Generation of an infectious *Beet mosaic virus* (BtMV) full-length clone based on the complete nucleotide sequence of a German isolate

Von dem Fachbereich Gartenbau

der Universität Hannover

zur Erlangung des akademischen Grades eines

Doktors der Gartenbauwissenschaften

- Dr. rer. hort. -

genehmigte Dissertation

von

M.Sc. Hana'a Hasan

geboren am 15.6.1970 in Syrien

Referent:	**Prof. Dr. Edgar Maiß**
Koreferent:	**Prof. Dr. Mark Varrelmann**
Tag der Promotion:	**13. Dezember 2004**

ABSTRACT

Beet mosaic virus (BtMV) is a member of the genus *Potyvirus* within the large and economically important family *Potyviridae*. BtMV occurs worldwide in major beet-growing areas, especially in temperate regions. The host range of BtMV includes all cultivated sugar beet and near relatives. BtMV infects mainly plants in the families *Chenopodiaceae*, *Solanaceae* and *Leguminosae*. It shows clearly visible mosaic disease symptoms on the leaves, whereas the infected plants are often of normal size. Damage and yield reduction due to BtMV infection has been reported for *Beta vulgaris*.

Limited information is available about its molecular properties and variability. The aim of this study was to determine the complete nucleotide sequence of a German isolate of BtMV (BtMV-G) and to compare the sequence with other potyvirus sequences. In addition an infectious full-length clone of BtMV-G was constructed in order to provide a possibility to study the virus multiplication cycle and to obtain an improved understanding of the molecular biology of potyviruses.

Ribonucleic acid was extracted from purified BtMV-G (DSMZ; PV-0065) or BtMV-G infected *Nicotiana benthamiana* plants and used as a template for cDNA synthesis. BtMV-specific oligonucleotides were designed and used together with a 26mer oligonucleotide, containing a random hexamer sequence at its 3'-end, for synthesis and amplification of cDNA fragments by reverse transcription-polymerase chain reaction (RT-PCR). The 5'-terminus of the genome was determined by reverse transcription of viral RNA using a specific primer, tailing of the cDNA with dGTP and PCR. All PCR fragments were cloned into the pGEM®-T Easy vector and subsequently the complete sequence of BtMV-G was determined. In addition, four cDNA clones generated by RT-PCR were used to assemble an infectious full-length clone of BtMV-G in a plasmid harbouring an enhanced *Cauliflower mosaic virus* 35S promoter.

The BtMV-G genome comprises 9592 nucleotides (nt) and contains one large open reading frame encoding a polyprotein of 3085 amino acid residues. The 5'- and 3'-untranslated regions were determined with 166 and 171 nt, respectively. Nine putative proteolytic cleavage sites were identified in the polyprotein resulting in ten mature proteins: P1, HC-Pro, P3, 6K1, CI, 6K2, NIa, VPg, NIb and CP, which are typical for all

members of the genus *Potyvirus*. Alignment of the predicted polyprotein sequence with a sequence of a BtMV isolate from the U.S.A. (BtMV-Wa) as well as with other potyviruses revealed amino acid sequence motifs typical of potyviruses. However, some motifs located in the HC-Pro, CI and NIb of BtMV-G contained different amino acids in comparison with other potyviruses. The highly conserved amino acid motif in the HC-Pro "Lys-Ile-Thr-Cys" involved in aphid transmission is diverged to the less common "Lys-Met-Ala-Cys" motif. Phylogenetic analysis clearly showed BtMV-G as a distinct member of the genus *Potyvirus*, sharing the highest amino acid sequence identity (55%) with *Peanut mottle virus* (PeMoV). The phylogenetic tree grouped BtMV-G, BtMV-Wa and PeMoV in one cluster located in the neighbourhood of the *Bean common mosaic virus* cluster.

The BtMV-G full-length clone leads to infectious virus in *N. benthamiana* after particle bombardment. Inoculated plants showed a delayed symptom development compared to the BtMV-G wild-type virus. Subsequent mechanical inoculation of *N. benthamiana* with BtMV-G generated from the full-length clone revealed indistinguishable symptoms from the wild-type virus. However, in *Atriplex hortensis* cv. 'Rheinische' BtMV-G generated from the full-length clone caused only small yellow blots on leaves compared with severe symptoms and stunting of the plants caused by the wild-type virus. In addition, BtMV-G from the infectious clone was not able to cause symptoms on some cultivars susceptible to the wild-type virus like *Spinacia oleracea* and *Beta vulgaris* (8T0015). It has still to be investigated, which genes of BtMV-G derived from the infectious full-length clone are involved in the different symptom expression.

The infectious cDNA clone of BtMV-G provides a powerful tool to study virus replication and could contribute towards a better understanding of the molecular biology of BtMV.

Keywords: Potyvirus, Beet mosaic virus, Infectious full-length clone

ZUSAMMENFASSUNG

Beet mosaic virus (BtMV) ist ein Mitglied des Genus *Potyvirus* aus der großen und wirtschaftlich wichtigen Familie *Potyviridae*. BtMV ist weltweit in den großen Rübenanbaugebieten vertreten, besonders in den gemäßigten Klimazonen. Der Wirtspflanzenkreis des BtMV umfasst alle kultivieren Rübenarten und nahe verwandte Pflanzen. BtMV infiziert hauptsächlich Pflanzen aus den Familien *Chenopodiaceae*, *Solanaceae* und *Leguminosae*. Das Virus ruft deutlich erkennbare Mosaiksymptome auf Blättern hervor, wobei die infizierten Pflanzen oftmals noch eine normale Größe aufweisen. Schäden und Ertragsverluste durch BtMV-Infektion wurden für *Beta vulgaris* beschrieben.

Es liegen nur begrenzte Informationen zu molekularen Eigenschaften und zur Variabilität des BtMV vor. Das Ziel der Arbeit war, die komplette Nukleotidsequenz eines deutschen Isolates des BtMV (BtMV-G) zu ermitteln und mit anderen Potyvirussequenzen zu vergleichen. Darüber sollte ein infektiöser Volllängenklon des BtMV-G erstellt werden, um so Möglichkeiten zu eröffnen, die Virusvermehrung zu untersuchen und ein verbessertes Verständnis der Molekularbiologie von Potyviren zu erhalten.

Ribonukleinsäure wurde aus BtMV-G (DSMZ; PV-0065) Reinigungen oder BtMV-G infizierten *Nicotiana benthamiana* Pflanzen extrahiert und als Template für cDNA-Synthesen eingesetzt. BtMV spezifische Oligonukleotide wurden zusammen mit einem 26mer Oligonukleotid, mit einer Zufallshexamersequenz am 3'-Ende, für die Synthese und Amplifikation von cDNA-Fragmenten mittels Reverser Transkription-Polymerase Ketten Reaktion (RT-PCR) verwendet. Der 5'-Terminus des Genoms wurde durch Reverse Transkription der viralen RNA mit einem spezifischen Primer, Tailing der cDNA mit dGTP und anschließender PCR erhalten. Sämtliche PCR-Fragmente wurden in den pGEM®-T Easy Vektor kloniert und anschließend wurde die vollständige Sequenz des BtMV-G ermittelt. Darüber hinaus wurden vier cDNA Klone durch RT-PCR erstellt und zu einem infektiösen Volllängenklon unter der Kontrolle eines verdoppelten *Cauliflower mosaic virus* 35S Promotors zusammengefügt.

Das BtMV-G Genom umfasst 9592 Nukleotide (nt) und besitzt einen großen offenen Leserahmen, der für ein Polyprotein mit 3085 Aminosäuren kodiert. Die 5'- und 3'-

nicht-translatierten Regionen wurden mit 166 nt bzw. 171 nt bestimmt. Neun Proteaseerkennungssequenzen wurden im Polyprotein identifiziert, so dass 10 für Potyviren typische Proteine gebildet werden können: P1, HC-Pro, P3, 6K1, CI, 6K2, NIa, VPg, NIb und CP. Durch Vergleich der Polyproteinsequenz mit einem BtMV Isolat aus den U.S.A. (BtMV-Wa) sowie mit anderen Potyviren wurden Potyvirus-typische Sequenzmotive lokalisiert. In einigen Motiven der Proteine HC-pro, CI und NIb des BtMV-G wurden Aminosäuresubstitutionen festgestellt. So besteht das hochkonservierte und an der Aphidenübertragung beteiligte Aminosäuremotiv „Lys-Ile-Thr-Cys“ im Fall des BtMV-G aus den Aminosäuren „Lys-Met-Ala-Cys“. Die phylogenetischen Analysen ordneten BtMV-G eindeutig als Spezies dem Genus *Potyvirus* zu, wobei die größte Aminosäureidentität (55%) mit dem *Peanut mottle virus* (PeMoV) bestand. Im Stammbaum wurden BtMV-G, BtMV-Wa und PeMoV einer phylogenetischen Gruppe zugeordnet, die eine enge Verwandtschaft zur *Bean common mosaic virus* Gruppe aufwies.

Vom Volllängenklon des BtMV-G wurden nach Partikelbombardment von *N. benthamiana* Pflanzen infektiöse Viren erhalten. Die inokulierten Pflanzen zeigten eine verzögerte Symptomentwicklung. Eine nachfolgende mechanische Übertragung auf *N. benthamiana* Pflanzen führte zu Symptomen, die nicht von Symptomen des BtMV-G Wildtyp-Virus zu unterscheiden waren. Allerdings verursachte BtMV-G, welches vom Volllängenklon abstammte, in *Atriplex hortensis* cv. ‘Rheinische’ nur kleine gelbliche und lokal begrenzte Chlorosen, während das Wildtyp-Virus zu schweren Symptomen verbunden mit einer sehr starken Stauchung der Pflanzen führte. BtMV-G, welches vom Volllängenklon abstammte, war nicht in der Lage Symptome auf einigen für das Wildtyp-Virus anfälligen Pflanzen wie *Spinacia oleracea* und *Beta vulgaris* (8T0015) hervorzurufen. Es bleibt noch zu untersuchen, welche Gene des vom Volllängenklon abstammenden BtMV-G in der unterschiedlichen Symptomausbildung beteiligt sind.

Der infektiöse Volllängenklon des BtMV-G eröffnet in Zukunft die Möglichkeit, die Virusreplikation zu untersuchen und könnte darüber hinaus zu einem verbesserten Verständnis der Molekularbiologie des BtMV beitragen.

Schlagworte: Potyvirus, Rübenmosaikvirus, Infektiöser Volllängenklon

TABLE OF CONTENTS

2. Materials and methods (continued)

2. Materials and methods (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF TABLES

LIST OF ABBREVIATIONS

aa	Amino acid
Abs	Absolute
AP	Ampicillin
AMV-RT	Avian Myelobastosis Virus Reverse Transcriptase
bp	Base pair
CI	Cylindrical inclusion protein
CP	Capsid protein
cDNA	Complementary deoxyribonucleic acid
dist	Distilled
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleotide triphosphate
dGTP	Deoxyguanine triphosphate
ds	Double strand
dsRNA	Double strand ribonucleic acid
DSMZ	German Collection of Microorganisms and Cell Cultures
E. coli	*Escherichia coli*
EDTA	Ethylene Diamine Tetracetic acid
ELISA	Enzyme-linked Immunosorbent Assay
EtOH	Ethyl alcohol
g	Gram
Fig	Figure
GSP	Gene specific primer
h	Hour
HC-Pro	Helper component protease
IC-RT-PCR	Immunocapture RT-PCR
IgG	Immunoglobulin G
IPTG	Isopropyl- β-D-thiogalactopyranoside
kDa	Kilo - Dalton
Kn	Kanamycin
LB	Lauria Bertoni Medium
L	Liter
LM-Agarose	Low Melting Agarose
mRNA	Messenger RNA
M	Molar
mM	Millimolar
ml	Milliliter
µg	Microgram
mg	milligram
min	minute
µl	Microliter
NaOAc	Sodium acetate
NIa	Nuclear inclusion protein a
NIb	Nuclear inclusion protein b

LIST OF ABBREVIATIONS (continued)

NTR	Non-translated region
nt(s)	Nucleotide(s)
ORF	Open Reading Frame
OD	Optical Density
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Round per minute
RTase	Reverse Transcriptase
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
RT – PCR	Reverse Transcription – PCR
SDS	Sodium Dodecylsulphate
sec	Second
ss	Single strand
Tab	Table
Tag	*Thermus aquaticus*
TE	Tris-EDTA
Tris	Tris (hydroxymethyl)-aminomethane
U	Unit
UN6	Random hexamer universal primer (oligonucleotides)
UN	Specific universal primer
W/V	Weight per volume
V	Volume
VPg	Viral protein genome-linked

Amino acids:

A	Alanine (Ala)
C	Cysteine (Cys)
D	Aspartic acid (Asp)
E	Glutamic acid (Glu)
F	Phenylalanine (Phe)
G	Glycine (Gly)
H	Histidine (His)
I	Isoleucine (Ile)
K	Lysine (Lys)
L	Leucine (Leu)
M	Methionine (Met)
N	Asparagine (Asn)
P	Proline (Pro)
Q	Glutamine (Gln)
R	Arginine (Arg)
S	Serine (Ser)
T	Threonine (Thr)
V	Valine (Val)

LIST OF ABBREVIATIONS (continued)

Amino acids:

W	Tryptophan (Trp)
Y	Tyrosine (Tyr)

Nucleotides:

A	Adenine
T	Thymine
C	Cytosine
G	Guanine
U	Uracil

Viruses:

BtMV-G	*Beet mosaic virus*- German isolate
BtMV-Wa	*Beet mosaic virus*- U.S.A isolate
CaMV	Cauliflower mosaic virus
LMV	*Lettuce mosaic virus*
LMoV	*Lily mottle virus*
MDMV	*Maize dwarf mosaic virus*
OYDV	*Onion yellow dwarf virus*
PLDMV	*Papaya leaf -distortion mosaic virus*
PRSV	*Papaya ringspot virus*
PSbMV	*Pea seed-borne mosaic virus*
PeMoV	*Peanut mottle virus*
PepMoV	*Pepper mottle virus*
PTV	*Peru tomato mosaic virus*
PPV	*Plum pox virus*
PVA	*Potato virus A*
PVV	*Potato virus Y*
PVY	*Potato virus Y*
ScaMV	*Scallion mosaic virus*
SrMV	*Sorghum mosaic virus*
SMV	*Soybean mosaic virus*
SCMV	*Sugarcane mosaic virus*
SPFMV	*Sweet potato feathery mottle virus*
TEV	*Tobacco etch virus*
TVMV	*Tobacco vein mottling virus*
TuMV	*Turnip mosaic virus*
WPMV	*Wild potato mosaic virus*
ZYMV	*Zucchini yellow mosaic virus*
LMV	*Lettuce mosaic virus*
LMoV	*Lily mottle virus*
MDMV	*Maize dwarf mosaic virus*
OYDV	*Onion yellow dwarf virus*
PLDMV	*Papaya leaf -distortion mosaic virus*
PRSV	*Papaya ringspot virus*

Viruses: (continued)

PSbMV	*Pea seed-borne mosaic virus*
PeMoV	*Peanut mottle virus*
PepMoV	*Pepper mottle virus*
PTV	*Peru tomato mosaic virus*
PPV	*Plum pox virus*
PVA	*Potato virus A*
PVV	*Potato virus V*
PVY	*Potato virus Y*
ScaMV	*Scallion mosaic virus*
SrMV	*Sorghum mosaic virus*
SMV	*Soybean mosaic virus*
SCMV	*Sugarcane mosaic virus*
SPFMV	*Sweet potato feathery mottle virus*
TEV	*Tobacco etch virus*
TVMV	*Tobacco vein mottling virus*
TuMV	*Turnip mosaic virus*
WPMV	*Wild potato mosaic virus*
ZYMV	*Zucchini yellow mosaic virus*

DECLARATION

I hereby declare that this work contained in this thesis has never been previously submitted for Ph.D. Degree or diploma at any other higher education institution. Conducted by me in the Institute of Plant Diseases and Plant Protection, Department of Molecular Plant Pathology / Virology, Hannover University. Furthermore, the thesis contains no previously published data or written by another person except where due reference is mentioned.

Hana'a Hasan

Date: 2.11.2004

Hannover University, Hannover, Germany

Eidesstattliche Erklärung

Hiermit versichere ich an Eides Statt, die vorliegende Arbeit selbstständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben, sowie die Arbeit noch nicht als Dissertation oder andere Prüfungsarbeit vorgelegt zu haben.

Hana'a Hasan

Hannover, den 2.11.2004

ACKNOWLEDGMENTS

I am most grateful to Prof. Dr. Edgar Maiss, whose supervision was a profound inspiration during my work, and I would like to thank him for his constant guidance, valuable suggestions and meticulous correction of the manuscript. Also, I deeply appreciate the molecular biology training received from him during the past three years, his patience, encouragement and his wise advices that go far beyond my study. I am so indebted to him there is no way to repay it.

I am grateful to Prof. Dr. Mark Varrelmann for co-refereeing the thesis.

Special thanks must also go to Dr. Stephan Winter and the virology group of the BBA for kind help in conducting virus purification.

Additionally, I would like to thank the technical assistant Ms. Jutta Zimmermann for her help in the technical part of the work.

Many thanks to all my colleagues of the laboratory for their immense support, cooperation and contribution in one way or another during the entire study period; namely Stephan Flatken, Reinhard Götz, Ilona Rolfes, Wulf Menzel, Rosana Blawid, Dennis Knierim, Jill Engelmann, Christof Dietrich and Dirk Stephan.

I deeply appreciate the financial support of this work that was awarded to me by the Islamic Development Bank as a long-term scholarship.

Finally, I extend my thanks to all my friends of the Arabic community in Germany for friendly behaviour and cooperation during the entire period of study in Germany.

Above all, my thanks to almighty God for his blessing.

Dedicated to my beloved family, and to

the special friend Mrs. Nouran Attar

CURRICULUM VITAE

Personal Data

Name	Hana'a Tawfik Hasan
Date of birth	15.6.1970
Place of birth	Hama / Syria
Nationality	Syrian

Education

1977-1982	Primary School, Hama / Syria
1983-1985	Elementary School, Hama / Syria
1986-1988	Secondary school, Hama / Syria
1989-1993	Bachelor of agriculture, Faculty of agriculture, Teshreen University, Lattakia / Syria
1994	Diploma, Department of Plant Protection, Faculty of agriculture, Teshreen University, Lattakia / Syria
1997-1999	Master of science, Department of Plant Protection, Faculty of agriculture, Aleppo University, Aleppo / Syria in cooperation with ICARDA
2001- 2004	Ph.D. student, Department of Molecular Plant pathology /Virology, Institute of Plant Diseases and Plant Protection, Hannover, University, Germany

Working

1994	Agricultural Directorate of Lattakia, Lattakia, Syria
1995/1996	Al-Ghab Agriculture Research Center, Hama/ Syria
2000/2001	

Conferences

1998	Plant Protection, American University, Lebanon
1999	Plant Protection, Agricultural Ministry, Damascus, Syria
2000	Plant Protection, Jordanian University, Jordan
2004	Plant Virology Meeting, Braunschweig, Germany
2004	Plant Pathology, Hamburg, Germany

1. Introduction

Potyviridae is the largest known family of plant viral pathogens. It contains 30% of all known plant viruses (Ward and Shukla, 1991). About 100 different members have been earlier described for this family (Edwardson, 1974; Hollings and Brunt 1981). Within this family, the genus-species concept is very well developed and proposed as the main system of classification (Ward and Shukla, 1991; Barnett, 1991; Barnett, 1992). Currently six genera have been established in the *Potyviridae*: *Bymovirus* with its type species *Barley yellow mosaic virus* (BaYMV), *Ipomovirus* with its type species *Sweet potato mild mottle virus* (SPMMV), Macluravirus with its type species *Maclura mosaic virus* (MacMV), *Rymovirus* with its type species *Ryegrass mosaic virus* (RGMV), *Tritimovirus* with its type species *Wheat streak mosaic virus* (WSMV) and *Potyvirus* with its type species *Potato virus Y* (PVY).

1. 1. The family *Potyviridae* and the genus *Potyvirus*

Potyvirus is the major genus in the *Potyviridae* with 91 species and 88 tentative species (Hull, 2002). This genus contains a large number of economically important viruses, e.g. PVY, *Bean yellow mosaic virus* (BYMV) and *Plum pox virus* (PPV). Potyviruses infect both dicotyledonous and monocotyledonous plants in many climatic regions. Most potyvirus species have a relatively narrow host range but their cumulative host range embraces a large number of plant species, including crop plants of economic importance from several families such as *Chenopodiaceae*, *Cucurbitaceae*, *Fabaceae*, *Poaceae* and *Solanaceae*. In addition potyviruses are frequently present as part of the natural pathogen spectrum causing chronic reductions in yield and quality (Hollings and Brunt 1981). Symptoms of potyvirus-infected plants depend on the virus, the host species and variety and can be observed on leaves, flowers, fruits, seeds, tubers and bulbs. The effects of potyvirus infections on a crop may vary from absence of any discernible injury or damage to complete crop failure. Damage is defined as the reduction in quantity or quality of the yield and loss is defined as the reduction in financial return due to damage. Potyvirus infections can reduce the quality of seed stock such as tubers in potato, stems, bulbs or seeds even if the infection in a current growing season does not affect yield. Losses can be

primarily due to a reduction of yield as in production of potatoes infected by PVY and melon infected by *Papaya ringspot virus* (PRSV). Some potyviruses cause cosmetic damages by decreasing crop value such as fruit distortion in cucumber by *Zucchini yellow mosaic viru*s (ZYMV) as reported in 1984 by Lisa and Lecoq, whereas the growth of potyvirus-infected plants is often retarded. The extent of losses depends on the proportion of infected plants, the moment at which the crop becomes infected, the virulence of the virus strain or isolate, the tolerance or resistance of the variety and also the climatic conditions. Therefore economic losses due to potyvirus infections can be significant in agricultural, pasture, horticultural and ornamental crops (Shukla and Ward, 1989a, 1989b).

