, T. H., and Vance, V. B. 1998. A viral suppressor of gene silencing in plants. *Proc. Natl. Acad. Sci.* 95, 13079–13084.

Anjos, R. J., Jarlfors, U., and Ghabrial, S. A. 1992. Soybean mosaic potyvirus enhances the titer of two comoviruses in dually infected soybean plants. *Phytopathology* 82, 1022-1027.

Argos, P., Kamer, G., Nicklin, M. J., and Wimmer, E. 1984. Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families. *Nucleic Acids Res.* 12, 7251-7267.

Barry, J. B., Daniel. R. S., Laura, J. C. and David, H. L. B. 1985. Evolution of Bunyaviruses by genome reassortment in dually infected mosquitoes (Aedes triseriatus). *Science* 230, 548-550.

Baulcombe, D.C., Chapman, S.N., and Santa Cruz, S. 1995. Jellyfish green fluorescent protein as a reporter for virus infections. *Plant J.* **7**, 1045-1053.

Bawden, F. C., and Kassanis, B. 1945. The suppression of one plant virus by another. *Ann. Appl. Biol.* 32, 52-57.

Bertens, P., Wellink, J., Goldbach, R., and van Kammen, A. 2000. Mutational analysis of the cowpea mosaic virus movement protein. *Virology* 267, 199–208.

Brigneti, G., Voinnet, O., Wan-Xiang, L., Ding, S.W., and Baulcombe, D.C. 1998. Viral pathogenicity determinants are suppressors of transgene silencing. *EMBO J.* 17, 6739–6746.

Buck, K. W. 1996. Comparison of the replication of positive-stranded RNA viruses of plants and animals. *Adv. Virus Res.* 47**,** 159-251.

Calvert, L. A., and Ghabrial, S. A. 1983. Enhancement by soybean mosaic virus of bean pod mottle virus titer in doubly infected soybean. *Phytopathology* 73, 992-997.

Carette, J. E., Stuiver, M., Van Lent, J., Wellink, J., and Van Kammen, A. 2000. Cowpea mosaic virus infection induces a massive proliferation of endoplasmic reticulum but not Golgi membranes and is dependent on de novo membrane synthesis. *J. Virol.* 74, 6556-6563.

Carette, J. E., Guhl, K., Wellink, J., and van Kammen, A. 2002a. Coalescence of the sites of cowpea mosaic virus RNA replication into a cytopathic structure. *J. Virol.* 76, 6235-6243.

Carette, J. E., van Lent, J., MacFarlane, S. A., Wellink, J., and Van Kammen, A. 2002b. Cowpea mosaic virus 32- and 60-kilodalton replication proteins target and change the morphology of endoplasmic reticulum membranes. *J. Virol.* 76, 6293–6301.

Carette, J. E., Verver, J., Martens, J., van Kampen, T., Wellink, J., and van Kammen, A. 2002c. Characterization of plant proteins that interact with cowpea mosaic virus '60K' protein in the yeast two-hybrid system. *J. Gen. Virol.* 83**,** 885-893.

Carvalho, C. M., Wellink, J., Ribeiro, S. G., Goldbach, R. W., and van Lent, J. W. M. 2003. The C-terminal region of the movement protein of Cowpea mosaic virus is involved in binding to the large but not to the small coat protein. *J. Gen. Virol.* 84, 2271-2277.

Carvalho, C. M., Pouwels, J., van Lent, J. W. M., Bisseling, T., Goldbach, R. W., and Wellink, J. 2004. The Movement Protein of *Cowpea Mosaic Virus* Binds GTP and Single-Stranded Nucleic Acid In Vitro. *J. Virol.* 78, 1591-1594.

Chen, X., and Bruening, G. 1992a. Cloned DNA copies of cowpea severe mosaic virus genomic RNAs: infectious transcripts and complete nucleotide sequence of RNA 1. *Virology* 191, 607-618.