Members of the *Potyviridae* can be transmitted by aphids (*Potyvirus, Macluravirus*), whiteflies (*Ipomovirus*), mites (*Tritimovirus, Rymovirus*), fungi (*Bymovirus*), seeds and pollen, mechanically or by dodder (Shukla *et al.*, 1991, 1994). Members of the genus *Potyvirus* are transmitted naturally by many aphid species in a non-persistent manner (non-circulative manner). This mode of transmission is characterized by short acquisition and transmission access periods. The absence of an incubation period in the vector is an important characteristic of this type of transmission. The aphid usually loses the ability to transmit the virus after the first or second probe or penetration of the plant cell by the stylet. Virus transmission by aphids is dependent upon two virus-encoded proteins: the coat protein (CP) and the helper component proteinase (HC-Pro). HC-Pro facilitates binding of virus particles to the aphid's maxillary stylets. Aphid specificity is also dependent upon the CP but the exact interaction of HC-Pro and CP in the potyvirus transmission process is not yet completely understood. The efficiency with which potyviruses are transmitted depends on aphid species, virus species and virus strain. Aphids can acquire the virus after only a brief feeding contact with an infected host plant. They usually retain the virus for less than an hour however a few virus particles retain infectivity up to 40 hours. This long retention period can explain, at least in part, the wide spread of these viruses as demonstrated by Zeyen and Berger (1990) for *Maize dwarf mosaic virus* (MDMV).

In addition, the spread of viruses is a process depending on the interaction between the host plant, the vectors, the virus sources and the environmental conditions. For example after an aphid deposits PVY in potato the virus enters the cell and the CP is removed. After

translation of the viral RNA, replicase-complexes are assembled and the viral RNA is copied many times. A primary infection occurs when the virus moves into adjacent cells and the vascular system to ascend into the upper regions of the plant without causing serious harm. Virus particles can remain in the upper regions throughout the vegetative cycle of potato. Upon completion of the plant cycle the virus descends to the tubers for hibernation. Tubers harboring PVY are harvested, transported and perhaps grown in other places. Once the tubers germinate the virus is activated and secondary infection takes place throughout the whole plant. This secondary infection causes more physical damage to the plant and the virus is now capable of being transmitted to other healthy plants by its vectors.

1. 2. Criteria to place virus species in the *Potyviridae*

The symptomatology and host range played a significant role in the delineation of potyviruses and their strains in the past and even today they remain the first phenotypic criteria for recognition of strains. Since most potyviruses have restricted host ranges, they can often be identified on the basis of the characteristic symptoms they produce in certain hosts. Host range and symptomology can play a significant role in the identification and classification of potyviruses, if the biochemical properties of the viruses are established first and then efforts are made to find differential hosts which can distinguish one virus from another similar virus. This approach was used successfully to differentiate four potyviruses infecting *Gramineae* on the basis of their reactions on selected sorghum inbreds. However, reliance on this criteria has created a lot of confusion in the identification of potyviruses infecting members of some plant families such as *Leguminosae* (Bos, 1970), *Gramineae* (Shukla *et al*., 1989b) and *Solanaceae* (Blanko-Urgoiti *et al*., 1996). Different potyviruses caused similar symptoms in some hosts, while different climatic conditions and different genotypes of the same plants species can have profound effects on the susceptibility and symptoms of potyvirus infections, which make it often difficult to compare results obtained in different laboratories (Bos, 1970). The host reactions are still used for the recognition of strains in spite of considered as not very precise and reliable (Hollings and Brunt, 1981). Recent data suggest that the symptomatology may not be a reliable marker of genetic relatedness. Data showed that a

single point mutation in the CP gene of *Tobacco mosaic virus* (TMV) completely changed the symptom phenotype from a systemic mosaic to appearance of local lesions (Knorr and Dawson, 1988). A similar information has also been reported for *Cucumber mosaic virus* (CMV) (Shintaku and Palukaitis, 1989).

All potyvirus members share a property, which is considered as another phenotypic criterion for assigning viruses to the *Potyviridae* (Shukla *et al*., 1989). It is the ability to induce the formation of characteristic "pinwheel" cylindrical inclusions (CI) in the cytoplasm of infected cells (Edwardson, 1974). According to the nature of the aggregates attached to the central protein of the characteristic potyvirus pinwheel inclusion (Edwardson and Christie, 1984) the viruses were assigned to subgroups. However, difficulties exist in the identification and application of the correct subgroup, especially if mixed infections occur. If these subgroups are used to establish hierarchical classifications within the potyvirus group, the subgroups do not appear to correlate with other accepted criteria. Lesemann (1988) suggested that the capacity to form cylindrical inclusions reflects conservatism of the CI coding region of the viral genome. This could be extended to imply that within each CI subdivision there would be higher sequence identity than between subdivisions. However, the question of the value of cytoplasmic inclusions as taxonomic markers remains open.

Serological tests are useful for establishing the identity of a new virus strain and for the quantitation of the degree of relationships among distinct members of a genus (Hollings and Brunt, 1981). Double antibody sandwich ELISA (DAS-ELISA) was used to differentiate between strains of the same virus (Clark and Adams, 1977). However, serological relationships among distinct members of potyviruses are extremely complex and sometimes unreliable (Shukla *et al*., 1992a). Hollings and Brunt (1981) suggested that there is no simple pattern of antigenic or consistent paired relationship among members in the family and serological relationships do not correlate with biological properties. Shukla *et al.,* (1992a) found that *Bean yellow mosaic virus* (BYMV) was serologically related to *Lettuce mosaic virus* (LMV) and *Bean common mosaic virus* (BCMV), but they did not observe any serological relationship between LMV and BCMV. Moreover, different strains of one potyvirus may be as serologically distantly related to strains of other potyviruses as

to one another, implying that the species and strain concept has limitations when applied to potyviruses.

Cross-protection was also used successfully for classification of some potyviruses since it was accepted that one strain of a virus can protect against the infection by another strain of the same virus (Shukla *et al*., 1991). Bos (1970) indicated that distinct strains of BYMV protected completely against one another, while strains of MDMV were not able to protect against one another (Paulsen and Sill, 1970). The former strains of MDMV are now regarded as distinct potyvirus species (Shukla *et al.*, 1992b). Therefore, cross-protection can be used for taxonomic purposes, but it should kept in mind that there are also limitations of this technique.

Thus, all the traditional criteria of symptomatology, host range, cytopathology, serology and even cross protection failed to provide consistent tools for taxonomy of potyviruses. However, particle morphology and molecular data like the nucleotide composition of the 3'- non translated region (NTR), used by Frenkel *et al.* (1989) and Habera *et al.* (1994) for virus classification, the amino acid sequence of the CP and - if available - additional molecular data concerning the entire genomic sequence are the most useful criteria, which are now widely used for assigning a virus to a genus and for differentiating closely related viruses (van Regenmortel *et al.*, 1997). These criteria can also be used to distinguish viruses from strains and to establish evolutionary relationships between groups of distinct potyviruses (Shukla and Ward, 1989b; Shukla and Ward, 1989c). This led to the revised classification of some viruses and strains and to the clarification of previously conflicting and inconsistent biological properties; it also provided a sound basis for subgrouping of potyviruses. Structure and sequence analyses have revealed why serological data can be confusing. Shukla and Ward (1988, 1989b) showed that the N- and C-terminal regions of the potyvirus CP are surface located and that the N-terminus is the immunodominant part of the virus particle. In addition potyvirus CPs were found to consist of a conserved core sequence, which elicits highly cross-reactive antibodies. Virus specific epitopes are located in the strongly immunogenic N-terminal region which tends to vary among most viruses and which elicits potentially discriminatory antibodies, while potyvirus group-specific epitopes reside in the core region of the coat protein. A complication is that the N-terminal

sequence of the coat protein is readily lost by proteolysis when sap extracts are prepared although the virus particles remain intact. Nevertheless, when antibodies are appropriately raised and when serological tests are well conducted the results do help to differentiate between species of potyviruses.

Analysis of the 136 possible pairings of the complete CP amino acid sequences from 17 strains of 8 distinct potyviruses revealed a bimodal distribution of sequence homology (Shukla and Ward, 1989c). In this analysis the sequence homology between distinct members ranged from 38 to 71% (average 54%) while that between strains of one virus ranged from 90 to 99% (average 94,5%). The evaluation of sequence data resolved the confusion concerning the identity of viruses in the potyvirus group, especially which appear similar to each other in host range and symptomatology. The sequence data confirmed that PPV-R is a strain of PPV and that PVY-1 and PVY-N are strains of PVY. In addition the data of four BYMV isolates revealed that three isolates have a high sequence identity (88 to 90%) and are true strains of BYMV, while BYMV-30 showed a much lower sequence identity (70 to 73%), suggesting that it is a distinct member of the potyvirus group. The sequence data for the MDMV and *Sugarcane mosaic virus* (SCMV) isolates resolved the problems associated with the large group of viruses known to infect maize, sugar cane, sorghum, and other plant species throughout the world (Pirone, 1972). The CP is the only virus product, which shows no significant sequence identity with the corresponding protein of other virus groups. Finally, the recent molecular data did not change the potyvirus group status of individual viruses apart from confirming the membership of the non-aphid transmitted viruses and underpinning the development of a group-specific serology. In contrast the molecular studies had a dramatic effect on the development of criteria to distinguish potyviruses from potyvirus strains. Complete genome sequences of many potyviruses (38 viruses) are available and there are more than 220 potyvirus coat protein sequences available in the GenBank database. Extensive sequence comparisons have shown that the application of such quantitative taxonomy to all members of the family *Potyviridae* results in a clear-cut distinction between each of the different taxonomic levels: strains, species and genera (Shukla *et al.,* 1994). Moreover, other characters can be deduced from the sequence data. The expression of the potyvirus RNA genome leads to the synthesis of a polyprotein, which is digested at specific sites by different virus-encoded proteases to form

the mature virus proteins. Comparison between the amino acid sequences at the cleavage sites in the polyproteins of two viruses can thus be a sensitive measure of similarity between the respective protease (van Regenmortel *et al.*, 1997). It is clear that not a solely criterion has an absolute supremacy over others, some are more informative and discriminate better or they are easier to acquire, but it is the sum of the information accumulated that has built up a clear and generally accepted taxonomy for potyviruses.

1.3. Potyvirus genome structure and expression

The genus *Potyvirus* counts currently for approximately 180 members making it the largest recognized group of plant viruses. It has one of the widest host ranges of all virus groups. Potyviruses are quite similar, in terms of their genomic structure and strategy of expression, to the plant bipartite como- and nepoviruses and to the animal picornaviruses. In addition, the genome of these differing species contains a region of conserved gene order that encodes non-structural proteins involved in RNA replication. For these reasons it has been proposed that the como-, nepo- and potyviruses may be arranged in one supergroup of picorna-like plant viruses (Goldbach, 1986; Goldbach, 1992). All members of the genus potyvirus share common features especially of the virus particles or virions. Potyviruses virions are flexuous, non-enveloped (Langenberg and Zhang, 1997) filamentous particles 680-900 nm in length and 11-15 nm in diameter (Dougherty and Carrington, 1988) composed of RNA and about 2000 copies of CP units (Martin and Gelie, 1997). Previous studies indicated that all members of the genus *Potyvirus* have a similar genome structure and organization but differ somewhat in detail.

The genetic information is contained in a positive sense single-stranded RNA (+ssRNA) approximately 10 kilobases (kb) in length (Allison *et al.*, 1986; Brakke and van Pelt, 1970; Domier *et al.*, 1986). The RNA genome carries a VPg (genome-linked viral protein) covalently attached to the 5'-end (Riechmann *et al.*, 1989; Hari, 1981; Murphy *et al.*, 1990) and a poly (A) tail at the 3'-end (Hari *et al.*, 1979). The positive sense genome can act directly as a messenger RNA, with the 5' non-coding region functioning as an enhancer of translation (Carrington and Freed, 1990). The genome contains a long single open reading frame (ORF), which translates into a large 340-370 kilodalton (kDa) polyprotein (Urcuqui-

Inchima *et al.*, 2000; Riechman *et al.*, 1992; Riechmann *et al.*, 1995) which is proteolytically processed into precursor proteins or the functional proteins. To date (30.12.2003) 35 entire genomes of potyviruses have been sequenced completely and published in the GenBank, all these sequenced viruses were assigned to the genus *Potyvirus* according to the CP amino acid sequence and other genome sequence data. The genetic map of most potyviruses indicated that the virus encoded proteins are P1 protein, helper component proteinase (HC-Pro), P3 protein, cylindrical inclusion protein (CI), nuclear inclusion protein a (NIa), viral genome-linked protein (VPg), nuclear inclusion protein b (NIb) and the capsid protein (CP), as well as two small putative proteins named as 6K1 and 6K2 (Fig. 1.). Within the potyvirus genome there are conserved and variable regions. The most conserved regions incorporate the HC-Pro and the NIb, while the most variable regions consist of P1, P3 and the CP (Aleman-Verdaguer *et al.*, 1997).

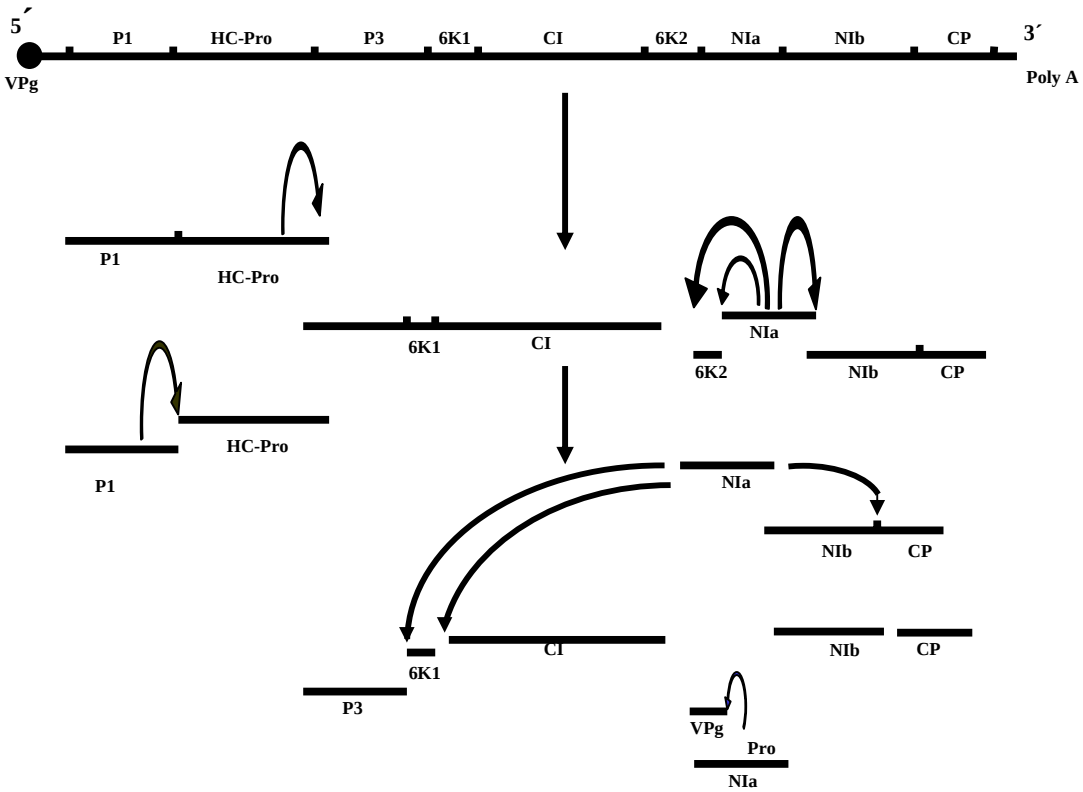

Figure 1: Genome organization and gene expression of the genus *Potyvirus*.

Proteolysis of the potyvirus polyprotein reveals ten mature proteins by three virus-encoded proteinases: the N-terminal protein P1 (Verchot *et al.*, 1991), the HC-Pro (Carrington *et al.*, 1989b) and the nuclear inclusion protein a (NIa-Pro) (Ghabrial *et al.*, 1990; Garcia *et al.*, 1989; Hellman *et al.*, 1988; Chang *et al.*, 1988). The 35-kDa P1 is cleaved from the P1-HC-Pro product at a Phe-Ser junction by P1 which contains a serine at the protease active site in its C-terminal region. This protease has proved to be difficult to analyze, as it does not appear functional in a rabbit reticulocyte lysate in vitro translation system (Riechmann *et al.*, 1992; Shukla *et al.*, 1994). The cleavage site was identified for some potyviruses like *Tobacco vein mottling virus* (TVMV) and *Tobacco etch virus* (TEV). The enzymatic activity responsible for the HC-Pro-P3 cleavage site is in a 20-kDa domain in the C-terminal half of the HC-Pro protein (Carrington *et al.*, 1989a), releasing the 52-kDa HC-Pro by cleavage at a C-terminal Gly-Gly junction. This proteinase has cystein and histidine residues at its active site.

The NIa has a two domain structure, the N-terminal genome-linked protein VPg (22 kDa) and the C-terminal protease (27 kDa). The 27 kDa protease domain of the nuclear inclusion body a (NIa) is a serine protease responsible for cleavage at most, if not all, Gln-(Ser/Gly) junctions. This protease is the major proteinase of potyviruses (Dougherty and Carrington, 1988). NIa is autocatalytically released from the polyprotein and then catalyses the cleavage of the various junctions to release the proteins P3, 6K1, cylindrical inclusion body (CI), 6K2, nuclear inclusion body b (NIb) and CP. NIa autoproteolysis results in the VPg but the cleavage of the VPg domain from the NIa protein is much less efficient (Fig. 1.). The cleavage sites have been identified for several viruses and the primary sites are considered to be QS, QG, and QA with the surrounding motif V-X-X-Q/(A, S, G or V) thought to be common to all potyvirus. However, there are some variations of this motif and especially at the site releasing the VPg from the NIa protein, which obviously differs with E/(S, A, G) residues from most of the others. There is no full information on the sequential order of the processing events. For instance the cleavage at the C-terminus of HC-Pro takes place, giving P1-HC-Pro, before P1 is cleaved off (Carrington *et al.*, 1989b). The requirement for this processing represents a control mechanism that is not yet understood. The processing and function of the potyviral proteins is still controversial, but it is has be shown, e.g. for the CP, that they are multifunctional.

1.3.1. P1 protein

This protein is encoded by the 5'-terminal gene of the potyvirus genome and encodes a proteinase, which is responsible for the cleavage of P1 from the polyprotein (Riechmann *et al.*, 1992). Among different potyviruses, the P1 protein is the least conserved region of the entire polyprotein (Domier *et al*., 1987; Vance *et al*., 1992) except for conserved amino acids found at the C-terminal part of all potyvirus P1 proteins and which correspond to the proteinase catalytic domain. In fact, the N-terminal half of P1 is hypervariable both in length and in sequence. A high degree of polymorphism in the P1 coding region was observed for PVY and ZYMV isolates. Deletion and mutational analyses have shown that P1 is not strictly required for viral infectivity, even though it enhances amplification and movement of the virus. On the contrary, cleavage at the boundary P1 and HC-Pro is essential for viability. Two TVMV P1 insertion mutants were produced (Klein *et al.*, 1994). Insertion in the C-terminus between the conserved Asp and Ser of the catalytic triad was lethal, whereas insertion in the N-terminus had no effect on virus viability. Deletion and mutations that did not affect the proteinase activity of P1 were also viable in case of TEV. The mutations in the proteolytic domain were lethal but could be rescued by replacement of the P1/HC-Pro cleavage site by using the NIa specific cleavage sequence (Verchot and Carrington, 1995a). Verchot and Carrington (1995b) pointed out that P1 acts in trans and stimulates genome amplification because a TEV ß-glucuronidase (GUS) mutant with the entire P1 gene deleted was able to replicate at low levels and to move from cell to cell and through the plant at reduced rates. All these studies indicated that the P1 proteinase activity is not required in itself, and also that processing between P1 and HC-Pro can be delayed in time until expression of NIa occurs.

The P1 of TVMV (Brantley and Hunt, 1993), *Turnip mosaic virus* (TuMV; Soumounou and Laliberte, 1994) and *Potato virus A* (PVA; Merits *et al.*, 1998) was shown to exhibit nonspecific RNA-binding activity, although the domain required for this function is not known. The precise function of P1 in viral infection has yet to be established. Non-specific RNA binding has often been attributed to involvement of the protein in viral movement. Localization of PVY P1 with cytoplasmic inclusion bodies involved in virus cell to cell movement could be in agreement with such a function (Arbatova *et al*., 1998).

1.3.2. Helper component proteinase (HC-Pro)

Extensive researches were carried out to identify the helper component protein and to determine its functions. Very recently, Plisson *et al.* (2003) suggested that the helper component proteinase (HC-Pro) is a key protein encoded by plant viruses of the genus Potyvirus and it is involved in different steps of the viral cycle. The helper component proteinase is a multifunctional protein with important roles in post-transcriptional gene silencing (PTGS) suppression (Llave *et al*., 2002), polyprotein processing (Carrington *et al*., 1989b) and aphid transmission (Maia *et al*., 1996; Pirone and Blanc, 1996; Flasinski and Cassidy, 1998). Recent studies supported the hypothesis that it functions in virus cell-to-cell and systemic movement and genome amplification (Kasschau *et al*., 1997; Klein *et al*., 1994). HC-Pro also influences the symptom severity and accumulation of heterologous viruses, probably as an indirect consequence of PTGS suppression (Kasschau and Carrington, 2001). Some mutations showed that HC-Pro plays an important role in virus life cycle. Mutations H23G, C25G and C53G of TVMV HC-Pro were lethal as some deletions in the PPV HC-Pro sequence (Guo *et al.*, 1998). In contrast Dolja *et al.* (1993) found that similar deletions in the TEV HC-Pro were not lethal.

Structural knowledge of HC-Pro played an important role for a deeper understanding of its multiple functions. HC-Pro activity is exchangeable among potyvirus species, also with different degrees of specificity (Lecoq and Pitrat, 1985; Pirone, 1981; Sako and Ogata, 1981). However, HC-Pro expressed in bacteria or in insect cells is unable to support aphid transmission, but purified biologically active HC-Pro was obtained by expression from a *Potato virus X* (PVX) vector in plants (Sasaya *et al.,* 2000). Studies suggested that HC-Pro mediates retention of virions at specific aphid mouthpart sites and that virions are subsequently released into plant cells (Maia *et al*., 1996; Pirone and Blanc, 1996). Blanc *et al.,* (1997) suggested that HC-Pro may act as a bridge between aphid mouthparts and virus particles facilitating retention in the mouthparts. HC-pro is probably active as a dimer (Thornbury *et al*., 1985) and as such it would also have a dimerization domain. HC-Pro binds CP monomers through a triad of amino acids: Pro-Thr-Lys (PTK) located near the C-terminus of HC-Pro (Peng *et al*., 1998). An amino acid motif consisting of Lys-Ile-Thr-Cys (KITC) at the N-terminal region of the HC-Pro is critical for virus retention in the stylet of

the aphid vector (Blanc *et al.*, 1998) Amino acid substitutions that affect any of these both motifs result in the loss of aphid transmission (Dolja *et al.*, 1993).

The N-terminus of the HC-Pro contains a cysteine (Cys)-rich region in which also the KITC motif is located. This Cys-rich domain could form a zinc finger-like motif, which may play structural and functional roles (Robaglia *et al.*, 1989) similar to those reported for many zinc-finger proteins (Berg, 1990). Mutations in the Cys-rich region of HC-Pro of some potyviruses lead to deleterious effects on viral infectivity and aphid transmissibility (Urcuqui-Inchima *et al.*, 1999). Additional substitutions involving other amino acids in this region also render HC-Pro incapable of mediating aphid transmission in TuMV, PVY and TEV (Blanc *et al.*, 1998; Canto *et al.*, 1995; Nakashima *et al.*, 1993). However, there is no information whether particular amino acid residues within the Cys-rich domain play similar roles in all potyviruses. Dolja *et al.*, (1993) reported that the N-terminal amino acid residues of TEV HC-Pro, which include the Cys-rich domain, are not necessary for systemic infection as deletion mutants lacking this region can move systemically at the same rate as wild-type virus. The N-terminus HC-Pro amino acid sequence is variable across different members of the potyviruses, however, in spite of this characteristic variability, the Cys-rich region presents a number of amino acid positions occupied by either highly conserved residues or residues with similar chemical properties. Llave *et al.* (2002) confirmed that the Cys-rich domain play a highly conserved role in aphid transmission among potyviruses. It was confirmed that HC-Pro of PVY was present in the food canal by using immunogold labeling and transmission electron microscopy. Flasinski and Cassidy (1998) indicated that the highly conserved motif (CCC) of HC-Pro of *Peanut stripe virus* (PStV) could play an important role in the specific interaction between the HC-Pro and the coat protein (DAG) motif. Additionally, mutations in this motif reduced genome amplification and movement of the virus, respectively (Cronin *et al.*, 1995).

1.3.3. P3 protein

The P3 protein is not as well characterized as other potyvirus proteins. Its size ranges from 38 to 50 kDa (Goldbach, 1992; Robaglia *et al.*, 1989). Most of the studies indicated that P3 is a very variable protein of potyviruses and displays only a low similarity between species

(Shukla *et al.*, 1991). Despite the variation observed between P3 proteins of distinct potyviruses, a few differences were noticed within the P3 of *Yam mosaic virus* (YMV) strains (Alleman-Verdaguer *et al*., 1997). Very little is known about the real function(s) of the P3 protein. Some studies suggested that P3 must play an important role, since it was conserved between strains. Most of the studies tried to assign its function by identifying the P3 localisation in the plant cell. Rodriguez-Cerezo and Shaw (1991) showed the presence of TVMV P3 protein in the membrane fraction of virus-infected cells. With immunogold labelling techniques, Rodriguez-Cerezo *et al*., (1993) determined the subcellular location of TVMV P3 in infected *Nicotiana tabacum* cells. They found the P3-specific gold label almost exclusively associated with the cylindrical inclusions, typically formed in the cytoplasm of potyvirus infected cells. However, similar experiments were conducted by Langenberg and Zhang (1997), who used TEV to determine the location of P3. The result showed that gold labels were associated with nuclear inclusions. Despite these different results, it can be concluded that P3 may be involved in virus amplification, since it was found associated with membranes and proteins involved in RNA replication (CI, NIb, VPg). Discrepancies between localization of potyviral proteins may result from different localizations of the functional protein compared to the non-functional forms, which occur in large excess (Riedel *et al.*, 1998). P3 has no RNA binding activity (Merits *et al*., 1998) suggesting that involvement of P3 in virus replication occurs through its interaction with CI (Rodriguez-Cerezo *et al.*, 1993), which is a part of the replication complex. The hypothesis of the involvement of P3 in virus replication is strongly supported by the discovery that three random insertional mutations located in the P3 of TVMV were unable to support viral replication in plants and protoplasts (Klein *et al.*, 1994).

Other functions were suggested for P3 depending on some indicators. An involvement of P3 in early stages of viral replication, due to the presence of P3 in nucleoli and nuclear inclusions has been postulated (Langenberg and Zhang, 1997). Riechmann *et al.* (1992) also suggested that P3 acts as a proteinase cofactor. Dougherty and Semler (1993) proposed that P3 is a movement protein. However, many studies supported the role of P3 in plant pathogenicity and Riechmann *et al.* (1995) confirmed that mutations in P3 of PPV resulted in symptomless infections. Mutations in the cleavage site between P3 and 6K1 led to variation in the appearance and strength of the symptoms.

1.3.4. Cylindrical inclusion protein (CI)

The cylindrical inclusion bodies are pinwheel type cytoplasmic inclusion bodies, which are made up of a single protein (Hiebert and McDonald, 1973). The CI from several potyviruses was isolated and characterised with a molecular weight ranging from 65 to 75 kDa. As mentioned before, the identification of CI proteins in infected plant cells is useful for assigning viruses to the genus *Potyvirus*.

Several sources of information led to the hypothesis on the indispensable involvement of the CI protein activity in the life cycle of potyviruses. The cylindrical inclusion protein contains a so-called nucleotide binding site, an amino acid sequence motif present in proteins encoded by most positive-strand RNA viruses. Lain *et al.* (1991) reported for the first time that the purified CI of PPV interacts with RNA and ATP and co-purifies with a nucleic acid-stimulated ATPase activity. It was also reported that CI may acts as an RNA helicase (Lain *et al.*, 1990). Domier *et al.*, (1987) pointed out that CI, the principal protein component of the cylindrical inclusions that form in the cytoplasm of infected cells, is similar in sequence to the membrane-bound and replication-associated protein of picornaviruses. Further sequence analysis classifies the cylindrical inclusion proteins of potyviruses as members in the superfamily of helicase-like proteins.

David *et al.* (1991) stated that the CI protein was found mostly as pinwheels in the cytoplasm and also in association with cell membranes. Electron microscopic studies of thin sections of leaves infected by some potyviruses show a relationship between the CI and the plasmodesmata of plant cells suggested that this protein may be involved in viral transport (Lawson and Hearon, 1971; Andrews and Shalla, 1974). Roberts *et al.* (1998) also reported that the CI of *Pea seed-borne mosaic virus* (PSbMV) was located at the entrance of the plasmodesmata. Another function of CI suggested by Langenberg (1986) is a possible involvement in cell-to-cell movement of potyvirus particles. Virus particle were found to be attached to the cytoplasmic inclusions and cytoplasmic inclusions aggregates were shown to be associated with plasmodesmata. The presence of ATPase activity in plasmodesmata of infected cells also suggests that cell-to-cell movement of the viral RNA

requires ATP and helicases, which are linked to ATPase, thus, CI protein may contribute in cell-to cell movement of the virus.

An important suggestion was reported by Lain *et al*. (1990) that the possible function of CI may include the unwinding of replicative intermediates. It may also be involved in genome recombination or unwinding of single strand RNA secondary structures to facilitate transcription or translation. The functions of CI protein is not completely resolved, especially its role in virus replication. Klein *et al*. (1994) investigated the possible role of CI in this process by creating mutations in an infectious full-length clone TVMV. Mutations in the coding region of the genome covering the CI resulted in mutants unable to replicate in the host plants. Extensive studies were carried out by Fernandez *et al.* during 1995, 1996 and 1997 for more information about helicase activities and RNA binding of CI. They found after introducing a series of deletions within the amino acid sequence of the PPV-CI protein (deletions upstream of the RNA binding domain) dramatically impaired the helicase and NTPase activities, but retained RNA binding capacity indicating that RNA binding can occur in the absence of helicase activity. In 1996, they characterised a second RNA binding domain and they determined it between amino acids 73 and 143. In 1997, they introduced other mutations in the CI of PPV. The mutants were deficient in RNA helicase activity, replicated only to low levels in protoplasts and did not infect host plants, indicating that the mutations altering the domain involved in NTPase activity, which is required for the unwinding process. Rojas *et al.* (1997) found that the CI is not a movement protein in the same sense as CP and HC-Pro, because it is not able to move to the neighbouring cells and remains in the cytoplasm of a cell where it forms aggregates. Rojas *et al*. (1997) and Rodriguez-Cerezo *et al.* (1997) suggested that the CI is located close to plasmodesma connections at early stages of infection that was known by using immunogold labelling procedures. They also indicated that the CI might be involved in cell-to-cell passage of the viral RNA-protein complex. In addition to all the suggested functions mentioned above, the CI may be involved in wilting symptoms. Chu *et al.* (1997) found two genetic determinants of TEV-HAT involved in wilting, whereas one of them covers the 3'-end of the CI, the 6K2 and the 5'-end of the VPg-NIa coding regions, the other one is located at the 3' one-third of the P3 encoding region.

1.3.5. The two small proteins 6K1 and 6K2

Despite extensive studies of potyviruses on the molecular biology level, very little information is available about the two 6 kDa proteins 6K1 and 6K2. Their functions are still not understood well and only suggestions are made, which are not fully supported by experiments. No specific function has been reported for the 6K1 peptide or for the P3-6K1 protein during the potyviral infection cycle before 2002. Some older studies were carried out to determine the role of these two proteins by mutation of the junction sites with the other proteins. It was shown that the partial processing at the cleavage site present between the P3 and 6K1 cistrons in PPV by the NIa protease occurs in vitro. Mutations constructed by Riechmann *et al*. (1995) prevented the cleavage at the proteolytic site between P3 and 6K1 but did not abolish virus infectivity, when an infectious full-length PPV cDNA clone was used. Transcripts containing a histidine for glutamine substitution in the cleavage site sequence were able to infect plants, indicating that the processing at P3-6K1 junction is not required for virus viability. The virus accumulation occurred after a second site mutation was introduced into the 6K2 cistron during replication. Another mutation was also engineered and it affected the time course and severity of the symptom induction process. However, the processing at the P3-6K1 junction plays a relevant role in the life cycle of potyviruses, and the presence of the presumed 6K1 peptide in infected plants has not yet been observed for any potyvirus.

Previous studies of Carrington *et al.* (1988) suggested that the cleavages at both sites flanking the 6K2 protein are processed by the NIa protein. Efficient processing of the cleavage site between CI and 6K2 was observed only if proteolysis of the site between 6K2 and VPg was prevented, which indicates that the cleavage between CI and 6K2 takes place in cis, as proposed for TEV. The two small proteins 6K1 and 6K2, predicted from the amino acid sequence of the potyviral polyprotein, could play a role in RNA replication, also they have not yet been identified *in vivo*. The presence of a stretch of hydrophobic amino acids in the 6K1 and in the 6K2 peptides, and the relative position of their cistrons in the potyviral genetic map, suggest that they could be analogous to the picornavirus 2B and 3A peptides, respectively (Lain *et al*., 1989). Therefore, it is conceivable that 6K2 could anchor the NIa/VPg to membranes and that processing at the 6K2-NIa junction might be related to

some step of the RNA replication process, although cleavage at this site might not be the last processing event to release the VPg from the polyprotein precursor. No enzymatic function was reported for the 6K2 peptide. But its intracellular localization was studied in more detail. Restrepo-Hartwig and Carrington (1992) demonstrated that the 6K2 peptide of TEV prevents transport of NIa to the nucleus, and it seems to override nuclear translocation of NIa. Mutations that led to improper maturation of the 6K2 peptide debilitated the virus or were lethal (Restrepo-Hartwig and Carrington, 1994). However, another mutation that prevents processing of the 6K1 peptide from P3 protein, or the 6K2 peptide from CI, were not lethal to TEV. It has been proposed that the 6K2 peptide is required for genome amplification and associates with large vesicular compartments deriving from the endoplasmic reticulum. Garcia *et al.* (1990) reported that a confident analysis of the 6K-NIa sites was not possible, because the efficiency of processing at the NIa-NIb site was very poor and free 6K could not be identified. It was illustrated by Rajamaki and Valkonen (1999) that the 6K2 protein is a virulence determinant of PVA-M and suggested a possible coordinated function of 6K2 in the vascular movement of PVA thus, they supported the hypothesis of Schaad *et al.* (1996) that the 6K2 peptide may function as a membrane anchor for CI and VPg during replication. The replication complex may be released from the membrane by catalysis of cleavage at the polyprotein proteolytic site between 6K2 and VPg. A recent study of Merits *et al*. (2002), using plant cells and a baculovirus expression system, suggested that both 6K peptides are indispensable for virus replication because deletion of either the 6K1 or 6K2 encoding regions rendered PVA non infectious. The cleavage process of the 6K2 protein from the adjacent proteins by NIa proteinase is very important for the rate of virus replication and virus movement. Deletion of the genomic region encoding the 6K1 protein prevented proper proteolytic separation of P3 from CI, but did not affect processing of VPg, NIa-Pro, NIb or CP from the polyprotein. The 6K2 encoding sequence could be removed without any detectable effect on polyprotein processing. However, the two small proteins seem to be important for the PVA life cycle.

1.3.6 Viral protein genome-linked (VPg)

VPg is present at the 5'-end of potyvirus genomic RNA as a typical structure instead of a cap. The VPg plays a role to protect mRNA from attack by exonucleases. In addition, it is probably involved in RNA replication acting as a primer for nucleic acid synthesis (Takeda *et al.*, 1986) or cleaving a replicative form of the RNA (Tobin *et al.*, 1989). Many experiments were performed to determine the role of potyvirus VPg. Destroying of VPg by using proteinase K increased infectivity of TEV RNA (Hari, 1981), but digestion of VPg with proteolytic enzymes diminished the infectivity of PPV RNA (Riechmann *et al.*, 1989). The VPg is not required for infectivity when mechanical inoculation is done with potyviruses transcripts. This demonstrates that paternal VPg is not essential for infectivity, but the low efficiency of the in vitro transcripts suggests that the VPg may be relevant for infection. Murphy *et al.* (1990) suggested that the VPg of TEV RNA is the entire 49-kDa proteinase (NIa) or the N-terminal 24-kDa part of the proteinase. To examine the possible roles of the VPg protein, several mutants were studied. Hong *et al.* (1995) indicated that four VPg mutants displayed markedly reduced interactions with the NIb protein, this interaction may be important for the initiation of viral synthesis in infected cells. The VPg can be a virulence determinant in incompatible potyvirus-host combinations in which the resistance prevents systemic virus infection. Substitution of amino acid residues in the central part of the VPg allowed isolates of TVMV (Nicolas *et al.*, 1997) and PVY (Masuta *et al.*, 1999) to overcome a resistance. Furthermore, Rajamakai and Valkonen (2002) indicated that virus strain-specific resistance to systemic infection with PVA in *Solanum commersonii* is overcome by a single amino acid substitution in the VPg.

Infection with the isolate PVA-M of PVA is restricted to the inoculated leaves of *Nicandra physaloides*, whereas the isolate PVA-B11 infects plants systemically. Site-directed mutagenesis of the cDNA clones of PVA-B11and PVA-M showed that the 6K2 protein and the VPg are determinants of systemic infection in *Nicandra physaloides* (Rajamaki and Valkonen, 1999). It was confirmed by Borgstrom and Johansen (2001) that mutational analysis of VPg of PSbMV isolates representing pathotypes P1 and P4 using recombinant infectious clone (P1: line 269818 is resistant to PSbMV pathotype P1 and P4: susceptible to

pathotype P4) revealed that changes affecting amino acids 105 to 117 in the central region of VPg influenced virulence on P1 line 269818.

To understand more about the role of VPg in the potyviral replication complex, the VPg interaction with other potyviral proteins were studied. It was found that the TVMV VPg is able to stimulate the polymerase activity of the NIb. Moreover, the VPg was able to stimulate a mutant NIb (Fellers *et al.*, 1998).

The VPg protein may also have a possibly coordinated function together with 6K2 in the vascular movement of PVA in different host species (Rajamaki and Valkonen, 1999). Recent studies have been demonstrated that the multifunctional VPg of PVA was found to be phosphorylated as a part of the virus particle by a cellular kinase activity from tobacco. Thus shows that the virion-bound VPg interacts with host proteins and that phosphorylation of VPg may play a role during the infection cycle of potyviruses (Puustinen *et al*., 2002).

1.3.7. Nuclear inclusion protein a (NIa)

All potyviruses code for two proteins that comprise the nuclear inclusion bodies: the first protein is the NIa with an apparent molecular weight of 49 kDa, which is also a viral encoded proteinase responsible for the cleavage of the polyprotein at five locations characterized by highly conserved sequences including its own autocatalytic excision (Carrington and Dougherty, 1988; Hellmann *et al.*, 1988; Garcia *et al*., 1989). The cleavage sites of the NIa were characterised for several potyviruses and showed differences, which might affect the speed of the proteolytic cleavage.

The bifunctional NIa polyprotein consists of an N-terminal VPg, not relevant for its proteolytic activity in vitro and a C-terminal proteinase region (Carrington *et al*., 1993; Garcia and Lain, 1991). As shown before, the VPg is covalently attached to the genomic RNA. Analysis of products expressed from the NIa protein genes, which were altered by deletion mutagenesis, revealed that only the carboxyl-terminal half was required for proteolytic activity (Carrington and Dougherty, 1987a). NIa makes up more than 95% of the total inclusion protein (Dougherty and Hiebert, 1980). The majority of TEV NIa molecules are localized to the nuclei of infected cells, however a portion of NIa is attached

covalently as VPg to the viral RNA in the cytoplasm. Several studies showed that NIa has a proteinase activity. Site-directed mutations that alter amino acid residues within the NIa protein eliminate the proteolytic activity (Carrington and Dougherty, 1987b). NIa proteases catalyse post-translational proteolysis at an internal cleavage site, which releases VPgs and protease domains of the NIa proteins. This proteolytic cleavage seems to happen in some but not all of the NIa molecules (Dougherty and Parks, 1991). The cleavage might be related to some steps of the viral RNA replication process. It is suggested that the processing intermediates of the proteolytic pathway may have functions, but which are still not known. NIa is related to the trypsin-like family of cellular serine proteases but with substitution of Cys at the active site (Gorbalenya *et al.*, 1989). Garcia *et al.* (1990) indicated that substitution in the PPV NIa proteases resulted in anomalous proteolytic activities, suggesting that there are different structural requirements in the protease for processing at the various cleavage sites. The low amino acid similarity of the C-terminal end of NIa proteases of potyviruses indicated that this region of the protease may be important in forming the substrate-binding pocket which allows each of the potyviral NIa proteases to recognize its own unique cleavage sequences (Ghabrial *et al*., 1990). Different mutations were conducted in the C-terminal part of NIa proteases, which affected the proteolytic cleavage processes of various cleavage sites whereas the effect depends on the selected cleavage site (Garcia *et al*., 1990; Carrington and Dougherty, 1987a). These results indicated that the extreme carboxyl end of the protease could be involved in maintenance of a proper structure that would allow other parts of the protein to interact correctly with the various cleavage sites. Two studies were carried out by Garcia and Lain (1991) and Parks and Dougherty (1991) to delimit protease regions involved in catalysis and in substrate recognition through the construction of hybrid NIa proteases. The results suggested that the catalytic and recognition sites of the NIa proteases are closely interlinked and that the main determinants for substrate specificity lie in the C-terminal one-third of the protease. Most of the researches showed that the NIa-catalytic maturation of the potyviral polyprotein is a regulated process involving cis and trans cleavages at sites which are recognized with different efficiencies and processed at different rates. NIa of TVMV interacts with itself and with NIb, and this interaction does not belong to the NIb domain since it was not affected by mutations in the conserved motif (GDD) of NIb. It was abolished by mutation

in the VPg domain, indicating NIa-NIb interaction arises through interaction with the VPg domain of NIa (Hong *et al.*, 1995).

1.3.8. Nuclear inclusion protein b (NIb)

NIb is the most conserved protein expressed from a potyvirus genome (Shukla *et al.*, 1991). It is the second and large nuclear inclusion protein (58 kDa) (Knuhtsen *et al.*, 1974) that might act as a RNA-dependent RNA polymerase or virus replicase (Alisson *et al*., 1986; Domier *et al*., 1986). NIb and NIa proteins were found to accumulate in the nucleus of TEV-infected cells at all times during which viral antigens could be detected, indicating that they contain nuclear transport signals (Restrepo *et al.*, 1990). The structural integrity of NIb is involved in controlling the transport of the protein to the nucleus, but its function in the nucleus is still not quite clear. However, localisation may vary for different potyviruses (Riedel *et al*., 1998) and accumulation occurs also in the cytoplasm. The NIb is required in the cytoplasm or at membranes associated with replication complexes during RNA synthesis (Urcuqui-Inchima *et al*., 2001). Most of the studies indicated that NIb is the RNA-dependent RNA polymerase of potyviruses due to a conserved sequence motif 'GDD' found in replicases (Hong and Hunt, 1996). Merits *et al.* (1998) found that NIb binds RNA in a sequence unspecific manner. Insertional mutations in the NIb of TVMV demonstrated that NIb is required for genome amplification (Klein *et al*., 1994). Li et al. (1997) showed the interaction of several TEV NIb mutants with the protease domain of NIa. In addition two independent nuclear localization signals of TEV NIb were identified. The NIb-NIa interaction was not affected by mutation in the GDD motif, but this mutation diminished the NIb-CP interaction (Hong *et al.*, 1995) and resulted in a dominant-negative phenotype. Loss of NIa-NIb interaction occurred by mutation of the TEV proteinase domain, thus, this result supports the hypothesis that interaction between NIa and NIb is important during TEV genome replication (Daros *et al*., 1999). Hong *et al.*, (1995) suggested that NIa-NIb interaction may be important for initiation of viral RNA synthesis in infected cells and they concluded that NIb-CP interaction is sensitive in changes in the highly conserved GDD motif of NIb, however the role of this interaction in the function of the NIb protein is not clear, but it is assumed that it may be involved in regulation of viral RNA synthesis in the infected cells.