Chen, X., and Bruening, G. 1992b. Nucleotide sequence and genetic map of cowpea severe mosaic virus RNA 2 and comparisons with RNA 2 of other comoviruses. *Virology* 187, 682-692.

Chen, Z., Stauffacher, C., Li, Y., Schmidt, T., Bomu, W., Kamer, G., Shanks, M., Lomonossoff, G., and Johnson, J. E. 1989. Protein-RNA interactions in an icosahedral virus at 3.0 Å resolution. *Science* 245,154-159.

Cogoni, C., and Macino, G. 2000. Post-transcriptional gene silencing across kingdoms. *Genes Dev*. 10, 638-643.

Demski, J. W. 1969. Local reaction and cross protection for strains of squash mosaic virus. *Phytopathology* 59, 251-252.

Dessens, J. T., and Lomonossoff, G. P. 1991. Mutation analysis of the putative catalytic triad of the Cowpea mosaic virus 24K protease. *Virology* 184, 738-746.

De Zoeten, G. A., Assink, A. M., and A. Van Kammen. 1974**.** Association of cowpea mosaic virus-induced double-stranded RNA with a cytopathological structure in infected cells. *Virology* 59**,** 341-355.

Di, R., Purcell, V., Collins, G. B., and Ghabrial, S. A. 1996. Production of transgenic soybean lines expressing the bean pod mottle virus coat protein precursor gene. *Plant Cell Rep*. 15, 746-750.

Di, R., Hu, C-C., and Ghabrial, S. A. 1999. Complete nucleotide sequence of bean pod mottle virus RNA1: Sequence comparisons and evolutionary relationships to other comoviruses. *Virus Genes*. 18,129-137.

Eggen, R., Kaan, A., Goldbach, R., and Van Kammen, A. 1988. Cowpea mosaic virus RNA replication in crude membrane fractions from infected cowpea and *Chenopodium armaranticolor*. *J. Gen. Virol*. 69**,** 2711-2720.

Fehr, W. R., Caviness, C. E., Burmood, D. T., and Pennington, J. S. 1971. Stages of development descriptions for soybeans, *Glycine max* (L.) *Merr*. *Crop Sci*. 11, 929-931.

Fraile, A., Alonso-Prados, J. L., Aranda, M. A., Bernal, J. J., Malpica, J. M., and Garcia-Arenal, F. 1997. Genetic exchange by recombination or reassortment is infrequent in natural populations of a tripartite RNA plant virus. *J. Virol*. 71, 934-940.

Franssen, H., Leunissen, J., Goldbach, R., Lomonossoff, G., and Zimmern, D. 1984. Homologous sequences in non-structural proteins from cowpea mosaic virus and picornaviruses. *EMBO J*. 3, 855-861.

Fraser, R. S. S. 1998. Introduction to classical crossprotection. *In* "Methods in molecular biology, Plant virus protocols" (G.D. Foster and S. C. Taylor, Eds), vol. 81. p. 13-24. Humana Press, Totowa, NJ, U.S.A.

Fulton, R. W. 1980. Biological significance of multicoponent viruses. *Ann. Rev. Phytopathol.* 18, 131-196.

Gerber, M., and Sarkar, S. 1989. The coat protein of tobacco mosaic virus does not play a significant role for cross-protection. *J. Phytopathol.*124, 323-331.

Gergerich, R. C., and Scott, H. A. 1996. Comoviruses: Transmission, epidemiology, and control. In *The Plant Viruses*, vol. 5, *Polyhedral Virions and Bipartite RNA Genomes*, pp. 77-97. Edited by B. D. Harrison & A. F. Murant. NewYork: Plenum Press.

Ghabrial, S. A., Pickard, C. M., and Stuckey, R. E. 1977. Identification and distribution of virus diseases of soybean in Kentucky. *Plant Dis. Rep.* 61, 690-694.

Ghabrial, S. A., and Schultz, F. J. 1983. Serological detection of bean pod mottle virus in bean leaf beetles. *Phytopathology* 73, 480-483.