1.3.9. Coat protein (CP)

Potyvirus genomes are encapsidated by around 2000 molecules of a single type of CP. It is the most extensively characterized gene product of potyviruses and selected for studies of genetic diversity and taxonomy. Many potyviruses and potyvirus isolates have been determined and classified on the basis of their coat protein sequence data (Higgins *et al*., 1998; Fuji *et al*., 2003). Comparison of CP sequences led to classification of members of the family *Potyviridae* into genera, distinct species, related species and strains (Xiao et al, 1994; Shukla *et al*., 1994). Potyvirus coat proteins vary considerably in size due to differences in sequences near the N-termini (Aleman *et al*., 1997). The N-terminus of the CP contains virus-specific epitopes (Shukla *et al*., 1989a; 1989b) and is involved in host-vector-virus interactions. The C-terminal end of the CP is also a variable domain, while the central part is a more conserved domain. This part is important for the structure and assembly of CP molecules. The CP is involved in aphid transmission (Pirone and Thornbury, 1983; Atreya *et al.,* 1990) as found for CPs in other plant virus groups. A specific conserved amino acid motif (DAG) present at the N-terminus of the CP is involved in virus transmission by aphids (Maiss *et al*., 1989). This hypothesis was confirmed experimentally by mutating full–length clones of some potyviruses to detect the importance of this conserved motif. Changes of residues in or adjacent to this amino acid triad resulted in loss or greatly reduced transmissibility. Atreya *et al.,* (1990) demonstrated that site directed mutations in the DAG motif of TVMV eliminated or decreased aphid transmissibility. However the amino acid that follows this conserved motif (DAG) is not strictly conserved in potyviruses but mutations at this position in TVMV reduced also aphid transmissibility (Atreya *et al.,* 1995). Recently this motif has been extended to DAGX (Flasinki and Cassidy, 1998). In addition it was shown that the removal of the N-terminal domain from PVY CP and ZYMV CP rendered these viruses non aphid-transmissible (Salomon, 1989; Salomon and Raccah, 1990). Mutations introduced into the PStV aphid transmission motif decreased aphid transmissibility. Flasinnki and Cassidy (1998) suggested that additional CP motifs influence aphid transmission or the potyvirus CPs could interact with other motifs from another protein like the HC-Pro.

Sequence comparisons among HC-Pro of potyviruses define a region of the HC-Pro that may interact with the CP. The highly conserved motif (PTK) of the HC-Pro was proposed to be involved in bridge formation with the N-terminal DAG motif of the CP (Huet *et al*., 1994). The CP is believed to be responsible for the formation of movement-competent complexes with viral RNA (Atabekov *et al*., 2000), but the mechanism of viral RNA transport is poorly understood. Some studies proposed that CP affects the movement of potyviruses, mutations in the DAG motif in the CP of TVMV prevented the systemic infection of tobacco plants. A change in the first residue of the DAG motif of TEV-CP near the CP N-terminus caused that the mutant failed to infect tobacco plants systemically and showed a limited cell-to cell movement (Lopez and Pirone, 1998).

Fedorkin *et al.* (2001) reported that cell-to-cell movement of PVX involves distinct functions of the CP. Coat protein of PVY expressed in a transgenic plant can complement movement-deficient CP of PVX emphasizing a direct involvement of potyvirus CP in cell-to-cell movement. N- and C-terminal deletions of TEV-CP were constructed in order to determine the minimum requirements for protein self-assembly. All gene constructs were expressed in *E. coli* and encoded fusion proteins of the expected size (Voloudakis *et al*., 2004). Removal of 28, 63 or 112 amino acids from the N-terminal part of the TEV CP did not prevent the formation of potyvirus-like particles. However deletion of 175 amino acids abolished particle formation. It was suggested that the highly conserved domains in the core region of potyvirus CPs are important for virus assembly (Dolja *et al*., 1991). Varrelmann and Maiss (2000) reported that mutations in the PPV-CP suppress particle assembly and indicated that mutation in the assembly motifs were solely responsible for the absence of systemic infection. Two mutants of *Johnson grass mosaic virus* (JGMV) CP produced by site-directed mutagenesis failed to assemble into virus-like particles (Jagadish *et al*., 1991; 1993) and the virus became unable to move from cell to cell. In addition to all the above-mentioned functions, Hong *et al.* (1995) suggested that CP is involved in regulation of viral RNA synthesis due to the mutation constructed in the conserved GDD motif of NIb, which affected the interaction between CP and NIb. Finally, the CP was widely used to produce pathogen-derived resistance in transgenic plant species.

1.3.10. 5’ and 3’ Non-translated regions (NTRs)

The 5’- and 3’-NTRs are located at the ends of the potyvirus genome. The 5’-NTRs of potyviruses are very rich in adenine residues (Gallie *et al*., 1987). Carrington and Freed (1990) stated that 5’-NTR of TEV is associated with translation enhancing activity when fused with a reporter gene. Nicolaisen *et al*. (1992) demonstrated that the 5’-NTR of PSbMV can serve as a general translational enhancer in future genetic transformations of plants. Riechmann *et al*. (1992) suggested that the 5’-NTR of some potyvirus genomes as PPV, PVY, TVMV and TEV contain some conserved sequence regions which may play an important roles in replication and translation. In contrast, Simon-Buela *et al*. (1997). found that the 5’-NTR of PPV is not necessary for virus infectivity but contribute to viral competitiveness and pathogenesis.

The 3’-NTRs of potyviruses differ in length and display no significant sequence homology. It serves as an aid for identification and classification of potyviruses. Frenkel *et al*. (1989) suggested that the sequence of the 3’-NTRs may be an accurate marker of the potyvirus genome and it can be used as a region of genetic relatedness in the taxonomy of potyviruses. The 3’-NTR was used to distinguish among potyviruses and virus strains because the 3’-NTRs of related strains are very similar in length and nucleotide sequence. Rodriguez-Cerezo *et al*. (1991) indicated that the 3’-NTR of TVMV RNA contains a determinant of disease symptom severity which may have a direct effect on the induction of disease symptoms by an RNA virus. In addition the 3’-NTR is polyadenylated and attributes for virus infectivity (Tacahashi and Uyeda, 1999).

1.4. Description of *Beet mosaic virus* (BtMV) and aim of the study

Sugar beet crops (*Beta vulgaris*) are naturally affected by several viruses, like *Beet yellows virus* (BYV), *Beet mild yellowing virus* (BMYV), *Beet western yellows virus* (BWYV), *Beet cryptic virus* (BCV), *Beet chlorosis virus* (BChV), *Beet necrotic yellow vein virus* (BNYVV) and *Beet mosaic virus* (BtMV). BtMV is a member of the genus *Potyvirus* in the family *Potyviridae* (Milne, 1988). It was first reported in 1898 in garden beet in northern France and near Paris. Later it was found in *Beta vulgaris* L. from Braunschweig, Germany

(Schneider and Mundry, 1956). BtMV is distributed worldwide in major beet-growing areas, especially in temperate regions of the world (Russell, 1971).

Farmers knew it as a virus that does not cause serious economical problems. Watson and Watson (1953) suggested that BtMV causes only little damage. However, it is generally assumed that BtMV does not affect yield significantly and the infected plant is often of normal size. Little effect is expected on yield of naturally infected crops due to the low natural infection levels and the late time at witch the crop becomes infected. The spread of BtMV is a process depending on the interaction between the host crop, the vectors, the virus source and the environmental conditions. The quantification of these relations is still not understood well. Information on losses caused by BtMV is limited, but old and few studies were carried out to determine the quality and quantity of BtMV damage. Schmelzer and Hartleb (1977) pointed out that this disease reduces industrial sugar beet yield up to 30%, and beet seed crops infected in the first year resulted in 50% lower seed yields. Another research in former Yugoslavia showed that BtMV reduced yields of roots by 1.5 to 10.5%, green matter by 12%, sugar yield by 8 to 19% and sugar content by 0.65 to 1.06% (Nikolic, 1958). BtMV infection reduces germination by 37.4% and the length of primary roots by 26.6% (Staki and Jasnic, 1985). Most strains of the virus decrease potential root yield and assimilation rate of 10% (Watson and Watson, 1953). But the virulent strains may cause losses of about 20% (Heathcote, 1973). Dusi (1999) observed during the field experiments that BtMV effected the overall growth of infected sugar beet plants in comparison with the healthy plant but the effect on the plant development is small, especially with the late infection, and simulated damage by BtMV infection resulted in an estimated damage of approximately 3%. Dusi (1999) stated that the earlier infected crop results in a higher damage in yield, and the considerable reduction of yield can occur when disease incidence is high early in the season. The delay of BtMV introduction in the field reduces loss by reducing damage at the plant level and by decreasing disease incidence at the crop level. When late infection occurs the crop has already leaf material enough to sustain yield at commercial levels. BtMV affects the photosynthesis rate and causes a reduction in mature leaves due to the reduction of chlorophyll content in the plant. Infection with BtMV reduced the maximum rate of leaf gross photosynthesis in mature leaves by

16% and increased dark respiration by 85 to 90% for young as well as mature leaves (Dusi, 1999).

Host range of BtMV includes all cultivated sugar beet (Schmelzer and Hartleb, 1977) and near relatives. One hundred species in 26 dicotyledonous families have been infected experimentally (Grüntzig, 1988), but BtMV infects mainly plants in the families *Chenopodiaceae*, *Solanaceae* and *Leguminosae*. It causes a mosaic disease in *Beta vulgaris* and in *Spinacea oleracea* and shows clear symptoms on these two species.

The symptoms of BtMV distinctly differ from those evoked by other viruses attacking sugar beet. Initial symptoms are numerous small yellow spots and blotches on one or several central leaves. A light mosaic and mottle occurs on young leaves as disease develops. Leaflets with initial symptoms are stunted; also curling and rolling of leaf margins and leaf tip necrosis occurs. In several cases, diseased leaves roll into a tubular shape. In contrast to BYV no phloem necrosis or starch accumulation occurs in leaves of BtMV infected plants. It was found that BtMV produced mosaic symptoms on sugar beet 7-12 days after inoculation, and it caused local lesion necrotic spots on inoculated leaves of *Chenopodium amaranticolor*, but it failed to infect some *Nicotiana* species and soybean plants by sap inoculation (Rogov *et al*., 1991). Dusi (1999) found that BtMV shows symptoms in all sugar beet leaves that develop after the infection, all leaves already present on the plant prior to infection did not develop any symptoms. BtMV causes clear symptoms on spinach (*Spinacea oleracea*). Youngest leaves show small yellow flecks, often coalescing to form large chlorotic areas. Older leaves become progressively chlorotic and necrotic, and infected plants are usually stunted. Like other potyviruses, it can be mechanically transmitted to *Nicotiana benthamiana* to produce very clear symptoms and effects on leaves formation. Three plant species were found naturally infected with BtMV in sugar beet fields in Southern California: *Amaranthus retroflexus* L. (pigweed), *Melilotus indica* L. (yellow sweet clover) and *Trifolium incarnatum* L. (crimson clover) (Bennett, 1949). It was experimentally reported that there are more than 40 insusceptible species from different plant families as *Compositae, Cruciferae, Cucurbitaceae, Leguminosae, Polemoniaceae, Solanaceae* and *Umbelliferae*. Symptomless BtMV infected species were also reported (e.g. *Sterllaria media* Vill) (Heathcote, 1973). Bennett (1949) pointed out that

BtMV has few local lesion hosts, developing specific local symptoms (*Beta patellaris*, *Amaranthus retroflex* and *A. caudatus*). Russell (1971) mentioned that BtMV produces chlorotic local lesions on *Gomphrena globosa* and *Chenopodium quinoa.*

BtMV, as well as several other potyviruses is transmitted by several vector species, with varying degrees of efficiency (Sylvester, 1952). It can be spread also by aphid species that do not colonize the crop. BtMV infection may be predicted on the basis of the initial infection date and vector abundance (Sylvester, 1952; Dusi *et al*., 2000). BtMV is transmissible by more than 28 aphid species (Kennedy *et al.*, 1962) including *Acyrthosihpon pisum*, *Aphis appi*, *Macrosiphum euphorbiae*, *Metopolophium dirhodum*, *Rhopalosiphum padi* (Sylvester, 1952). However, *Myzus persicae* and A*phis fabae* are the principal vectors in the field. Transmission occurs in a non-persistent manner (Watson and Watson, 1953; Sylvester, 1952) and is improved by starving aphids for 2-5 min before acquisition feeding. Sylvester (1949) indicates that the acquisition and inoculation thresholds lie somewhere between 6-10 sec, with a mean transmission threshold value of approximately 1 min and no latent period. Persistence of the virus in the vector depends on the aphid species (Sylvester, 1952) but is probably less than 1 hr (in apterae). All instars transmit virus but alata are more efficient than apterae. Sylvester (1949) stated that fasting has a beneficial effect upon the ability of apterae to transmit BtMV. The virus is not passed to the offspring and it does not persevere through aphid molting (Russell, 1971). BtMV can be transmitted by mechanical inoculation utilizing infect plant sap. In beet sap, thermal inactivation point (10 min) is between 55 and 60° C, dilution end-point is up to 1/4000 and infectivity is retained for 24-48 hr at 20 °C (Russell, 1971). BtMV can be transmitted by grafting, but probably is not transmitted by seed and dodder, because three species of *Cuscuta* failed to transmit BtMV (Bennett, 1944). In addition BtMV is not transmitted by contact between plants or by pollen.

The disease caused by BtMV is polycyclic. The virus spread occurs within the field in several cycles during the development of the epidemic from the first source established. The latter infected plants play a major role in this epidemiological process. The shorter the latency period, the faster the epidemic can develop. Dusi (1999) indicated that the length of

the latency and incubation period increased with decreasing temperatures. He observed that the latency and incubation periods are shorter during the summer months.

Traditional methods were used to control BtMV spread, e.g. chemical control of vectors and elimination of weeds known to harbor BtMV, because BtMV is maintained in natural hosts, including biennial seed crops (wild beet) and winter spinach for year to year spread. Well-timed treatment of sugar beet root crops with systemic insecticides was used in England to control the aphid vector of the virus disease. In addition a decrease in the mangold crops acreage has also contributed to its decline. The systemic organophosphorous insecticides that became wildly used decreased the spread of BtMV within sugar beet root crops, but not the primary infections because the plants become infected during a feeding probe by an aphid lasting only a few seconds (Heathcote, 1973). Many studies did not support the chemical control of aphids, since the insecticides have a slow mode of action against incoming alatae (Eckel and Lampert, 1993; Raccah, 1986; Loebenstein and Raccah., 1980). When the aphid numbers are very high the high spraying frequency is not effective (Pirone *et al*., 1988; Roberts *et al*., 1993).

Electron microscopy studies were carried out in order to determine the type and the place of amorphous inclusion bodies of BtMV in infected plant cells. Fujisawa *et al.* (1967) found that BtMV infected cells contain vesicular X-bodies produced in the cytoplasm and appeared as a typical inclusion body, located near the nucleus and were encountered in almost all epidermal cells. The X-bodies were denser than the nucleus, plastids, mitochondria, and cytoplasmic ground substances. Vacuole-like spaces were found throughout the X-bodies and crystalline inclusions occur in the chloroplasts. Nucleoli are enlarged and distorted in beet (Bos, 1969) and in *Gomphrena globosa* (Martelli and Russo, 1969). Hoefert (1969) and Rogov (1991) pointed out that cytoplasm contains inclusions which appear in section as pinwheels and bundles and are digestible by pepsin but not by ribonuclease, thus this criterium supports that BtMV is a member of the genus *potyvirus*. Jafri (1972) found by using a standard leaf-dip technique that BtMV has flexuous, filamentous particles with a mean length of 730 nm long and 13 nm in diameter. The progressive degeneration of the substructure of some of the organelles were the most prominent features of BtMV infected tissue, and pinwheels were commonly found in the

cytoplasm of mesophyll cells, crystal formation was one of the prominent features of some of the degenerating chloroplasts from mesophyll cells. Frequent disintegration of mitochondria was also observed in degenerated cells.

Rogov (1991) analysed the CP of BtMV by electrophoresis polyacrylamide gels in order to determine the molecular weight, which was about 34700 ± 700 Da. BtMV is immunogenic and serologically related to BYMV, PVY and SMV while BYV and PPV displayed no serological relation to BtMV (Bercks, 1960; Grüntzig and Fuchs, 1979; Rogov, 1991).

Despite its clear position in the genus *Potyvirus* BtMV is not well characterized at the molecular level. Therefore the major objective of this research is the molecular characterization of BtMV including cloning and sequencing of the entire virus genome. The subsequent sequence analyses should help to establish a precise position of BtMV in the genus *Potyvirus* and in the family *Potyviridae*. Moreover, the construction of a BtMV infectious full-length clone will provide the basis for further analyses of the BtMV encoded proteins.

2. Materials and methods

2.1. Materials

2.1.1. Plants

Plant	Source
Atriplex hortensis cv. Rheinische	Gartenmelde, Gärtner Linie®
Beta vulgaris, (6B22840), (8T0015)	Planta / KWS
Chenopodium quinoa	Reismelde, IPP
Beta vulgaris ssp. *vulgaris* var. *cicla* cv. Lukullus	Weißstieliger Mangold, Gärtner Linie®
Beta vulgaris ssp. *vulgaris* var. *conditiva* cv. Rote Kugel 2	Rote Rübe, Gärtner Linie®
Nicotiana benthamiana	IPP
Spinacia oleracea cv. Matador	Spinat, Gärtner Linie®

2.1.2. Virus isolate

Virus name	Number of isolate	Source
Beet mosaic virus (BtMV-G)	PV-0065	DSMZ, Braunschweig

2.1.3 Antibodies

Antibody	Source
BtMV specific antibody (IgG)	DSMZ AS-0143
IgG Goat-anti-Rabbit-Antibody	Sigma

2.1.4. Bacterial strain

Bacterium	Strain	Source and description
E. coli	NM522	Pharmacia; F' lacI^{Q}Δ(lacZ)M15 proA^{+}B^{+} /supE thi Δ(lac-proAB) Δ (hsdMS-mcrB)5 ($r_k^-m_k^-$McrBC$^-$)

2.1.5. Plasmids

Plasmid	Resistance	Company
p442 pe35StupA	Ampicillin	Provided by E. Maiss
pBluescript II KS (-)	Ampicillin	Stratagene
pGEM T-Easy vector	Ampicillin	Promega
V81 pNEB193	Ampicillin	New England Biolabs
V100 pSPORT1	Ampicillin	Gibco BRL
V206 pNEB193NheI	Ampicillin	New England Biolabs, modified by E. Maiss
V208 pNEB193EM	Ampicillin	New England Biolabs, modif. by E. Maiss
V209 pe35Stu paII	Ampicillin	Provided by E. Maiss
V217 pBlueKSP XBNMX	Ampicillin	Stratagene, modif. by E. Maiss
V220 pDRIVE	Kanamycin	Qiagen
V223 pDRIVE2	Kanamycin	Qiagen, modified by E. Maiss

2.1.6. Antibiotics

Antibiotic	Company
Ampicillin	Serva
Kanamycin	Serva

2.1.7. Enzymes

Enzyme	Source	Concentration
AMV Reverse Transcriptase	Promega	10 u/µl
Expand Reverse Transcriptase	Roche	50 u/µl
Klenow Fragment (exo$^-$)	MBI	5 u/µl
Pfu DNA Polymerase (proof reading)	Promega	2-3 u/µl
Platinum *Taq* Polymerase	Invitrogen	10 u/µl
Proteinase K	Roth	20 mg/ml
RNase A	Fluka	75 u/µl
RNase Inhibitor	Promega	400 u/µl
RQ1 DNase	Promega	1 u/µl
Superscript II Reverse Transcriptase	Invitrogen	200 u/µl
T4 DNA ligase	MBI	1 u/µl
Taq DNA Polymerase	PeqLab	5 u/µl
Taq DNA Polymerase	Promega	5 u/µl
Terminal Transferase (TdT)	BioLabs	20 u/µl

2.1.8. Restriction enzymes

Enzyme	Company	Concentration
*Bam*H I	MBI	10 u/µl
*Bsp*E I	BioLab	10 u/µl
Dra I	BioLab	20 u/µl
*Eco*R I	MBI	10 u/µl
*Eco*R V	Promega	10 u/µl
*Hin*d III	MBI	10 u/µl
Kpn I (*Acc*65 I)	MBI	10 u/µl
Mlu I	MBI	10 u/µl
Nco I	MBI	10 u/µl
Nhe I	MBI	10 u/µl
Pst I	MBI	10 u/µl
Pvu II	MBI	10 u/µl
Sal I	MBI	10 u/µl
Stu I (*Eco*147 I)	MBI	10 u/µl
Sty I (*Eco*130 I)	MBI	10 u/µl
Xba I	Promega & MBI	10 u/µl
Xho I	Promega	10 u/µl

2.1.9. Chemicals

Material	Company
Acetic acid 99.8%	Roth
Agarose	Biozym
Bovine serum albumin	Loewe
Bromophenol blue	Serva
Calcium chloride	Merk
Celite 545	Serva
CF11 (Cellulose)	Whatmann
Chloroform	Merck
D(+)-Glucose	Roth
dATP, dGTP, dTTP, dCTP, (100 mM)	Roth
Diethanolamine	Sigma
Diethylpyrocarbonate (DEPC)	Sigma
Dimethyl sulphoxide (DMSO)	Serva
EDTA	Merck
Ethanol	Roth
Ethidium bromide	Sigma
Glycerin 100%	Roth
$HACoCl_3$	Sigma
HEPES	Roth
Hydrochloric acid (HCl)	Merck
IPTG	BTS
Isoamylalcohol	Merck
Isopropanol	Merck
KH_2PO_4	Merck
Lithium chloride	Roth
LM-Agarose	Biozym
$MgCl_2$	Merck
$MnCl_2$ x 2 H_2O	Merck
$MgCl_2$ x $6H_2O$	Merck
Na_2CO_3	Merck
Na_2HPO_4	Merck
$NaHCO_3$	Merck
Ovalbumin	Loewe
Phenol	Roth
Polyvinyl polypyrrolidone (PVPP)	Sigma
Polyvinyl pyrrolidone	Sigma
Potassium chlorid (KCl)	Roth
SDS	Roth
Select agar	Gibco BRL
Select pepton140	Gibco BRL
Select yeast extract	Invitrogen
Sodium acetate (NaAc, or NaOAc)	Merck
Sodium chloride (NaCl)	Merck

2.1.9. Chemicals (continued)

Material	Company
Sodium hydroxide (NaOH)	Merck
ß-Mercaptoethanol	Merck
Tris, pH 7.0	Roth
Tris-base, pH 8.0	Roth
Triton X-100	Serva
Tungsten particles M-17 (1.1 µm Ø)	BioRad
Tween 20	Fluka
X-gal	BTS

2.1.10. Equipment

Equipment	Company
Autoclave	Sanoclav
Balance 1205 MP	Sartorius
Centrifuge 4K10	Sigma
Centrifuge RC5B Plus	Du Pont, Sorvall
Electrophoresis chambers	Pharmacia & Biozyme
ELISA-Microtiter Plate	Greiner BioOne
Heating block DB-3	Techne Inc
Laminar flowhood lamina Air HA 2448 GS	Heraeus
Micro 20 Centrifuge	Hettich
Microcomputer electrophoresis power supply E443 and E417	Consort
Microplate Reader Model 550	BioRad
Minispin Centrifuge	Eppendorf
Mortar	Haldden-wanger
pH Meter (523)	Landgraf Laborgeräte
Pistil	Haldden-wanger
Speedvac	Heto, Hetovac VR-1
T3 Thermocycler	Biometra
UV-Transilluminator TVL-312-A	Spectroline
UV-Transilluminator Type N-90 M	Intas
Vacuum pressure pump for particle bombardment model Millipor	Millipore Corporation
Vortex MS 2 Minishaker	IKA® Works Inc
Water Bath F3	Haake

2.1.11 Acronyms of potyviruses and accession numbers of sequences used in multiple alignments

Virus name	Acronym	Accession
Bean common mosaic virus	BCMV	NC_003397
Bean common mosaic necrosis virus	BCMNV	NC_004047
Bean yellow mosaic virus	BYMV	NC_003492
Beet mosaic virus	BtMV-Wa	NC_005304

2.1.11 Acronyms of potyviruses and accession numbers of sequences used in multiple alignments (continued)

Virus name	Acronym	Accession
Clover yellow vein virus	ClYVV	NC_003536
Cocksfoot streak virus	CSV	NC_003742
Cowpea aphid borne mosaic virus	CABMV	NC_004013
Dasheen mosaic virus	DsMV	NC_003537
Japanese yam mosaic virus	JYMV	NC_000947
Johnsongrass mosaic virus	JGMV	NC_003606
Leek yellow stripe potyvirus	LYSV	NC_004011
Lettuce mosaic virus	LMV	NC_003605
Lily mottle virus	LMoV	NC_005288
Maize dwarf mosaic virus	MDMV	NC_003377
Onion yellow dwarf virus	OYDV	NC_005029
Papaya leaf-distortion mosaic virus	PLDMV	NC_005028
Papaya ringspot virus	PRSV	NC_001785
Pea seed-borne mosaic virus	PSbMV	NC_001671
Peanut mottle virus	PeMoV	NC_002600
Pepper mottle virus	PepMoV	NC_001517
Peru tomato mosaic virus	PTV	NC_004573
Plum pox virus	PPV	NC_001445
Potato virus A	PVA	NC_004039
Potato virus V	PVV	NC_004010
Potato virus Y	PVY	NC_001616
Scallion mosaic virus	ScaMV	NC_003399
Sorghum mosaic virus	SrMV	NC_004035
Soybean mosaic virus	SMV	NC_002634
Sugarcane mosaic virus	SCMV	NC_003398
Sweet potato feathery mottle virus	SPFMV	NC_001841
Tobacco etch virus	TEV	NC_001555
Tobacco vein mottling virus	TVMV	NC_001768
Turnip mosaic virus	TuMV	NC_002509
Wild potato mosaic virus	WPMV	NC_004426
Zucchini yellow mosaic virus	ZYMV	NC_003224

2.2. Methods

2.2.1. Virus source

The German *Beet mosaic virus* isolate (BtMV-G) used throughout this work was isolated from *Beta vulgaris* ssp. *vulgaris* var. *altissima*. The isolate is available from the German Collection of Microorganisms and Cell Cultures (PV-0065).

2.2.2. Planting of *Nicotiana benthamiana* in the greenhouse

Nicotiana benthamiana was selected as a suitable propagation host of BtMV-G. It can be easily inoculated with virus sap (mechanical inoculation) and with full-length clones (particle bombardment). In addition it shows very clear symptoms of virus infection.

The line 27/1 of *Nicotiana benthamiana* was used according to Timpe *et al.* (1992*). N. benthamiana* seeds were sown in soil in a tray and each seedling was transferred to a pot (10-12 cm in diameter) when they reached a height of 3-5 cm. The conditions of the greenhouse were maintained at temperatures of 20-25^{o}C at a moisture of about 70%-80%. In the winter season extra light was given for plant assimilation.

2.2.3. Propagation of BtMV-G

2.2.3.1. Mechanical inoculation

Mechanical inoculation was done in young *N. benthamiana* plants grown in the greenhouse at the 5-10 leaf stage. The surface of young leaves was dusted with a little amount of Celite 545 to scratch the leaves during inoculation. Virus infected leaves were grinded and homogenized with mortar and pestle in 0.03 M HEPES (pH 7.0). Subsequently the leaves were wiped with the virus buffer suspension. About 2 weeks later clear mosaic symptoms were developed. The infected plant leaves were used for virus purification as well as viral RNA and dsRNA extraction.

2.2.3.2. Aphid transmission

Aphids *(Myzus persicae)* previously starved for 30 min, were placed on leaves of BtMV-G infected *N. benthamiana* plants for 30 min for virus acquisition, thereafter they were collected and transferred in groups of 20-30 aphids to each assay healthy plant and allowed to stay 48 h before being killed with an insecticide. This experiment was performed for 3 times using ten plants each time. Once virus symptoms had appeared, each inoculated plant was tested by plate trapped ELISA.

2.2.3.3. Particle bombardment

High-velocity particle bombardment was used as a highly efficient system for plant infection using pe35BtMV a plasmid containing the full-length clone of BtMV-G with an enhanced *Cauliflower mosaic virus* (CaMV) 35S promoter and a CaMV terminator. The plasmid was bombarded directly onto young tobacco plants according to Gal-On *et al*. (1995) and Rasco-Gaunt *et al*. (1999) with a self-made particle gun device (Varrelmann, 2000).

Components required for particle bombardment

DNA (Plasmid approx. 1µg/µl)	3 µl
Tungsten particles	3 µl
1 M $Ca(NO_3)_2$, pH 10	3 µl

The components were mixed well by vortexing for 30 sec and incubated at room temperature for 5 min to precipitate the DNA to the tungsten particles. 6 µl of the supernatant were discarded and the sediment, which contains the DNA bound to the tungsten particles, was used for bombardment. 3 µl were transferred to the bombardment filter located in the bombardment chamber, the pressure was controlled to about 3 bar and the vacuum was set to 800 mbar.

2.2.4. Detection of *Beet mosaic virus*

2.2.4.1. Enzyme-linked Immunosorbent Assay (ELISA)

Plate trapped ELISA was used to detect BtMV-G following a standard ELISA procedure (Hampton *et al*., 1990). 500 mg leaf material was grinded with 500 µl coating and extraction buffer (PBS-TPO) in an 1.5 ml Eppendorf tube with a pestle fitting in the tube. A healthy plant sample was used as a negative control. The samples were centrifuged at 13,000 rpm for 1 min and 100 µl of the supernatant were used to coat one well of a polystyrene microtiter plate. The plate was incubated at 4^{o}C overnight, washed 3 times for 5 min each with PBS-T followed by adding 100 µl of the antibody (IgG) specific to BtMV which was diluted in PBS-TPO (1 µl/ml). The plate was incubated at 4^{o}C for 4 hours and again washed 3 times for 5 min each with PBS-T followed by adding 100 µl of the

conjugate, a second antibody (goat anti-rabbit; Sigma A3687) 1: 4000 (1 µl /4 ml) diluted in PBS-TPO. The plate was incubated at 37°C for 4 hours. After 3 washing steps (each for 3 min) 100 µl of the substrate solution (1 mg of p-Nitrophenylphosphate in 1 ml substrate buffer) was added to each microtiter well and incubated at room temperature. After 30 min the absorbance was measured with an ELISA-photometer at a wavelength of 405 nm.

Phosphate buffered saline (PBS): 10x PBS pH 7.4
80.0 g NaCl
2.0 g KCl
2.0 g KH_2PO_4
29.0 g Na_2HPO_4
Adjust pH add distilled water up to 1 liter and autoclave.

Washing buffer: Phosphate buffered saline-Tween (PBS-T)
100 ml 10x PBS
0.5 ml Tween 20
Add distilled water up to 1 liter.

Coating and extraction buffer (PBS-TPO), pH 7.4
100 ml 10x PBS
20 g Polyvinyl Pyrrolidone (PVP)
2.0 g Bovine serum albumin
0.5 ml Tween 20
Adjust pH add distilled water up to 1 liter.

Substrate buffer: pH 9.8
97 ml Diethanolamine
0.2 g $MgCl_2$ x 6 H_2O
Adjust pH add distilled water up to 1 liter.

2.2.4.2. RT–PCR (Reverse Transcription-Polymerase Chain Reaction)

RT-PCR was used to amplify a segment of RNA after converting the RNA into cDNA. Two primers flanking a known sequence region were used to amplify the cDNA.

Taq-DNA polymerase is extensively used to amplify DNA fragments. However, the error rate of the enzyme during synthesis of new DNA strands is quite high. In addition it also adds an A-residue at the end of both DNA strands. This A-overhang can be used for cloning of the DNA fragments into a so called T-vector. In cases were a high fidelity of the polymerase reaction is required (e.g. for amplification of fragments used for an infectious

full-length clone) enzymes with proof reading activity (e.g. *Pfu*-polymerase) were used in conjunction with *Taq*-polymerase.

10x PCR buffer
500 mM KCl
100 mM Tris-HCl (pH 8.3)
15 mM $MgCl_2$
The ingredients were dissolved in double distilled water and pH was adjusted to 8.3.

General RT-PCR reaction mix

Total RNA:	1 µl
10 x PCR buffer:	5 µl
$MgCl_2$ (25mM):	5 µl
Primer 1 (10 µM):	1 µl
Primer 2 (10 µM):	1 µl
dNTPs (25 mM):	0.5 µl
Taq-polymerase:	0.1 µl (5 u/µl)
Reverse Transcriptase:	0.05 µl (10 u/µl)

Double distilled water was added to a final volume of 50 µl.

Oligonucleotides used for RT-PCR and Immunocapture RT-PCR

No.	Name	Sequence
1	BtMVcol1	5'- GCTTTCTTTCCAGAAACGAGTAGTGC – 3'
2	BtMVcol2	5'-CGTGATATCCAATGTTTTGCTGCTA-3'
3	BtMVCPs	5'-AGCTATACATCCGAGTA-3'
4	BtMVCPas	5'-TGCTCGATTGTCCCACA-3'

If not otherwise stated the following PCR program was used:

Step	Action	Temperature	Time
1	Reverse transcription	42°C	45 min
2	Denaturation	94°C	4 min 30 sec
3	Denaturation	94°C	30 sec
4	Annealing	55°C	30 sec
5	Extension *	72°C	1 min 30 sec
6	Final extension	72°C	5 min

* 32 cycles programmed from step 5 back to step 3.

2.2.4.3. Immunocapture-RT-PCR

Immunocapture was carried out according to Lanneau and Candresse (1993).

Standard 0.5 ml micro Eppendorf tubes for PCR were coated with 100 µl IgG specific for BtMV in sodium carbonate coating buffer pH 9.6 (Clark and Adams, 1977) further diluted

(1:100 v/v). Tubes were incubated for 3 hours at 37°C before washing 2-3 times with PBS-T buffer (at this stage the tubes can be frozen for several weeks at –20°C).

The plant samples were prepared by grinding infected tissue in PBS-TPO buffer at a ratio of 1:500 in 1.5 ml Eppendorf tubes. After a brief centrifugation for 5 min 100 µl of clarified plant sap was added to a pre-coated tube and incubated overnight at 4°C or for 2-3 hours at 37°C. The plant extracts were removed before washing 2-3 times with PBS-T buffer.

RT- PCR was performed as a one-tube protocol:

1.7% TritonX100	3.0 µl
10x PCR buffer	2.5 µl
dNTPs (25 mM each)	0.25 µl
100 µM Primers (each)	0.25 µl
Reverse Transcriptase	0.25 u
Taq-DNA polymerase	0.5 u
Double distilled water	up to 25 µl.

The tubes were subjected to thermal cycling, consisting of one cycle of reverse transcription at 42°C for 45 min followed by 92°C for 2 min and linked straight to 40 cycles of PCR denaturation 92°C for 30 sec, annealing 55°C for 30 sec, extension 72°C for 1 min and 10 min for a final extension at 72°C.

Carbonate buffer: pH 9.6
2.93 g $NaHCO_3$
1.59 g Na_2CO_3
Dissolve in 1 litre distilled water.

2.2.5. Viral RNA extraction

2.2.5.1. Phenol RNA extraction method modified after Verwoerd *et al.* (1989)

250 mg of infected leaf material was homogenised with 500 µl of 0.03 M HEPES buffer and the paste was transferred into a safety Eppendorf tube containing a mixed solution of equal amounts (300 µl) of RNA extraction buffer and phenol, preheated for 5 min in a water bath at 80°C. After shaking the mixture vigorously, it was centrifuged for 5 min at 13000 rpm. The upper phase was transferred to a new Eppendorf tube. An equal amount of

chloroform/isoamyl alcohol (24:1) was added and mixed very well. Centrifugation of the solution was done for 5 min at 13000 rpm. The upper phase was transferred to a new tube. An equal amount of 4M LiCl was added and incubated overnight at 4^{o}C. The solution was centrifuged for 20 min at 13000 rpm at 4^{o}C. The supernatant was discarded and the pellet was dried in a speedvac. The pellet was redissolved in 190 µl of RNase free water. Then 5 µl of RQ1 DNase (1u/µl) and 5 µl of 0.3 M $MgCl_2$ were added and the solution was incubated for 30 min at 37^{o}C. Afterwards 200 µl of proteinase K buffer and 20 µl of proteinase K (20 mg/ml) were added and the solution was mixed and incubated for 30 min at 37^{o}C. 420 µl of phenol was added to the solution, mixed vigorously and centrifuged for 5 min at 13000 rpm. The upper phase was transferred to a new Eppendorf tube and an equal amount of chloroform was added. After shaking well the mixture was centrifuged for 5 min at 13000 rpm. The upper phase was transferred to a new Eppendorf tube and 0.1 volumes of 3M NaOAc (pH 5.5) and 2.5 volumes of ice cold abs. EtOH were added to the solution and incubated for 20 min at -20^{o}C for precipitation. After centrifugation for 15 min at 4^{o}C and 13000 rpm the precipitate was recovered after discarding the supernatant. The pellet was washed with 70% EtOH and dried in a speedvac. The pellet was dissolved in 50 µl of TE buffer (10/0.1). The amount, integrity and size of the total RNAs were estimated by agarose gel electrophoresis.

RNA extraction buffer
0.1 M LiCl
100 mM Tris
10 mM EDTA
1% SDS
Adjust pH 8.0 and autoclave.

RNase free water
2 ml of DEPC (Diethylpyrocarbonate) was added to 1 L of double distilled water and kept overnight at room temperature with continuous stirring. The DEPC was destroyed by autoclaving.

4 M LiCl solution; 3 M Sodium acetate solution, pH 5.2; 0.03 M HEPES, pH 7.0

TE buffer (10/0.1)
10 mM Tris
0.1 mM EDTA
Adjust pH 7.5 and autoclave.

2.2.5.2. dsRNA extraction method according to Morris and Dodds (1979)

10 g of infected leaf material was grinded with liquid nitrogen, the powdered plant material was added to a mixture of 20 ml 2x STE buffer, 3 ml SDS (10%), 20 ml chloroform, 0.4 g PVPP and 0.4 ml ß-Mercaptoethanol solution. The mixture was stirred for 30 min and than centrifuged for 30 min at 6000 rpm at 4°C. 80% of the upper phase was collected and filled up to a final volume of 20 ml with 1x STE buffer. Then 2 g CF11 Cellulose and 4 ml of abs. EtOH was added slowly and mixed by stirring well for 10 min to allow dsRNA to bind to the cellulose. The suspension was centrifuged for 2 min and the cellulose pellet was washed 4 times with 40 ml STE/16.5% EtOH followed by centrifugation for 2 min after each washing step. After the last washing step 20 ml of 1xSTE/16.5% EtOH was added and mixed well. The solution was transferred into a column containing a cotton plug at the bottom. The column was washed with 100 ml 1xSTE/16.5% EtOH. 13 ml 1x STE buffer was added to elute the dsRNA. The first 5 drops of the elution liquid were discarded and the rest was collected in a beaker. 2.5 volume of abs. EtOH and 0.05 volumes of 3 M NaOAc (pH 5.5) were added followed by incubation on ice for 15 min. Afterwards the solution was centrifuged for 30-60 min at 6000 rpm at 4°C. The supernatant was discarded and the pellet was dissolved in 0.4 ml of 1x DNase buffer. The solution was transferred to an Eppendorf tube and centrifuged for 5 min at 13000 rpm to remove residues of cellulose powder. The supernatant was transferred to a new Eppendorf tube and 15 µl of RQ1 DNase **(**1u/µl**)** was added and mixed well followed by incubation for 30 min at 37°C. 200 µl of 1M $MgCl_2$ and 2 µl of RNase (10mg/ml) were added and mixed well followed by incubation for 30 min at 37 °C. 250 µl of phenol was added, mixed well and centrifuged for 5 min at 13000 rpm. The upper phase was collected carefully and 250 µl of chloroform was added and mixed well followed by centrifugation for 5 min at 13000 rpm. The upper phase was collected carefully, 2.5 volume of abs. EtOH and 0.1 volumes of 3 M NaOAc (pH 5.5) were added and centrifuged for 30 min at 4°C at 13000 rpm. The supernatant was discarded and the pellet was washed with 150 µl of 70% EtOH and dried in a speedvac. The pellet was dissolved in 30 µl of TE 10/0.1 buffer. The amount and the size of the dsRNA were estimated by gel electrophoresis.

10x STE buffer
0.1 M NaCl
0.1 M Tris-base
0.001 M EDTA
Adjust pH 7.0 and autoclave.

2.2.5.3. Partial purification of BtMV-G

The method of Rogov *et al.* (1991) was modified to be used in this study to purify BtMV-G for obtaining high virus quality and high viral RNA concentration for RT-PCR amplification. The purification was done at the Institute of Biochemistry and Plant Virology (BBA) as follows:

100 g of systemically infected fresh leaves (1:4 w/v) were homogenised in a blender with 400 ml of ice cold extraction buffer 0.1 M sodium acetate buffer pH 7.0 containing 0.02 M EDTA and 0.02 M sodium diethyldithiocarbamate (DIECA). 20% n-butanol (v:v) was added and blended for 30 sec. The paste was filtered through three layers of cheesecloth to remove plant debris. The plant sap was collected and transferred to centrifuge tubes. Centrifugation of the solution was done for 15 min at low speed of 8000 rpm at 4^{o}C. The aqueous upper phase was transferred into new tubes and agitated with 1.5% Triton X-100 for one hour on ice, afterwards a low-speed centrifugation of the extract was done at 8000 rpm for 15 min. The virus was pelleted from the supernatant by high-speed centrifugation for 90 min at 32 000 rpm at 4^{o}C. The supernatant was discarded and the pellet was dissolved in 0.01 M sodium citrate buffer pH 7.5 and clarified by low speed centrifugation at 10000 rpm for 10 min. The supernatant was transferred to new centrifuge tubes and another round of high-speed centrifugation was done at 32000 rpm for 90 min. The final pellet was dissolved in Aqua dest. and kept at -20°C until extraction of RNA.

2.2.5.4. Viral RNA extraction from purified virus

Viral RNA was deproteinized by mixing equal amounts of purified virus (150 µl) with proteinase K buffer and 10 µl proteinase K (20mg/ml). After incubation for 30 min at 37^{o}C an equal amount of phenol (310 µl) was added and the solution was mixed well. Centrifugation of the solution was done for 5 min at 13000 rpm. The upper phase was

transferred to a new tube and an equal amount of chloroform was added. After centrifugation for 5 min the upper phase was carefully transferred into a new tube and 2.5 volumes of abs. EtOH and 0.1 volumes of 3 M sodium acetate (pH 5.5) were added and mixed well. After centrifugation at 4^{o}C for 30 min the pellet was washed by adding 70% EtOH, followed by another centrifugation at 4^{o}C for 10 min. The final pellet was dried in a speedvac and dissolved in 50 µl TE-buffer.

Proteinase K-buffer
100 mM Tris-HCl
12.5 mM EDTA
150 mM NaCl
1 % (w/v) SDS
Add 1 liter distilled water.

2.2.6. Virus measurements

Partially purified virus was analysed by electron microscopy and gel electrophoresis (SDS-PAGE). In addition the absorbance of the virus suspension was measured in a spectrophotometer for assessing the purity of the virus preparation and for estimating its concentration.

2.2.6.1. Estimation of virus concentration

The light absorbance of a virus suspension can be measured in a spectrophotometer. The UV absorbance in a cuvette containing the virus suspension in a suitable buffer is compared with the absorbance by a similar cuvette with the same buffer without the virus.

The yield of purified virus was estimated spectrophytometrically in the spectrophotometer using the following formula:

Virus yield (mg/ml) = O.D. $_{(260nm)}$ × Vol / Ext. coeff

O.D. $_{(260nm)}$: Optical Density units of purified virus suspension at 260 nm
Vol: volume of purified virus suspension (ml)
Ext. coeff: Extinction coefficient (the extinction coefficient $A_{260,1cm}$ = 2.6) (Fujisawa *et al.*, 1983)

2.2.6.2. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

In order to determine the molecular weight of the coat protein of the purified BtMV-G samples from the partially purified virus preparations were analysed by SDS-polyacrylamide gel electrophoresis (PAGE) using a 4% stacking gel on a 12% polyacrylamide gel. The capsid protein size was compared with a standard protein molecular weight marker.