Ghabrial, S. A., Hershman, D. E., Johnson, D. W., and Yan, D. 1990. Distribution of bean pod mottle virus in soybeans in Kentucky. *Plant Dis.* 74, 132-134.

Goldbach, R., Rezelman, G., and van Kammen, A. 1980. Independent replication and expression of B-component RNA of cowpea mosaic virus. *Nature* 286,297-300.

Goldbach, R., Martelli, G. P., and Milne, R. G. 1995. Family Comoviridae. In: *Virus Taxonomy*, pp 341-347. Murphy, F.A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D. eds. Springer-Verlag, New York.

Goldbach, R. W., and Wellink, J. 1996. Comoviruses: molecular biology and replication. In *The Plant Viruses: Polyhedral Virions and Bipartite RNA Genomes*, pp. 35-76. Edited by B. D. Harrison & A. F. Murant. New York:Plenum.

Gorbalenya, A. E., Koonin, E. V., and Wolf, Y. I. 1990. A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. *FEBS Letters* 262, 145-148.

Goodin, M. M., Dietzgen, R. G., Schichnes, D., Ruzin, S., and Jackson, A. O. 2002. pGD vectors: versatile tools for the expression of green and red fluorescent protein fusions in agroinfiltrated plant leaves. *Plant J.* 31, 375-83.

Groregaoker, S. P., Eckhardt, L.G., and Culver, J. N. 2000. Tobacco mosaic virus replicase-mediate cross-protection: contribution of RNA and protein-derive mechanisms. *Virology* 273, 267-275.

Gu, H., Clark, A. J., de Sa, P. B., Pfeiffer, T. W., and Ghabrial, S. A. 2002. Diversity among isolates of *Bean pod mottle virus*. *Phytopathology* 92, 446-452.

Hamilton, A. J., and Baulcombe, D. C. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286, 950–952.

Hamilton, A., Voinnet, O., Chappell, L., and Baulcombe, D. 2002. Two classes of short interfering RNA in RNA silencing. *EMBO J*. 21, 4671–4679.

Han, S. S., Yoshida, K., Karasev, A. V. and Iwanami, T. 2002. Nucleotide sequence of a Japanese isolate of Squash mosaic virus. *Arch. Virol*. 147, 437-443.

Hartman, G. L., Sinclair, J. B., and Rupe, J. C. 1999. Bean pod mottle virus. Pages 61-62 in: *Compendium of soybean diseases*, 4th edition, American Phytopathological Society, St Paul, MN.

Haseloff, J., Siemering, K. R., Prasher, D. C., and Hodge, S. 1997. Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proc. Natl. Acad. Sci. USA 94,* 2122-2127.

Haudenshield, J. S., and Palukaitis, P. 1998. Diversity among isolates of squash mosaic virus. *J. Gen. Virol*. 79, 2331-2341.

Hendenson, W. W., Monroe, M. C., St. Jeor, S. C., Thayer, W.P., Rowe, J. E., Peters, C. J., and Nichol, S. T. 1995. Naturally occurring Sin Nombre virus genetic reassortants. *Virology* 214, 602–610.

Hu, C.–C., Sanger, M., Ghabrial, S. A. 1998. Production of infectious RNA transcripts from full-length cDNA clones representing two subgroups of peanut stunt virus strains: mapping satellite RNA support to RNA1. *J Gen Virol*. 79, 2013-21.

Hull, R. 2001. Direct passage in living high plant material. In **"***Matthews' plant virology"* (4th eds), pp547-554. Academic press.

Johnsonsen, L. K., and Carrington, J. C. 2001. Silencing on the Spot. Induction and suppression of RNA silencing in the Agrobacterium-mediated transient expression system. *Plant Physiology* 126, 930-938.

Kashiwazaki, S., and Hibino, H. 1996. Genomic reassortment of barley mild mosaic virus: evidence for the involvement of RNA1 in pathogenicity. *J Gen Virol*. 77, 581–585.