The partially purified virus preparation was mixed with loading buffer (1:1;v:v) in a safety Eppendorf tube. Samples were boiled at 97^{o}C for 5 min to denature the proteins and immediately placed on ice. 20 µl of each sample were loaded and electrophoresis was first carried out at 80 V until the Bromphenol Blue front reached the resolving gel and continued at 120 V until it reached the bottom of the gel. The protein bands were observed after staining the gel with Coomassie Brilliant Blue for about 3 h and removing the excess stain by immersing the gel for 2 h in destaining solution, whereas the solution was changed each 20 min. The molecular weight of the viral protein was determined by comparison with known molecular weights of marker proteins.

Solutions required for SDS-PAGE:

<u>30% polyacrylamide solution (PAA)</u>
29 g acrylamide
1g N'N'-bis-acrylamide
Dissolve in 100 ml of distilled H_2O and keep at 4^{o}C in the dark.

Resolving gel buffer (4x RGB): 1.5 M Tris-HCl, pH 8.8. (18.15g /100ml)

TEMED: N,N,N'N'-tetramethylethlenediamine

Stacking gel buffer (4x SGB): 0.5 M Tris-HCl, PH 6.8. (3g / 50ml)

10% SDS: 10 g/ 100 ml;

Freshly prepared 10% ammonium persulfate $(NH_4)_2S_2O_3$ (APS) solution: 1g /ml

<u>Electrophoresis buffer</u>
25 mM Tris base, 192 mM glycine, 0.1% (w/v) SDS, pH 8.3

2x sample buffer, pH 6.8 (Laemmli, 1970)

Stacking gel buffer	20 ml
10 SDS	12.5 ml
Glycerol	10 ml
1% Bromophenol blue	2.5 ml
ß-Mercaptoethanol	5 ml

Staining solution (Coomassie Brillant Blue)

Coomassie Brilliant Blue R-250 (0.1%)	1 g
Methanol (50%)	500 ml
Acetic acid (10%)	100 ml
Deionised distilled H_2O	400 ml

Destaining solution

Methanol (25%)	250 ml
Acetic acid (10%)	100 ml
Deionised distilled H_2O	650 ml

SDS-PAGE preparation (12% resolving gel and 4% stacking gel)

Solutions	30 ml resolving gel (12%)	15 ml stacking gel (4%)
30 PAA	12 ml	2 ml
4x RGB	7.5 ml	-
4x SGB	-	3.75 ml
10% (w/v) SDS	300 µl	150 µl
10% (w/v) APS	300 µl	150 µl
TEMED	30 µl	15 µl
Dist. H_2O	9.9 µl	8.95 µl

2.2.7. Amplification and cloning of viral RNA

2.2.7.1. Random hexamer PCR (RH-PCR)

The method of Froussard (1992) was modified and used for amplification of unknown sequences of genomic BtMV-G RNA. Random hexamer PCR (RH-PCR) can be used to amplify a whole cDNA population, from 5’ to 3’ ends, derived from low amounts of RNA using an oligonucleotide primer containing a random hexamer sequence at its 3’ end (UN6-P: 5’-GCCGGAGCTCTGCAGAATTCNNNNNN-3’). The six undefined nucleotides (random hexamer) enable this primer to anneal at different positions of a template cDNA and leads to different fragment sizes in subsequent PCR amplification. Firstly, cDNA was synthesized by using Reverse transcriptase and a gene specific antisense primer (designed

according to the known sequence, see Tab. 1). Second strand synthesis was performed by using the UN6-primer (see 2.2.7.1.4.) and Klenow Fragment exo$^-$ (Exonuclease minus). This enzyme is a proteolytic product of DNA polymerase I, which retains polymerase activity, but has a mutation (D355A; E357A), which abolishes the 3´-->5´ exonuclease activity.

2.2.7.1.1. First strand cDNA synthesis

First strand cDNA was synthesized by mixing 8 µl pure viral RNA (0.28 µg/µl) with 2 µl gene-specific antisense-primer, e.g. BtMVcol2 (10 µM; Tab. 1; No.1 of Step 1) in 20 µl of double distilled water. The mixture was incubated at 99°C for 5 min and rapidly cooled on ice before 6.5 µl distilled water, 10 µl 5x reverse transcription buffer, 2 µl dNTPs (25 mM) and 1 µl AMV RT Reverse transcriptase (10 u/µl) were added and incubated at 42°C for 1 h. The reaction was then boiled at 99°C for 2 min and rapidly cooled on ice.

2.2.7.1.2. Second strand cDNA synthesis

Second strand cDNA (dscDNA) synthesis was done by adding 34 µl double distilled water, 10 µl 10x Klenow buffer, 2 µl dNTPs (25 mM), 2 µl Klenow Fragment (5 u/µl) and 2 µl UN6-primer (10 µM) to the first strand cDNA reaction. After careful mixing the reaction was incubated for 1h at 37°C. Subsequently, double stranded cDNA was purified by using the PCR purification kit (Qiagen) according to the manufacturer instructions to eliminate excess primers, enzymes and reaction buffers.

2.2.7.1.3. Purification of double stranded cDNA fragments

The QIAquick PCR purification protocol was used to purify dscDNA fragments from reaction materials before amplification of the target dscDNA fragments. 5 volumes of PB buffer were added to one volume of the reaction mixture containing the dscDNA and mixed carefully. The solution was applied to a QIAquick spin column and placed in a provided 2 ml collection tube, followed by centrifugation for 30-60 sec at 13000 rpm. The flow-through was discarded and the QIAquick spin column was placed back into the same tube.

Then 0.75 ml washing buffer (PE) was added and centrifuged for 30-60 sec at 13000 rpm. The flow-through was discarded and the QIAquick spin column was centrifuged for another minute to remove residual liquid of the washing buffer (PE). The QIAquick column was placed in a clean 1.5 ml microcentrifuge tube, 25 µl elution buffer was added to the center of the QIAquick membrane to elute dscDNA. The tube was held for 1 min at room temperature to increase the release of dscDNA, then centrifugation was done for 1 min. The flow-through was pipetted into a new 1.5 ml tube.

2.2.7.1.4. Amplification of double stranded cDNA fragments

Amplification of the synthesised double strand cDNA population was performed by mixing 26.5 µl double distilled water, 5 µl 10x PCR buffer, 5 µl $MgCl_2$ (100 mM), 1 µl dNTPs (25 mM), 0.5 µl *Taq* polymerase (1 u/µl) + 1 µl Universal primer (10 µM) (5'-GCCGGAGCTCTGCAGAATTC-3'), 1 µl gene specific antisense primer, e.g. BtMV5 (10 µM; Tab. 1; No. 2 of Step 1) and 10 µl purified dscDNA. The universal primer (UN-P), which was used in this step is identical to the UN6-primer at its 5'-end, but has no random hexamer at its 3'-end. Therefore it is able to anneal at the 20 nucleotides incorporated by UN6-P in the synthesis of double strand cDNA.

The tubes were subjected to thermal cycling consisting of one cycle for 2 min at 94°C and linked straight to 40 cycles of PCR (denaturation: 94°C for 20 sec, annealing: 55°C for 30 sec extension: 72°C for 3 min and 7 min for the final cycle. Final amplification products were analysed on agarose gels.

This method was applied four times to resolve most of the unknown BtMV-G sequence by using two gene-specific antisense primers (GSPs) and the two universal primers for each time. The GSPs were constructed from the known downstream sequence (Tab. 1). The first primer (1.) was always used for synthesizing cDNA and the second one (2.) was used for the PCR amplification step. This method was also used with total RNA from BtMV-G infected plants, extracted by the hot phenol extraction method (2.2.5.1).

Table 1: Oligonucleotides (antisense primers) used for gene walking to determine the complete nucleotide sequence of the RNA genome of BtMV-G. Oligonucleotides were synthesiszed by MWG-Biotech or Invitrogen.

Primer	Step	Sequence
1. BtMVcol2	1.	5'-CGTGATATCCAATGTTTTGCTGCTA-3'
2. BtMV5		5'-AGTCTAGAGCACATGATACCCAGTTGTCAA-3'
1. BtMVcol5	2.	5'-ATCCCTTTGATTGTCACACAAAAGCGACGGA-3'
2. BtMV2		5'-AATCTAGAGCTTTGTTTGAACGTTTATTTCTAA-3'
1. BtMVcol6	3.	5'-CCTTCCTCTTACTATTACCTCATCATA-3'
2. BtMV3		5'-CGCATGTTGATGCAATGTTCTTGGGCAGAA-3'
1. BtMVcol7	4	5'-AATTCCCGCTGTGATGTCACCTTCCAA-3'
2. BtMV4		5'-ATTCCTTCCTCTTACTATTATGGTGGGGTAA-3'

2.2.7.2. Amplification of the 5'ends of BtMV-G RNA

2.2.7.2.1. Adaptor ligation and PCR amplification

This method was modified according to Potgieter *et al.* (2002) to determine the extreme 5'-end of the BtMV-RNA.

2.2.7.2.1.1. cDNA synthesis and purification

1 µl pure viral RNA (0.28 µg) and 1 µl gene-specific antisense-primer BtMVcol6 (10 µM; Tab. 1) were mixed with 16 µl double distilled water, incubated at 99°C for 5 min and rapidly cooled on ice.

5 µl 5x Reverse transcription buffer, 1 µl dNTPs (25 mM), 1 µl AMV RT Reverse transcriptase (10 u/µl) were added and incubated at 42°C for 1 h. The cDNA was precipitated by adding 2.5 volume of abs. EtOH and 0.1 volumes of 3 M NaOAc (pH 5.5) and centrifuged for 10 min at 4^{o}C at 13000 rpm. The supernatant was discarded and the pellet was washed twice with 150 µl of 70% EtOH and dried in speedvac. The pellet was dissolved in 6 µl double distilled water.

2.2.7.2.1.2. Ligation of an adaptor oligonucleotide

6 µl cDNA was mixed with 2 µl of an adaptor oligonucleotide (100 µM) carrying a phosphate residue at its 5'-end and an amino group at its 3'-end (5'-P-GGATCCCGGGA-ATTCGGTACCTAAAAAAAAAAA-NH_2-3') and incubated over night at 4°C in the presence of 1 µl T4 ligase and 1 µl ligase buffer. Afterwards 1-2 µl of the ligation mixture was used for PCR amplification.

2.2.7.2.1.3. PCR amplification

An aliquot of the ligation mixture (1-2 µl) was directly subjected to PCR. The reaction mixture contains in a final volume of 50 µl: 1 µl gene specific antisense primer col7 (10 µM; Tab. 1), 1 µl primer (5'-AGGTACCGAATTCCCGGGATCC-3') complementary to the used adaptor oligonucleotide, 5 µl PCR buffer, 5 µl $MgCl_2$, 0.5 µl dNTPs (25 mM), Taq DNA polymerase and double distilled water. PCR was performed with denaturation at 94°C for 2 min, followed by 3 cycles of denaturation at 94°C for 30 sec, annealing at 37°C for 30 sec and primer extension at 72°C for 1.30 min. Subsequently 36 cycles were performed with denaturation and extension conditions as stated above but with annealing at 55°C for 30 sec followed by a final extension step at 72°C for 10 min.

2.2.7.2.2. Poly-G Tailing and amplification

2.2.7.2.2.1. First strand cDNA synthesis and cDNA purification

First strand cDNA was prepared by mixing 1 µl pure viral RNA (0.28 µg) and 1 µl gene-specific antisense-primer BtMVcol8 (10 µM; 5'-AAGAGTATATAATA-GGTCGAGCAA-3') were mixed with 6 µl of distilled water and incubated at 95°C for 5 min rapidly cooled on ice.

To the reaction mixture (8 µl) were added: 4 µl 5x reverse transcription buffer, 2 µl 0.1 M DTT, 1 µl dNTPs, 1 µl Superscript II Reverse Transcriptase (200 u/µl). The reaction was incubated at 42°C for 1 h.

After adding 4 µl freshly prepared $MnCl_2$ solution (10 mM) the reaction was incubated for 15 min at 42°C. 1 µl NaOH (1M) was added and incubated for 20 min at 70°C.

2.5 volumes of abs. EtOH and 0.1 volumes of 3 M NaOAc (pH 5.5) were added and centrifuged for 20 min at 4°C at 13000 rpm. The supernatant was discarded and the pellet was washed twice with 150 µl of 70% EtOH and dried in speedvac. The pellet was dissolved in 10 µl double distilled water. Afterwards 5 µl cDNA was used for poly G tail construction.

2.2.7.2.2.2. Poly G tailing and PCR amplification

The poly-G tailing approach was used for amplification of the cDNA 5'end as described by Götz and Maiss (2002).

G-tailing was performed by using 5 µl cDNA, 5 µl 5x NEB Buffer 4 (50 mM potassium acetate, 20 mM Tris-acetate 10 mM Magnesium acetate, 1 mM dithiothreitol), 5 µl $CoCl_2$, 1 µl TDT, 1 µl dGTP (10 mM) and 33 µl double distilled water. The components were mixed and incubated at 37°C for 15 min and then rapidly incubated at 70°C for 10 min to stop the reaction.

2.5 volume of abs. EtOH and 0.1 volumes of 3 M NaOAc (pH 5.5) were added and centrifuged for 20 min at 4°C at 13000 rpm. The supernatant was discarded and the pellet was washed twice with 150 µl of 70% EtOH and dried in speedvac. The pellet was dissolved in 10 µl double distilled water.

1 µl of G-tailed cDNA was used in the following PCR reaction. The PCR reaction mixture (50 µl) was prepared containing 5 µl PCR buffer, 5 µl $MgCl_2$, 1 µl of a gene specific antisense-primer BtMV5as (10 µM; 5'-AGGAGGCCATCATTTCTTTC-GTAACCAA-3'), 1 µl poly C primer (5'-AATCTAGACCCCCCCCCCCCCC-3'; 10 µM), 0.5 µl dNTPs (25 mM), 0.1 µl *Taq* DNA polymerase and 36.4 µl double distilled water. The PCR reaction was carried out as described under 2.2.7.2.1.3.

2.2.8. Agarose gel electrophoresis

Two kinds of agarose gels were used. Standard agarose gels were used to analyse RNA, PCR fragments or plasmids after control digestion. The preparation of the gel was carried out by dissolving 0.5 g agarose in 50 ml of 1x TAE buffer. The solution was boiled using a microwave oven until the agarose dissolves completely in the buffer, cooled down and supplemented with 1.5 µl ethidium bromide (10 mg/ml). Subsequently the mixture was poured into a tray and a comb was placed in the gel to form slots. The gel was placed in the electrophoresis unit, filled with electrophoresis buffer (1X TAE) to cover the gel and comb was carefully removed. 10 µl of nucleic acid containing samples were mixed with 5 µl gel-loading buffer and carefully loaded into one slot. 3-5 µl of a DNA size standard (*Pst* I digested λ-DNA) was placed in an slot to ensure later size comparison of DNA fragments.

Electrophoresis was performed for the first 5 minutes at 80 V and then at 120 V for about 30-60 minutes. The nucleic acids were visualised by ultraviolet light and a photograph was taken.

Low-melting agarose (LM) was used to separate DNA fragments for further purification and cloning steps. LM-gel preparation was done as agarose gel preparation, but electrophoresis was carried out at 60 V for the first 5 min followed by 80 V for 30-90 min.

<u>50x TAE-Electrophoresis-buffer</u>

Tris-base	242 g
Acetic acid 99.8%	57.1 ml
0.5M EDTA (pH 8.0)	100 ml

add to 1 litre with distilled water.

<u>Loading buffer</u>

Distilled water	3.44 ml
Glycerine	6.54 ml
Bromphenol blue	small amount (spatula tip)

<u>Ethidium bromide solution</u>

Ethidium bromide 10 mg /ml dist. water

λ-*Pst* I DNA size standard

λ-DNA (0.4 mg/ ml)	140 µl
Pst I (10 u/µl)	8 µl
10x buffer (MBI orange)	30 µl
Distilled water	122 µl

The reaction mixture was incubated for 3 h at 37°C followed by incubation for 15 min at 70°C in the heating block. Finally 100 µl loading buffer was added.

2.2.9. Purification of nucleic acids from low-melting agarose gels

Purification of DNA fragments was done using the QIAEX II Agarose Gel Extraction protocol according to the manufacturer instructions. Briefly, the DNA band was excised under UV light from ethidium bromide-stained low melting point agarose gels with a sterile and sharp scalpel and transferred to an Eppendorf tube. 300 µl of buffer QX1 was added to dissolve the gel, which contains DNA fragment less than 4 kb. QIAEX II solution (glass milk) was resuspended by vortexing for 30 sec and 8 µl QIAEX II solution was added. The solution was incubated for 10 min at 50°C in a heating block to solubilize the agarose and bind the DNA to the glass milk. The solution was mixed every 2 min to keep the QIAEX II in suspension. The solution was centrifuged for 30 sec at 13000 rpm and the supernatant was carefully removed by a pipette. The pellet was resuspended in 500 µl of buffer QX1 followed by centrifugation for 30 sec at 13000 rpm and the supernatant was removed by a pipette. The pellet was washed two times with 500 of buffer PE. Centrifugation was done for 30 sec with two washing steps. The residual supernatant was carefully removed with a pipette. The pellet was air dried for 10-15 min until it became white. 26 µl of 10 mM Tris-Cl (pH 8.5) was added to elute the DNA from the glass milk followed by incubation according to the size of DNA fragment.

Size of the DNA fragment	**Incubation temperature**	**Incubation period**
DNA fragment 4 kb	Room temperature	5 min
DNA fragment 4-10 kb	50°C	5 min
DNA fragment > 10 kb	50°C	10 min

The solution was centrifuged for 30 sec and the supernatant was carefully transferred to a new Eppendorf tube avoiding the carry over of any glass milk.

2.2.10. Cloning and analysis of DNA fragments

2.2.10.1. Ligation of DNA fragments into pGEM®T-Easy vector

The pGEM®T-Easy vector (Promega) is a convenient system for cloning of PCR fragments. The vector is prepared by digesting it with *Eco*R V to generate blunt ends. Subsequently terminal thymidine residues are attached to the 3'-ends at both ends, e.g. with Taq-Polymerase. These single 3'-T overhangs at the insertion site greatly improve the efficiency of ligation of a PCR product into the plasmid vector by providing compatible ends to the insert, which carries adenosine residues at its ends. In addition the pGEM®T-Easy vector has a multiple cloning site allowing easy characterisation of the insert by digestion with several restriction enzymes (Fig. 2.).

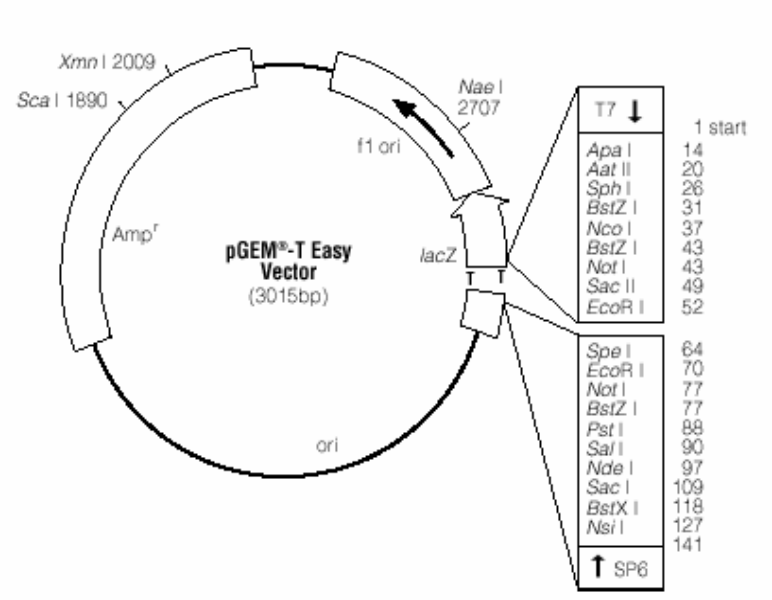

Figure 2: pGEM®T-Easy Vector

After purification of the PCR product, it was directly ligated to the cloning vector in a ligation mixture. Ligation was done to introduce the ligated vector into *E. coli* to increase the DNA fragment in multiple folds.

2 X Ligation buffer
60 mM Tris - HCl (pH 7,8)
20 mM MgCl2
20 mM DTT
2 mM ATP
10% Polyethylene Glycol MW 6000

Ligation mixture

Ligation buffer	5 µl
DNA fragment	3 µl
pGEM-T Easy vector (10 ng/µl)	0.5 µl
T4-Ligase	1 µl
A.bidest	0.5 µl

The ligation mixture was incubated overnight at 4°C.

2.2.10.2. Preparation of competent *Escherichia coli* NM 522 cells

A protocol according to Hanahan (1983) was used for preparation of competent *E. coli* NM522 cells. A few bacterial cells from a stock culture were incubated at 37^{o}C with shaking overnight in 30 ml LB. 1 ml of the overnight culture was transferred to 30 ml SOB supplemented with 300 µl 2 M Mg^{++}. The bacteria were incubated at 37^{o}C with vigorously shaking at 180 rpm for 1-1.5 h until the OD_{550nm} reached 0.45- 0.55. The suspension was transferred to 30 ml sterile ice cold centrifuge tubes, which were placed on ice for 10 min to stop cell division. The suspension was centrifuged for 12 min at 1000 rpm at 4^{o}C. The supernatant was discarded and the pellet was resuspended in 10 ml TFB buffer and incubated on ice for 10 min, centrifuged again for 12 min at 1000 rpm at 4^{o}C. The supernatant was discarded and the pellet was resuspended by adding 4 ml of TFB buffer and 140 µl DND. After incubation on ice for 15 min another 140 µl DND was added and the mixture was incubated on ice for additional 15 min. The competent bacterial cells were directly used for transformation.

SOB Medium (1 litre)

Select Peptone 140	10 g
Select Yeast Extract	5 g
NaCl (10 mM)	0.584 g
KCl (2.5 mM)	0.186 g

Distilled water was added to 1L and the pH was adjusted to 6.0-7.05 followed by autoclaving.