Kasteel, D., Wellink, J., Verver, J., van Lent, J., Goldbach, R., and van Kammen, A. 1993. The involvement of cowpea mosaic virus M RNA-encoded proteins in tubule formation. *J. Gen. Virol*. 74, 1721-1724.

Kasteel, D. T. J., Perbal, C.-M., Boyer, J.-C., Wellink, J., Goldbach, R. W., Maule, A. J., and van Lent, J. W. M. 1996. The movement proteins of cowpea mosaic virus and cauliflower mosaic virus induce tubular structures in plant and insect cells. *J. Gen. Virol*. 77, 2857-2864.

Krengiel, R., Vincente, A. C., Weyne, M., Shindo, N., Brioso, P. S., Felix, D. B., Villaroel, R., de Oliveira, D. E. and Timmerman, B. 1993. Molecular cloning and sequence analysis of a segment from Andean potato mottle virus B RNA encoding the putative RNA polymerase. *J. Gen. Virol*. 74, 315-318.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.

Lin, H-X., Rubio, R., Smythe, A. B., and Falk, B.W. 2004. Molecular population genetics of *Cucumber mosaic virus* in California: evidence for founder effects and reassortment. *J. Virol*. 78, 6666-6675.

Lomonossoff, G. P., and Shanks, M. 1983. The nucleotide sequence of cowpea mosaic virus B RNA. *EMBO J*. 2, 2253-2258.

Lomonossoff, G., Shanks, M., and Evans, D. 1985. The structure of cowpea mosaic virus replicative form RNA. *Virology* 144, 351-362.

Lomonosssoff, G. P. 2001. Comovirus. In *The Springer index of viruses*, pp 250-258. Edited by C. R.Tidona and G. Darai. Springer-Verklag Berlerg New York Press.

Lomonossoff, G. P., and Ghabrial, S. A. 2001. Comoviruses. In: *Encyclopedia of Plant Pathology,* Vol. 1. pp 239-242.O. C. Maloy, and T. D. Murray, eds. John Wiley & Sons, New York.

Lu, B., Stubbs, G., and Culver, J. N. 1998. Coat protein interactions involved in tobamovirus cross-protection. *Virology* 248, 188-198.

Lütcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F., and Scheele, G. A. 1987. Selection of AUG initiation codons differs in plants and animals. *EMBO J.* 6, 43-48.

MacFarlane, S. A., Shanks, M., Davies, J. W., Zlotnick, A., and Lomonossoff, G. P. 1991. Analysis of the nucleotide sequence of bean pod mottle virus middle component RNA. *Virology* 183:405-409.

Matthews, R. E. F. 1991. In"*Plant Virology*", pp. 145, 439-450, and 503-510. Academic Press, New York.

Memelink J, Swords, K. M. M., Staehelin, L. A., and Hoge, J. H. C. 1994. Southern, Northern and Western blot analysis. In: *Plant Molecular Biology Manual*, pp F1:1-23 2nd edition. S. B. Gelvin and R. A. Schilperoort, eds. Kluwer Academic Publishers, London.

Miranda, G. J., Azzam, O., and Shirako, Y. 2000. Comparison of Nucleotide Sequences between Northern and Southern Philippine Isolates of Rice Grassy Stunt Virus Indicates Occurrence of Natural Genetic Reassortment. *Virology* 266, 26-32.

Murphy, B. R., and Webster. R. G. 1990. Orthomyxoviruses, In B.N. Fields and D. M Kniper (ed) *Virology,* pp.1091-1152. Raven Press, New York, N. Y.

Pringle, C. R. 1996.Genetics and genome segment reassortment. In: *The Bunyaviridae*, pp. 189–226. Editor, Elliott, R. M. Plenum, New York.

Nelson, M. R., and Knuhtsen, H. K. 1973. Squash mosaic virus variability: review and serological comparisons of six biotypes. *Phytopathology* 63, 920-926.

Oxelfelt, P., Shanks, M., Widmark, A. K., and Lomonossoff, G. P. 1992. Identification and characterization of pseudo-recombinants of red clover mottle comovirus. *J. Gen. Virol.* 73, 2121-2124.

Palukaitis, P., and Zaitlin, M. 1984. A model to explain the cross-protection phenomenon shown by plant viruses and viroids. *In* ''Plant–microbe interactions: Molecular and genetic perspectives'', pp. 420. Eds., T. Kosuge and E. W. Nester. Macmillan, New York.

Peden, K. W. C., and Symons, R. H. 1973. Cucumber mosaic virus contains a functionally divided genome. *Virology* 155, 487-492.

Peters, S. A., Voorhorst, W. G., Wery, J., Wellink, J., and van Kammen, A. 1992. A regulatory role for the 32K protein in proteolytic processing of cowpea mosaic virus polyproteins. *Virology* 191, 81-89.

Peters, S. A., Verver, J., Nollen, E. A., van Lent, J. W., Wellink, J., and van Kammen, A. 1994**.** The NTP-binding motif in cowpea mosaic virus B polyprotein is essential for viral replication. *J. Gen. Virol.* 75, 3167-3176

Pitre, H. N., and Patel, V. C. 1975. Characteristics of bean pod mottle virus transmission by *Cerotoma trifurcata* (Forster). J. Miss. Acad. Sci. 19,186.

Posnette, A. F., and McA, T. J. 1955. Virus disease in West Africa. *Annu. Rev. Phytopathol.* 24, 355-381.

Pouwels, J., Carette, J. E., van Lent, J., and Wellink, J. 2002a**.** Cowpea mosaic virus: effect on host cell processes. *Mol. Plant Pathol.* 3, 411–418.

Pouwels, J., Van Der Krogt, G. N. M., van Lent, J., Bisseling, T., and Wellink, J. 2002b. The cytoskeleton and the secretory pathway are not involved in targeting the cowpea mosaic virus movement protein to the cell periphery. *Virology* 297,48-56.

Pouwels, J., Kornet, N., van Bers, N., Guighelaar, T., van Lent, J., Bisseling, T., and Wellink, J. 2003. Identification of distinct steps during tubule formation by the movement protein of Cowpea mosaic virus. *J. Gen. Virol.* 84, 3485 - 3494.

Powell-Abel, P.A., R.S. Nelson, Barun De, N. Hoffmann, S.G. Rogers, R.T. Fraley, and R.N. Beachy. 1986. Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science* 232, 738-743.

Qiu W, Park, J. W., and Scholthof, H. B. 2002. Tombusvirus P19-mediated suppression of virus-induced gene silencing is controlled by genetic and dosage features that influence pathogenicity. *Mol Plant Microbe Interact.* 15, 269–280.

Qu, F., and Morris, T. J. 2000. Cap-independent translational enhancement of turnip crinkle virus genomic and subgenomic RNAs. *J. Virol.* 74, 1085-1093.

Ratcliff, F. G., Harrison, B. D., and Baulcombe, D. C. 1997. A similarity between viral defense and gene silencing in plants. *Science* 276**,** 1558–1560.

Ratcliff, F. G., MacFarlane, S., and Baulcombe, D. C. 1999. Gene silencing without DNA: RNA-mediated cross-protection between viruses. *Plant Cell* 11, 1207-1215.

Reddy, M. S. S., Ghabrial, S. A., Redmond, C. T., Dinkins, R. D., and Collins, G. B.

2001. Resistance to the comovirus *Bean pod mottle virus* in transgenic soybean lines expressing the capsid polyprotein. *Phytopathology* 91, 831-838.

Rezelman, G., van Kammen, A., and Wellink, J. 1989. Expression of cowpea mosaic virus mRNA in cowpea protoplasts. *J. Gen. Virol.* 70**,** 3043 - 3050.

Ritzenthaler, C., Viry, M., Pinck, M., Margis, R., Fuchs, M., and Pinck, L. 1991. Complete nucleotide sequence and genetic organization of grapevine fanleaf nepovirus RNA1. *J. Gen. Virol.* 72, 2357-2365.