2 M Mg^{2+} solution

$MgCl_2$ x 6 H_2O	1M
$MgSO_4$ x 7 H_2O	1M

DND (10 ml)

DTT (Dithiothreitol)	1.53 g
DMSO (Dimethylsulphoxide)	9 ml
1M KAc (pH 7.5)	100 µl

TFB-transformation buffer

R-MES	10 mM, pH 6.3
RbCl	100 mM
$MnCl_2$ $4H_2O$	45 mM
$CaCl_2$ $2H_2O$	10 mM
$HACoCl_3$ (Hexamine cobalt chloride)	3 mM

2.2.10.3. Transformation of *E.coli*

200 µl of competent cells were gently mixed with 10 µl of the ligation mix in a 1.5 ml Eppendorf tube, then the cell-DNA mixture was mixed carefully by flicking the tube before incubating it on ice for 30 min. Subsequently then tube was incubated at 42^{o}C in a water bath for exactly 90 sec. The tube was placed immediately on ice for 2 min to allow chilling of the cells. 500 µl LB medium was added and incubated for 60 min at 37^{o}C with shaking to allow transformed bacteria to recover and express the antibiotic resistance gene. Then 100 µl aliquots of each transformation reaction were plated on LB-agar plates containing the appropriate antibiotic (see below) and 47 mg/L IPTG and 40 mg/L X-gal. The plates were incubated (upside down) at 37^{o}C overnight.

LB (Luria-Bertoni) Medium (modification of Sambrook *et al*., 1989)

Agar	15 g
Select Peptone 140	10 g
Yeast Extract	5 g
NaCl	10 g

Distilled water was added to 1L and the pH was adjusted to 7.5 followed by autoclaving. Plates were stored at 4^{o}C. LB liquid medium lacks the agar.

LB-AP Medium

Ampicillin	150 mg/L

Ampicillin was dissolved in 1 ml of autoclaved double distilled water and then added to 1 L of autoclaved LB medium (at 50^{o}C)

LB-Kn Medium
Kanamycin 50 mg/L
Kanamycin was dissolved in 1 ml of autoclaved double distilled water and then added to 1 L of autoclaved LB medium (at 50°C).

2.2.10.4. Isolation of recombinant plasmid DNA from transformed *E. coli*

Extraction and purification of plasmids was done according to Birnboim and Doly (1979). For isolation of plasmid DNA from transformed *E. coli*, a single colony was picked up by a toothpick from the plate (2.2.10.3) and inoculated into 3 ml LB-liquid medium containing the appropriate selective antibiotic. The bacteria were incubated for 6-8 hours at 37°C with vigorous shacking (180 rpm), afterwards, 2 ml of the bacteria suspension was transferred to a 2 ml Eppendorf tube. The bacterial cells were harvested by centrifugation for 2 min at 12000 rpm at room temperature. The supernatant was removed completely and 200 µl of solution A was used to resuspend the pellet, followed by an incubation for 10-15 min at room temperature. Subsequently 400 µl of solution B was added and mixed gently. The suspension should become clear and 300 µl of solution C was added and mixed carefully by inverting the tubes 5-6 times. The tubes were placed immediately on ice for 15-30 min followed by centrifugation for 13000 rpm at room temperature for 10 min. The supernatant was decanted into a new 1.5 ml microfuge tube and centrifugation was repeated again with the same conditions as before. The supernatant was transferred again to a new 1.5 ml tube and 600 µl isopropanol (-20°C) was added and mixed carefully. Centrifugation was done for 10 min at 13000 rpm at room temperature and the supernatant was removed completely. The pellet was resuspended in 200 µl of solution D by incubating the sample for 5 min at room temperature and gently flicking the tube. For precipitation, 400 µl of EtOH (-20°C) was added and mixed. The sample was centrifuged for 10 min at 13000 rpm at room temperature. The supernatant was removed completely and the pellet was dried in a speedvac for 5-15 min. Finally the pellet was redissolved in 60 µl of TE-RNase buffer and stored at –20°C.

Solution A
Tris HCl 15 mM, pH 8.0
EDTA 10 mM
Glucose 50 mM

Solution B
NaOH 0.2 mM
SDS 1%

Solution C
NaOAc 3 M, pH 4.8

Solution D
NaOAc 0.1 M, pH 5.5
Tris-HCl 0.05 M, pH 8.0

RNase/TE buffer
RNase 10 µg/ml TE 10/0.1, pH 8.0

2.2.10.5. Restriction enzyme digest of recombinant DNA clones

Digestion of plasmid DNA was done by using various restriction endonucleases. The enzyme mostly used was *Eco* RI for single digestion of the pGEM®T-Easy vector. Other enzymes were used in double digestions and also for subcloning steps and for generation of the full-length clone (2.1.8.).

After isolation of plasmid DNA (as described in section 2.2.10.4.) aliquots were used for preparation of the digestion mixture:

DNA (lysate or purified) 2-6 µl
Appropriate buffer: 1-4 µl
Restriction enzyme: 1.0 µl
Double distilled water: final volume of 15 µl
The mixture was incubated at 37°C for 1-2 hours.

Digested DNA was separated by agarose gel electrophoresis and the fragments were visualised by ethidium bromide staining and UV irradiation according to the procedure described in section 2.2.8. The plasmid mini preparations containing DNA fragments of the expected size were selected for further analysis.

2.2.10.6. Colony and plasmid PCR

E.coli transformants were picked from plates and each was resuspended in 100 µl of LB by vortexing for 10 sec and 1 µl was used for PCR .

In addition 0.2 µl plasmid mini preparations were also analysed for the presence of DNA inserts. A standard PCR was carried out using the vector primer and GSP flanking the expected fragment. PCR reaction mix was prepared as the following:

10x PCR buffer	5 µl	
$MgCl_2$ (25mM)	5 µl	
Primer 1 (10 µM)	1 µl	M13R: 5'-GGAAACAGCTATGACCATG-3'
or	1 µl	M13F: 5'-GTAAAACGACGGCCAGT-3'
Primer 2 (10 µM)	1 µl	(GSP)
dNTPs (25 mM)	0.5 µl	
Taq-Polymerase	0.1µl (5 u/µl)	
Bacterial suspension	1 µl	
or	0.2 µl plasmid DNA of mini preparation	

Double distilled water was added to a final volume of 50 µl.

The PCR was performed with denaturation at 94^{o}C for 4.5 min, followed by 40 cycles of denaturation at 94^{o}C for 30 sec, annealing at 55^{o}C for 30 sec, primer extension at 72^{o}C for 1 min and a final extension step at 72^{o}C for 10 min.

10 µl of the PCR products were analysed by electrophoresis in a 1% agarose gel (2.2.8) for the presence of DNA fragments.

2.2.10.7. Preparations of permanent cultures

Permanent cultures of bacterial cells were prepared for clones, which gave positive results after analysis for presence of DNA fragment. Small amount of bacterial cells were mixed with glycerine by using 500 µl bacterial cells and 675 µl sterile glycerin (86%). The suspension was vortexed and stored at -20^{o}C.

2.2.10.8. Purification and precipitation of plasmid DNA

Plasmid DNA from mini preparations was purified prior sequencing according to a modified protocol of the QIAprep Spin Miniprep Kit (Qiagen). For this purpose 5 volumes of PB buffer were added to the extracted plasmid (2.2.10.4.) and mixed carefully, the mixture was transferred by pipetting to the QIAprep spin column which was placed in a 2 ml collection tube, followed by centrifugation for 30-60 sec. The flow-through was discarded and 0.75 ml washing buffer (PE) was added and centrifuged for 30-60 sec for

washing the QIAprep spin column. The flow-through was discarded and the QIAprep spin column was centrifuged for additional one minute to remove residual wash buffer. The QIAprep column was placed in a clean 1.5 ml Eppendorf tube, 50 µl elution buffer was added to the center of the QIAprep column. After incubation at room temperature for 1 min centrifugation was done for 1 min and the flow-through was pipetted into a new 1.5 ml tube.

DNA Precipitation was done by adding 2.5 volumes of ice cold abs. EtOH (-20^{o}C) and 0.1 volumes of 3M sodium acetate (NaOAc, pH 5.5). The mixture was centrifuged at 13000 rpm for 20 min at 4^{o}C and the supernatant was discarded. The pellet was washed by adding 300 µl 70% EtOH (-20^{o}C) and centrifuged at 13000 for 10 min at 4^{o}C. The supernatant was discarded and the pellet was dried for 3-5 min either at room temperature or in a speedvac.

2.2.10.9. Sequencing of DNA

Sequencing of plasmids was done by MWG Biotech AG, Anzinger Strasse 7a, D-85560 Ebersberg. For this purpose 3-5 µl of purified and dried plasmids (2.2.10.8.) were used as a template. To prime the sequencing reactions, standard primers M13 uni (-43): 5'-AGGGTTTTCCCAGTCACGACGTT-3' or M13 rev (-49): 5'-GAGCGGATAACAA-TTTCACACAGG-3' or sequence specific primers were used. For the entire sequencing of the BtMV-G full-length clone the oligonucleotides listed in Tab. 2 were adjusted to 10 µM and used by MWG-Biotech for sequencing.

Table 2: Oligonucleotides used for sequencing of the BtMV-G full-length clone

Primer	**BtMV position**	**Sequence**
BtMV_1	265-283	5'-AATGATGGCCTCCTTCGCG-3'
BtMV_2	566-587	5'-GGTGACATCACAGCGGAATTT-3'
BtMV_3	858-879	5'-GCAGGGAACTACATCCACAACG-3'
BtMV_4	1156-1175	5'-GCTTGTTCCAGCTCAGAAAG-3'
BtMV_5	1466-1485	5'-TTGACACAAGGGCACAAGGC-3'
BtMV_6	1774-1792	5'-CGATGAAGTTAACTCCTGGG-3'
BtMV_7	2053-2074	5'-TGGAACAAGTGGAGATCCCAAG-3'
BtMV_8	2369-2390	5'-GGCTCATTGACAACTGGGTATC-3'
BtMV_9	2681-2699	5'-AGTGCAATCTTCGCGATGC-3'
BtMV_10	2962-2983	5'-GGAATGGCGAGGTTTAAGCTTG-3'

Table 2: Oligonucleotides used for sequencing of the BtMV-G full-length clone (continued)

Primer	BtMV position	Sequence
BtMV_11	3283-3301	5'-CATGATTGCAGCCAGTATAT-3'
BtMV_12	3552-3571	5'-CAGTCGCTCTTATGGCTTTG-3'
BtMV_13	3857-3875	5'-ACGAGGCACACAGCAGCTA-3'
BtMV_14	4157-4178	5'-CCACAGAGGCTTGAAGAGTTTG-3'
BtMV_15	4479-4499	5'-TGGAGAAGGCTACCTGGTCA-3'
BtMV_16	4769-4789	5'-CGCATTGGCCATACTGAGACA-3'
BtMV_17	5078-5099	5'-CGATGGAAGACAGTCGGAGAAT-3'
BtMV_18	5388-5406	5'-CAAGCTTTTCGCTAGCAGG-3'
BtMV_19	5675-5695	5'-GGTTGCTGGATGCTATGGAG-3'
BtMV_20	5982-6003	5'-CAAAGGATGAGAGCGTGCAAAC-3'
BtMV_21	6286-6307	5'-AGAGGTTGTCAGGGAAGAGGGA-3'
BtMV_22	6587-6608	5'-CCGATGAAAGCGGAAGTTTAGAG-3'
BtMV_23	6894-6912	5'-CGCTTGAATGGGTGAAACA-3'
BtMV_24	7171-7191	5'-TGCAACACATGAGGAAGCAGC-3'
BtMV_25	7479-7500	5'-CTGCAGAGCGCGAGAATTTACT-3'
BtMV_26	7770-7791	5'-GGATCTATTGTGATGCCGATGG-3'
BtMV_27	8081-8102	5'-CGGTTCTTCGCAAATGGTGATG-3'
BtMV_28	8366-8387	5'-GCATGGGGATACCCAGAACTTT-3'
BtMV_29	8692-8713	5'-GCAGAGTGTCATCAAGCAGAT-3'
BtMV_30	8996-9017	5'-CCAAATCTAGTGGAGACTGGG-3'
BtMV_31	9294-9312	5'-TGGCAAGCGTGTCCAACAA-3'

2.2.11. Construction of a BtMV-G full-length clone

2.2.11.1. Amplification of the genomic RNA of BtMV-G

A full-length cDNA clone of the German isolate of BtMV was constructed from four overlapping cDNA clones in the plasmid pe35StupAII harboring an enhanced *Cauliflower mosaic virus* (CaMV) 35S promoter and a CaMV terminator. The 35S promoter provides high efficient gene expression and can function in a wide range of organisms, e.g. also in tobacco plants (Mushegian and Shepherd, 1995). Oligonucleotides used for the construction of the BtMV-G full-length clone are shown in Tab. 3.

Table 3: Oligonucleotides used for construction of the BtMV-G full-length clone. Primers were designed based on the BtMV complete sequence. Restriction sites used for cloning are shown in bold letters. s = sense primers; as = antisense primers; C1-C4 = antisense primers

Primer	**Sequence**
B1s-*Dra* I	5'-AA**TTTAAA**ATTAAAACATCTCAATACAACACATGC-3'
B1as-*Xba* I	5'-AAGGGGTTCCGTCA**TCTAGA**GTGA-3'
B2s-*Xba* I	5'-AATGTCAC**TCTAGA**TGACGGAACCC-3'
B2as-*Mlu* I	5'-AATG**ACGCGT**CCTACTCTACCAAG-3'
B3s-*Mlu* I	5'-AATTGGTAGAGTAGGA**CGCGT**CAA-3'
B3as-*Nco* I	5'-AAACAGA**CCATGG**ACCCTTGAGGT-3'
B4s-*Nco* I	5'-AAAAACCTCAAGGGT**CCATGG**TCT-3'
B4as-*Bsp*E I	5'-AAA**TCCGGA**TTTTTTTTTTTTTTTTTTTTTTGTAAACC-3'
C1	5'-CGTGATATCCAATGTTTTGCTGCTA-3'
C2	5'-CAGCCATTAAAATGAACTAG-3'
C3	5'-TAACAATTCGTTCCACCCTC-3'
C4-dt20	5'-TTTTTTTTTTTTTTTTTTTT-3'

Four specific antisense primers were designed based on the BtMV-G complete sequence for cDNA synthesis (Tab. 3): C1 (2643-2667), C2 (4961-4980), C3 (7731-7550) and C4, which is a oligo-dT(20) complementary to the poly-A tail at the 3'end of the virus. These four primers were used separately to synthesis the cDNA starting from the purified RNA as follows:

1 µl pure viral RNA and 1.5 µl (10 µM) gene-specific antisense-primer were mixed with 8.9 µl of double distilled water and incubated at 94^{o}C for 5 min, rapidly cooled on ice. 4 µl 5x reverse transcription buffer [(250 mM Tris-HCl, 200 mM KCl, 25 mM $MgCl_2$, 2.5% Tween 20% (v/v), pH 8.3 (25 °C)] 2 µl DTT (100 mM), 2 µl dNTPs (10 mM), 0.5 µl RNase Inhibitor (400 u/µl) and 0.1 µl Expand Reverse Transcriptase (50 u/µl) were added and incubated at 43^{o}C for 1 h. Thereafter, the cDNA was used for amplification of the four fragments.

Eight primers were designed based on the BtMV-G complete sequence including single restriction sites to amplify four fragments of BtMV-G genomic cDNA; every primer contains the suitable restriction site as shown in Tab. 3.

The PCR reaction mixture (50 µl) was prepared containing 5 µl 10x PCR buffer (2.2.4.2), 5 µl $MgCl_2$ (25 mM), 1 µl (10 µM) of gene specific antisense-primer and 1 µl (10 µM) gene specific sense primer, 0.5 µl dNTPs (25 mM), 0.1 µl mixture of (10 µl platinum Taq (10 u)

and 1u *Pfu*-polymerase), 35.4 µl double distilled water and 2 µl cDNA. Special PCR conditions were programmed to amplify the four fragments (B1, B2, B3, B4):

Number of step	Program Step	Temperature	Time
1	Denaturation	94°C	2 min
2	Denaturation	94°C	15 sec
3	Annealing	50°C	30 sec
4	Extention*	68°C	1 min
5	Denaturation	94°C	15 sec
6	Annealing	50°C	30 sec
7	Extention	68°C	3 min + 5 sec for each cycle
8	Final extension	72°C	7 min

* 10 cycles programmed from step 4 back to step 2.

The four overlapping fragments were extracted from the LM gel purified (2.2.9), ligated into the pGEM-T vector (2.2.10.1) and transformed in *E. coli* NM 522 (2.2.10.3). Afterwards, the plasmids were extracted (2.2.10.4) and a control digestion was done using *Eco*R I (2.2.10.5) in order to confirm the DNA inserts.

2.2.11.2. Construction of the required Plasmids:

The sequences of the original plasmids were modified due to the requirements of the cloning steps by insertion of required restriction site sequences either from other plasmids or by using oligonucleotides containing the suitable restriction sites.

2.2.11.2.1. Construction of the plasmid pe35StupAII

The plasmid V81-pNEB193 was modified by inserting four nucleotides (TAGC) inside the restriction site sequence of *Hin*d III (5'-AAGCTT-3') using a Klenow fill-in reaction to obtain a *Nhe* I site (5'-CTAGCT-3') creating a new plasmid named V206-pNEB193NheI. This plasmid was digested with *Sal* I/*Pst* I to remove 19 unwanted nucleotides (5-ACGACTGTTTAAACCTGCA-3'). The digested plasmid was ligated with a fragment of 31 nucleotides (5'-TCGACCCGGGAATTC-CGGACCGGTACCTGCA-3') containing multiple restriction sites of plasmid V100-pSPORT1 giving plasmid V208-pNEB193EH. The plasmid p442-pe35StupA, which contains the 35S promoter was digested with *Bam*H I / *Xba* I to remove 6 nucleotides (5'-GATCCT-3') and replace them by 62 nucleotides (5'-

GATCCTTAATTAAGTCTAGAGTCGACCCGGGAATTCCGGTACCTGCAGGCATGC AAG-3') derived from digestion of the constructed plasmid V208 pNEB193EH with *Bam*H I / *Nhe* I. This ligation created the plasmid V209 pe35StupAII containing an enhanced 35S promoter (e35S), a 35S termination sequence (35S Term), a ß-lactamase gene (ApR) and suitable restriction sites for cloning of the BtMV-G fragments. The plasmid is shown in Fig.3.

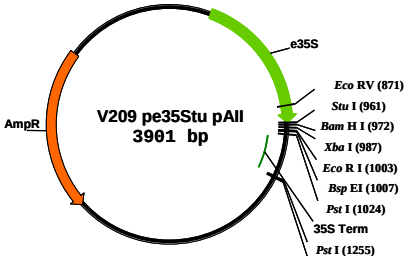

Figure 3: Plasmid pe35StupAII with selected restriction endonucleases sites. e35S = enhanced 35S CaMV promoter; 35S Term = CaMV Termination signal; AmpR = ß-lactamase gene.

2.2.11.2.2. Construction of the plasmid V217-pBlueKSP

The plasmid pBlueScriptII(-)KS was modified by digesting with *Eco*R V and annealing to the oligonucleotides V1(5'-AATCTAGAAAAAACGCGTAAAACCATGGAAAAT-CCGGAAA-3') and V2 (5'-TTTCCGGATTTTCCATGGTTTTACGCGTTTTTT-CTAGATT-3') to create plasmid V217-pBlueKSP containing *Nco* I, *Xba* I and *Bsp*E I sites (Fig. 4.).

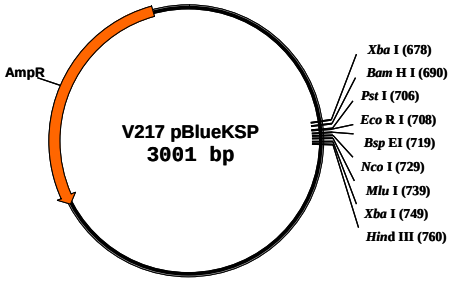

Figure 4: Plasmid V217-pBlueKSP with selected restriction endonuclease sites.
AmpR = ß-lactamase gene.

2.2.11.2.3. Construction of the plasmid V223-pDRIVE2

The plasmid V220-pDRIVE was digested with *Bsp*120 I / *Pst* I to remove 91 nucleotides and ligated to a short fragment of 84 nucleotides 5'-GGAATTCGATTTTCCGGATTTTCCATGGTTTTACGCGTTTTTTCTAGATTATCAA GCTTATCGATACCGTCGACCTCGAGGGGG-3', which derived from digestion of V217pBlueKSP with *Bsp*120 I and *Pst* I. The ligation resulted in plasmid V223-pDRIVE2 containing *Nco* I and *Mlu* I restriction sites and a gene conferring resistance to kanamycine (Fig. 5.).

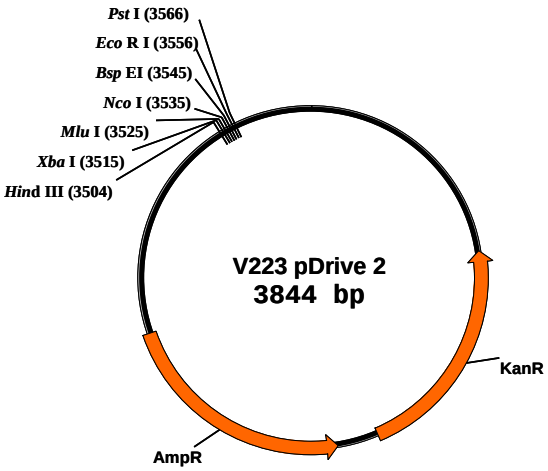

Figure 5: Plasmid V223-pDRIVE2 with selected restriction endonuclease sites.
AmpR = ß-lactamase gene; KanR = neomycin phosphotransferase

2.3. Sequence analysis

The retrieved sequences of BtMV-G fragments were assembled by DNASTAR in order to establish the complete nucleotide sequence. Vector-NTI (Demo version) was used to determine the open reading frames (ORF) and to translate the BtMV sequence into proteins. In addition Vector-NTI was used to design oligonucleotides for RT-PCR. Alignments of deduced amino acids sequences with other potyvirus amino acid sequences were done by CLUSTALX (Vers. 1.83; Thompson et al., 1997). Phylogenetic trees were generated by CLUSTALX and visualized with TreeView using standard parameters.