Rodriguez, L. L., Owens, J. H., Peters, C. J., and Nichol S.T. 1998. Genetic reassortment among viruses causing hantavirus pulmonary syndrome. *Virology* 242, 99–106.

Ross, J. P. 1963. Transmission of bean pod mottle virus in soybean by beetles. *Plant Dis. Rep.* 47, 1044-1050.

Ross, J. P. 1968. Effect of single and double infections of soybean mosaic and bean pod mottle virus on soybean yield and seed characters. *Plant Dis. Rep.* 52, 344-348.

Ross, J. P., and Butler, A. K. 1985. Distribution of bean pod mottle virus in soybeans in North Carolina. *Plant Dis.* 69, 101-103.

Rott, M. E., Tremaine, J. H., and Rochon, D. M. 1991. Comparison of the 5' and 3' termini of tomato ringspot virus RNA1 and RNA2: evidence for RNA recombination. *Virology* 185, 468-472.

Salaman, R. N. 1933. Protective inoculation against a plant virus. *Nature* 131, 468.

Sambrook, J., and Russell, D. W. 2001. Molecular Cloning: A Laboratory Manual. Third Edition, Cold Spring Harbor Laboratory Press, New York.

Schmitthenner, A. F., and Kmetz, K. 1980. Role of *Phomopsis* spp. in the soybean seed rot problem. Pages 355-366 in: *Proc. World Soybean Res. Conf. 2nd*.

Shanks, M., Stanley, J., and Lomonossoff, G. P. 1986. The primary structure of Red clover mottle virus middle component RNA. *Virology* 155, 697-706.

Shanks, M., and Lomonossoff, G. P. 1990. The primary structure of the 24K protease from red clover mottle virus: implications for the mode of action of comovirus proteases**.** *J. Gen. Virol.* 71, 735-738.

Shanks, M., and Lomonossoff, G. P. 1992. The nucleotide sequence of red clover mottle virus bottom component RNA. *J. Gen. Virol.* 73**,** 2473 - 2477.

Shank, M., and Lomonossoff, G. P. 2000. Co-expression of the capsid proteins of *Cowpea mosaic virus* in insect cells leads to the formation of virus-like particles. *J. Gen. Virol*. 81**,** 3093 - 3097.

Sherwood, J. L., and Fulton, R. W. 1982. The specific involvement of coat protein in tobacco mosaic virus (TMV) cross protection. *Virology* 119, 150–158.

Sherwood, J. L. 1987. Demonstration of the specific involvement of coat protein in tobacco mosaic virus (TMV) cross protection using a TMV coat protein mutant. *J. Phytopathol*. 118, 358–362.

Shindo, N., Vicente, A. C., Krengiel, R., and de Oliveira, D. E. 1993. Nucleotide sequence analysis of an Andean potato mottle virus middle component RNA cDNA clone: comparisons of the encoded proteins with those of other comoviruses. *Intervirology* 36,169-180.

Silva, M. S., Wellink, J., Goldbach, R. W., and van Lent, J. W. M. 2002. Phloem loading and unloading of *Cowpea mosaic virus* in *Vigna unguiculata*. *J. Gen. Virol*. 83, 1493-1504.

Simmonds, J. H. 1959. Mild strain protection as a means of reducing losses from the Queensland woodiness virus in the Passion vine. *Queensland J. Agric. Sci.* 16, 371-380.

Skotland, C. B. 1958. Bean pod mottle virus of soybeans. *Plant Dis. Rep*. 42, 1155-1156.

Stuckey, R. E., Ghabrial, S. A., and D. A. Reicosky, D. A. 1982. Increased incidence of *Phomopsis* spp. in seeds from soybean infected with bean pod mottle virus. *Plant Dis*. 66, 826-829.

Swofford, D. L. 2000. *PAUP*: Phylogenetic analysis using parsimony and other methods* (software). Sinauer Associates, Sunderland, MA.

Tumer, N. E., O'Connell, K. M., Nelson, R. S., Sanders, P. R., and Beachy, R. N. 1987. Expression of alfalfa mosaic virus coat protein gene confers cross-protection in transgenic tobacco and tomato plants. *EMBO J*. 6, 1181-1188.

Ushijima, H., Clerx-van Haaster, C. M., and Bishop, D. H. L. 1981. Analysis of Patois group bunyaviruses: evidence for naturally occurring recombinant bunyaviruses and existence of immune precipitable and nonprecipitable nonvirion proteins induced in bunyavirus-infected cells. *Virology* 110, 318–332.

Valkonen, J. P. T., Rajamäki, M-L., and Kekarainen, T. 2002. Mapping of viral genomic regions important in cross-protection between strains of a potyvirus. *Mol. Plant-Microbe Interact.* 15, 683-692.

van Bokhoven, H., van Lent, J. W., Custers, R., Vlak, J. M., Wellink, J., and van Kammen, A. 1992**.** Synthesis of the complete 200K polyprotein encoded by cowpea mosaic virus B-RNA in insect cells. *J. Gen. Virol.* 73, 2775-2784.

van Bokhoven, H., Le Gall, O., Kasteel, D., Verver, J., Wellink, J., and Van Kammen, A. B. 1993a. *Cis*- and *trans*-acting elements in cowpea mosaic virus RNA replication. *Virology* 195, 377–386.

van Bokhoven, H., Verver, J., Wellink, J., and van Kammen, A. 1993b. Protoplasts transiently expressing the 200K coding sequence of cowpea mosaic virus B-RNA support replication of M-RNA. *J. Gen. Virol.* 74, 2233-2241.

van Lent, J., Wellink, J., and Goldbach, R. 1990. Evidence for the involvement of the 58K and 48K proteins in intracellular movement of cowpea mosaic virus. *J. Gen. Virol.* 71, 219-223.

van Lent, J., Storms, M., Van Der Meer, F., Wellink, J., and Goldbach, R. 1991. Tubular structures involved in movement of cowpea mosaic virus are also formed in infected cowpea protoplasts. *J. Gen. Virol.* 72, 2615–2624.

Valverde, R. A., and Fulton, J. P. 1996. Comoviruses: identification and diseases caused. In *The Plant Viruses*, vol. 5, *Polyhedral Virions and BipartiteRNA Genomes*, pp. 17-33. Edited by B. D. Harrison & A. F. Murant. New York: Plenum Press.

Verver, J., Goldbach, R., Garcia, J. A., and Voc., P. 1987. In vitro expression of a full-length DNA copy of cowpea mosaic virus B-RNA: identification of the B-RNA encoded 24-kilodolaton protein as a viral protease. *EMBO J.* 6, 549-554.

Verwoerd, T. C., Dekker, B. M., Hoekema, A. 1989. A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Res.* 17, 2362.

Voinnet, O. 2001. RNA silencing as a plant immune system against viruses. *Trends Genet.* 17**,** 449–459.

Walkey, D. G. A. 1992. Studies on the control zucchini yellow mosaic virus in courgettes by mild strain protection. *Plan Pathol.* 41, 762-771.

Wellink, J., van Lent, J. W., Verver, J., Sijen, T., Goldbach, R. W., and van Kammen, A. 1993. The cowpea mosaic virus M RNA-encoded 48-kilodalton protein is responsible for induction of tubular structures in protoplasts. *J. Virol.* 67, 3660-3664.

Wellink, J., Verver, J., van Lent, J., and van Kammen, A. 1996. Capsid proteins of cowpea mosaic virus transiently expressed in protoplasts form virus like particles. *Virology* 224, 352-355.

Welink, J., Le Gall, O., Sanfacon, H. Ikegami, K., and Jones, A. T. 2000. Family *Comoviridae*. In: *Virus taxonomy*, Seventh ICTV report, pp, 691-701. Edited by van Regenmortel M. H. V., Fauquet, C. M., Bishop, D. H. L., Carstens, E. B., Estes, M. K., Lemon, S. M., Maniloff, J., Mayo, M. A., McGeoch, D. J., Pringle, C. R., Winckner, R. B. Academic Press, San Diego.

Wen, F., Lister, R. M., and Fattouh, F. A. 1991. Cross protection among strains of barley yellow dwarf virus. *J. Gen. Virol.* 72, 791-799.

Yeh, S. D., Wang, H. L., and Chiu, R. J. 1986. Control of papaya ring spot virus by seedling inoculation with a mild virus strain. In *Plant Virus Disease of Horticultural Crop in the Tropics and Subtropics,* pp. 169-178. Food and Fertilizer Techonology Center. Taiwan,

Zaitlin, M. 1976. Viral cross protection: More understanding is needed. *Phytopathol.* 66,382-383.

Zaumeyer, W. J., and Thomas, H. R. 1948. Pod mottle, a virus disease of beans. *J. Agr. Res.* 77, 81-96.

VITA

Name: Hongcang Gu

Date/Place of Birth: March 21, 1967

Shandong, China

Education

1989-1999: M.S, Shandong Agricultural University (China).
1984 –1989: Diploma, Shandong Medical School (China).

Professional Experience

1994-1999: Assistant Professor, Shandong Academy of Agricultural Sciences (China).
1992-1994: Research Assistant, Shandong Academy of Agricultural Sciences (China).
1987-1989: Clinical Technician, Shandong Cancer Hospital (China).

Awards and Honors

2003-2004: Dissertation Year Fellowship.
2002-2003: Kentucky Opportunity Fellowship.
2003: The American Society for Virology Travel Award.
2002: The American Society for Virology, Plant Virology Club Travel Award.
2002: University of Kentucky Graduate School Travel Award for Attending Annual meeting of American Phytopathological Society.

Publications

I. Papers in Refereed Journals

Gu, H., Clark, A. C., De Sa, P. B., Pfeiffer, T. W., Tolin, S., and Ghabrial, S. A. 2002. Diversity Among Isolates of *Bean pod mottle virus*. *Phytopathology*, 92:446-452.
Gu, H., Yan, D., Wang, J., Qiu, B., and Tian, B. 1996. The cloning and sequencing of coat protein gene from GFV. *Journal of Shandong Biotechnology*, 1:9-13.
Gu, H., Yan, D., Liu, H., Zhu, H., Qiu, B., and Tian, B. 1994. Detection of GFV by dot-blot hybridization with biotin-labeled GFV-cDNA probes. *Virologica Sinica*, 9:48-53.
Gu, H., Yan, D., Liu, H., Zhu, H., Qiu, B., and Tian, B. 1994. Detection of GFV by ELISA. *Journal of Shandong Agricultural University,* 25:82-86.

II. Book Chapter

Gu, H. 1994. New fungicides. Pages 53-73 in: New Pesticide Guide in Orchard. C. Yin, and Y. Zhang, eds. China Agricultural Publishing House.

III. Published Abstracts

Gu, H and Ghabrial, S. A 2003. Mapping symptoms determinant in Bean Pod Mottle Virus. The American Society for Virology 22st Annual Meeting.

Gu, H., Zhang, and C., Ghabrial, S. A. 2002. Molecular characterization of genetically distinct strains of *Bean pod mottle virus*. Phytopathology, 92:S32.

Gu, H., Zhang, C., and Ghabrial, S. A. 2002. The naturally-occurring reassortant strains of *Bean pod mottle virus* induce severe symptoms on their soybean host plants. The American Society for Virology 21st Annual Meeting Abstracts, P41-10.

Gu, H., Clark, A. C., P. B., Pfeiffer, T. W., and Ghabrial, S. A. 2000. Diversity and cross-protection among Isolates of *Bean pod mottle virus*. Phytopathology, 90:S31.

Membership in Professional Societies

The American Society for Virology.
The American Phytopathological Society.
Gamma Sigma Delta, the International Hon