3. Results

3.1. Symptoms of BtMV-G

3.1.1. Symptoms of BtMV-G on *N. benthamiana, B. vulgaris* and *Chenopodium quinoa* plants after mechanical inoculation

Mechanical inoculation (2.2.3.1.) of BtMV-G was successfully applied to infect *N. benthamiana* and *B. vulgaris.* One to two weeks after inoculation typical mosaic symptoms appeared on the leaves. Mosaic symptoms on *N. benthamiana* leaves were very prominent and the virus infected plants showed a tubular shape, curling and crinkling of leaves and leaf distortions (Fig. 6-A1 and Fig. 6-A2).

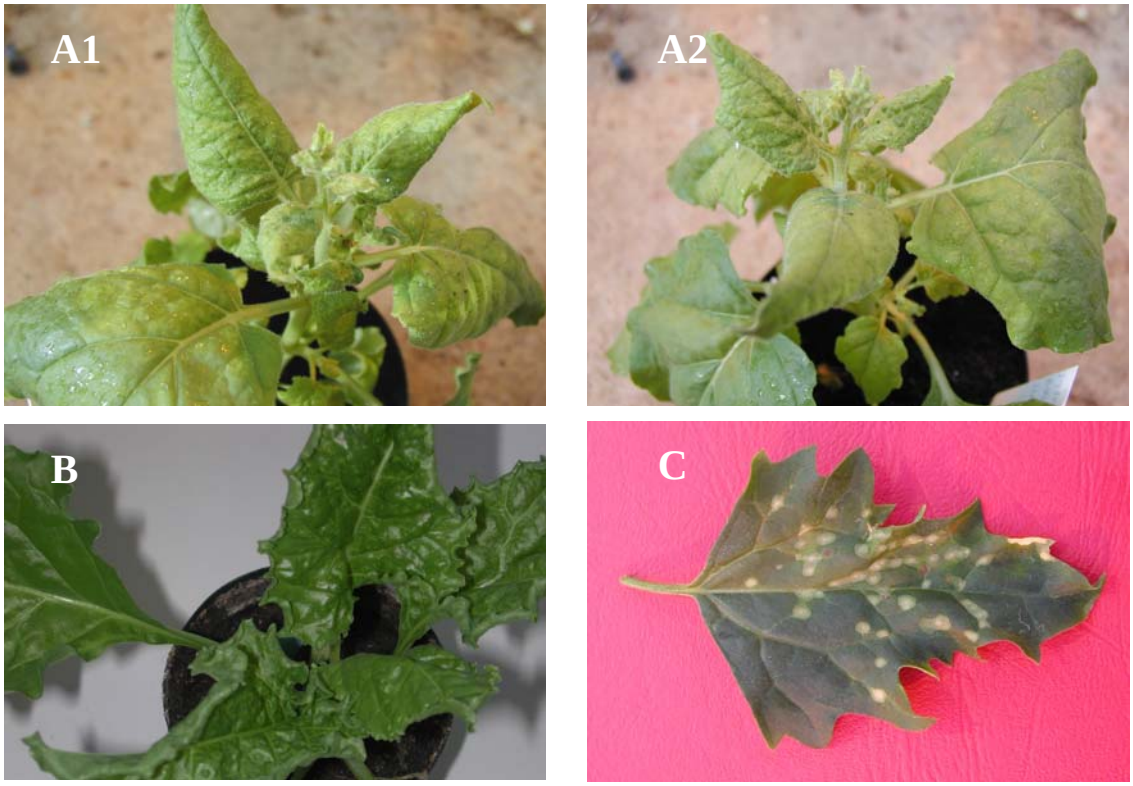

Figure 6: BtMV-G infected plants. (A1 to A2) Systemically infected leaves of *N. benthamiana*. (B) Systemically infected leaves of *B. vulgaris*. (C) Local lesions on *Chenopodium quinoa*. (Pictures were taken 14 - 21 dpi).

The infected *B. vulgaris* plants showed a light mosaic, curling and rolling of leaf margins, spots, blotches and mottle on young leaves (Fig. 6-B). *Chenopodium quinoa* plants inoculated with BtMV-G developed chlorotic local lesions (Fig. 6-C).

The presence of BtMV-G in *N. benthamiana* plants was confirmed by plate trapped ELISA (2.2.4.1.), RT-PCR (2.2.4.2.) and Immunocapture-RT-PCR (2.2.4.3.). ELISA readings showed clear differences between the infected samples (n = 4; O.D. = 2.172 ± 0.28) and samples from healthy plants (n = 4; O.D. = 0.0845 ± 0.0044). RT-PCR and Immunocapture RT-PCR with oligonucleotides BtMVCPs and BtMVCPas (2.2.4.2.) confirmed the presence of BtMV-G in the infected leaf material by amplification of a DNA band of 635 bp representing a part of the coat protein gene as shown in Fig. 7.

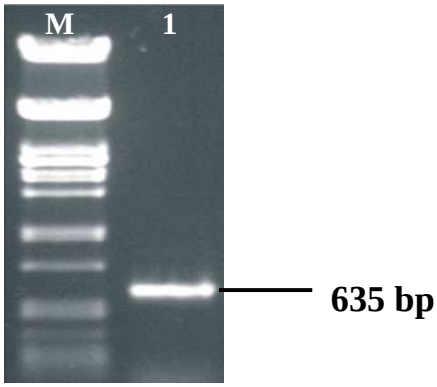

Figure 7: BtMV - CP DNA fragment amplified by Immunocapture RT-PCR
Lane M. DNA size standard: Lambda-DNA digested with *Pst* I. Lane 1. DNA fragment of 635 bp representing a part of the BtMV-G coat protein gene.

3.1.2. Symptoms of BtMV-G after aphid transmission on *N. benthamiana*

Myzus persicae was used for aphid transmission (2.2.3.2.) of BtMV-G on *N. benthamiana*. The experiment was repeated three times. Symptoms were observed about 10 days after aphid inoculation, whereas the symptoms developed on *N. benthamiana* were indistinguishable from symptoms developed after mechanical transmission of BtMV-G. In each of the three experiments the aphid transmission efficiency was 30 %.

3.2. RNA extraction from BtMV-G infected *N. benthamiana*

Three different methods were used to extract RNA from infected *N. benthamiana* leaves and from a partially purified BtMV-G preparation. The yield and quality of the extracted RNA generated with each method were compared. In addition the suitability of the different RNA preparations for RT-PCR assays was checked. RNA was extracted by using a hot phenol extraction method (2.2.5.1.), by a dsRNA extraction procedure (2.2.5.2.) and after partial purification of BtMV-G (2.2.5.3.; 2.2.5.4.).

The hot phenol extraction method revealed high RNA yields. Nucleic acids appeared as several bands and also as smear on the gel (Fig. 8-A, lane 1). The distinct bands show the presence of rRNA, however a band representing the viral RNA was not visible. The smear probably represents degradation products of RNAs. RT-PCR (2.2.4.2.) with an aliquot of the RNA preparation (1 µl) and oligonucleotides BtMVCPs and BtMVCPas (2.2.4.2.) reveals a DNA fragment of 635 bp, which appears as a clearly visible band in an agarose gel (Fig. 8-B, lane 2).

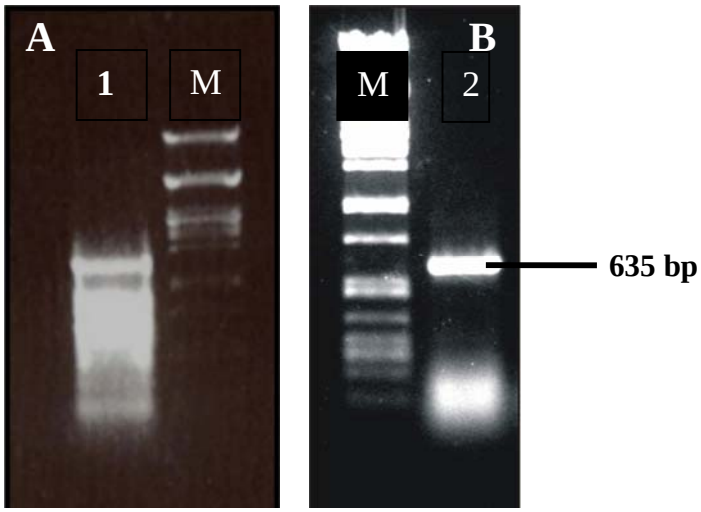

Figure 8: (A) Total RNA extract (5 µl) in lane 1 derived from the hot phenol extraction.
(B). DNA fragment of BtMV-G (10 µl) in lane 2 generated by RT-PCR
M = DNA size standard: Lambda-DNA digested with *Pst* I.

The hot phenol extraction method delivered RNA, suitable to conduct RT-PCR for generating DNA fragments from a known region of BtMV-G, indicating that this method can be also used for the generation of DNA fragments extending into the so far unknown region at the 5'-end of the BtMV-G genome.

The dsRNA extraction method delivered only sparse amounts of dsRNA, hardly visible as a nucleic acid band on the top of an agarose gel (Fig. 9-A, lanes 1 and 2). In addition a smear over the entire lane and putative degradation products of nucleic acids at the bottom of the lanes were visible. After extraction of the dsRNA from the gel a RT-PCR (2.2.4.2.) with an aliquot of the dsRNA preparation (2 µl) and oligonucleotides BtMVcol1 and BtMVcol2 reveals a DNA fragment of 385 bp as a clearly visible band (Fig. 9-B, lanes 1 and 2.). However, despite the general usability of the dsRNA for generating DNA fragments of BtMV-G by RT-PCR, the method was found to be very laborious and time consuming and was therefore only of limited value for preparation of BtMV-G RNA.

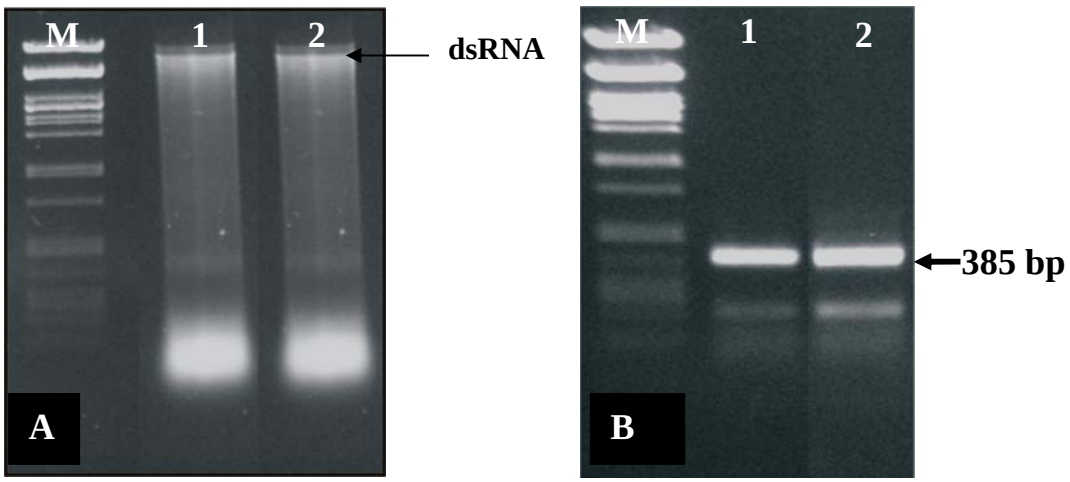

Figure 9: (A) dsRNA extracts (30 µl) in lane 1 and lane 2 derived from the dsRNA extraction method. (B) DNA fragments of BtMV (10 µl) in lane 1 and lane 2 generated by RT-PCR using oligonucleotides BtMVcol1 and BtMVcol2 and dsRNA.

3.3. Purification of BtMV-G and analysis of the BtMV-G coat protein by SDS-polyacrylamide gel electrophoresis

3.3.1. Purification of BtMV-G

BtMV-G was partially purified to extract subsequently high quantities of pure viral RNA. A purification procedure developed by Rogov (1991) was used with some modifications to purify BtMV-G in this study. The purification protocol was adopted by omitting the sucrose cushion centrifugation step to obtain higher amounts of the virus. The inclusion of DIECA during homogenization of infected plant material was essential, because in its absence virus yield decreased by at least a factor of 10. The concentration of the purified virus was determined with 5.5 mg/ml with a calculated RNA concentration of 0.28 µg/µl.

Viral RNA was extracted from the partially purified virus by viral RNA deproteinization and EtOH precipitation (2.2.5.4.). Purified RNA (5 µl) was separated on an agarose gel. Because of a too high concentration of the RNA sample a smear was visible in the agarose gel (Fig. 10-A). However, the RNA obtained with this method was found to be suitable in RT-PCR to generate a clearly visible DNA band of 635 bp by using primer BtMVCPs and BtMVCPas (Fig. 10-B).

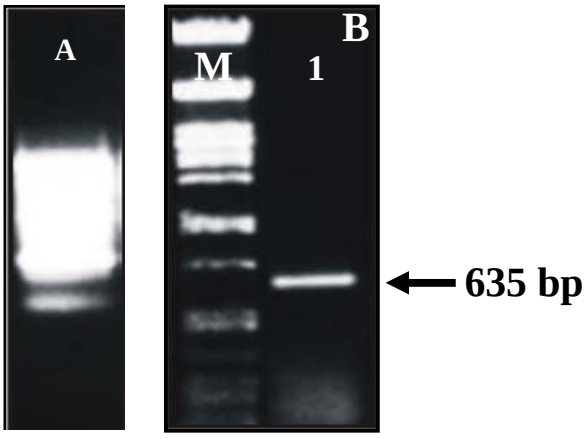

Figure 10: (A) RNA extracted from purified BtMV-G.
(B) DNA fragment of 635 bp generated from purified viral RNA by RT-PCR.
M = DNA size standard: Lambda-DNA digested with *Pst* I.

This partially purified viral RNA was subsequently used for RT-PCR to generate cDNA fragments covering the 5'-end of the BtMV-G genome.

3.3.2. SDS-polyacrylamide gel electrophoresis

An aliquot of the purified virus preparation was examined by SDS-PAGE to evaluate the purity of the preparation and to determine the molecular weight of the BtMV-G CP. SDS-PAGE revealed a single protein band representing the CP of BtMV-G. No further protein bands were observed, indicating no contaminant proteins in the virus preparation. The apparent molecular weight of the BtMV-G CP was determined with ~29 kDa (Fig. 11).

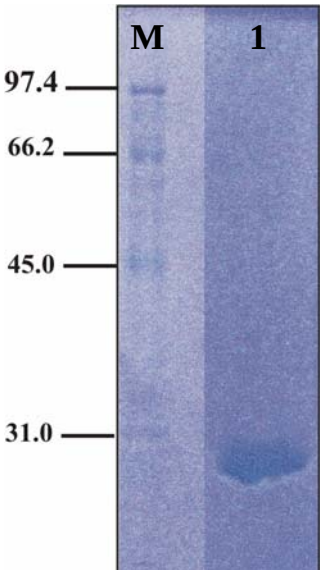

Figure 11: SDS-PAGE of partially purified BtMV-G. M = Molecular weight marker; the molecular weights of the marker proteins are indicated on the left in kDa; Lane 1 = CP of BtMV-G.

3.4. Cloning and sequencing of the BtMV-G genome

Reverse transcription and random hexamer PCR was used to generate cDNA fragments covering almost of the 5'-end of BtMV-G. In addition ligation of an adaptor oligonucleotide and poly-G tailing procedures were performed in order to amplify and determine the exact 5'-end of the BtMV-G genome.

3.4.1. Random hexamer PCR (RH-PCR)

A random hexamer PCR method developed by Froussard (1992) was used with modifications to amplify cDNA populations from the unknown sequence part of the BtMV-G genomic RNA. For this purpose, first strand cDNA synthesis was primed with an oligonucleotide designed to an already known region of the BtMV-G genome. Second strand cDNA synthesis was achieved with a universal oligonucleotide carrying a random hexamer sequence at its 3'-end and the Klenow fragment of *E. coli* DNA polymerase I. Products synthesized by the Klenow fragment were subsequently amplified using a universal oligonucleotide without the random hexamer sequence and a nested BtMV-G specific oligonuclotide.

10 µl of the final RH-PCR amplification products were analyzed on a 1% agarose gel. PCR amplification products appeared as dominant bands but also as a smear with band sizes ranging from 0.2 kb to 1.2 kb (Fig. 12-A and Fig. 12-B).

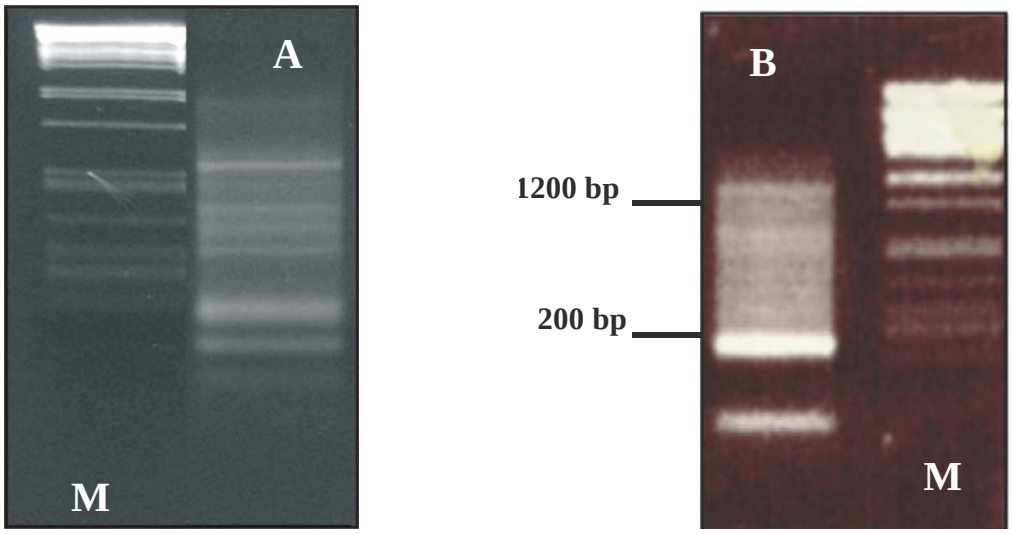

Figure 12: Agarose gel electrophoresis of RH-PCR products (A) First round amplification using primer BtMVcol2 and BtMV5. (B) Second round amplification using primer BtMVcol5 and BtMV2. M = DNA size standard: Lambda-DNA digested with *Pst* I.

The method was applied to generate in four steps DNA fragments covering parts located at the 5'-end of the BtMV-G genomic RNA. Each step includes the use of a BtMV-G specific primer for first strand cDNA synthesis, the use of the universal oligonucleotide carrying the random hexamer sequence at its 3'-end for second strand cDNA synthesis and a RH-PCR step with the universal primer (2.2.7.1.4.) and another BtMV-G specific primer (2.2.7.1.4.) to amplify synthesized double-stranded DNAs.

3.4.2. Adaptor ligation and PCR amplification

Template RNA was used to synthesize cDNA by using Reverse Transcriptase and BtMV-G gene-specific antisense-primer. The cDNA was precipitated and dissolved in double distilled water (2.2.7.2.1.1). Dissolved cDNA was ligated to an adaptor oligonucleotide carrying a phosphate residue at its 5'-end and an amino group at its 3'-end (2.2.7.2.1.2.). The ligation mixture was amplified using a BtMV-G gene specific antisense primer and a primer complementary to the used adaptor oligonucleotide (2.2.7.2.1.3). Final PCR products were analyzed on a 1% agarose gel. PCR amplification products appeared as 3 bands (Fig. 13). The upper band of approximately 900 bp was purified and used for cloning.

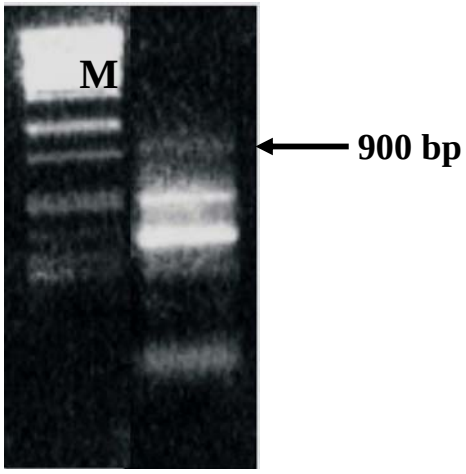

Figure 13: Agarose gel electrophoresis of PCR products after adaptor ligation using primer BtMVcol7 and a primer complementary to the used adaptor oligonucleotide.
M = DNA size standard: Lambda-DNA digested with *Pst* I.

3.4.3. Poly-G tailing and PCR amplification

Viral genomic RNA was used as a template to synthesize cDNA with Superscript II Reverse Transcriptase by using a BtMV-G specific antisense primer (2.2.7.2.2.1.). The cDNA was tailed with dGTP using Terminal Transferase followed by PCR amplification with a nested BtMV-G specific primer and an oligo-C15 primer (2.2.7.2.2.2.). This approach generated an additional DNA fragment of about 300 bp (Fig. 14) covering the extreme 5'-end of the BtMV-G genome.

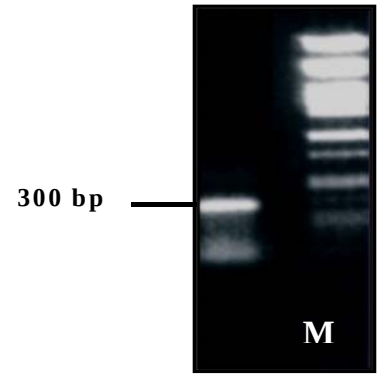

Figure 14: Agarose gel electrophoresis of PCR products after poly G tailing using primer BtMV5as and C15. M = DNA size standard: Lambda-DNA digested with *Pst* I.

3.4.4. Cloning of PCR products

In the first round of RH-PCR amplification a DNA band of approximately 900 bp was generated from purified BtMV-G RNA. The DNA fragment was extracted from a low melting agarose gel, purified (2.2.9) and ligated into the pGEM-T Easy vector (2.2.10.1). After transformation (2.2.10.3) 24 colonies were screened by colony PCR (2.2.10.6) and plasmid preparations (2.2.10.4) were used for a restriction endonuclease digest with *Eco*RI (2.2.10.5). As indicated in Fig. 15 for one plasmid DNA inserts of approx. 900 bp were found.

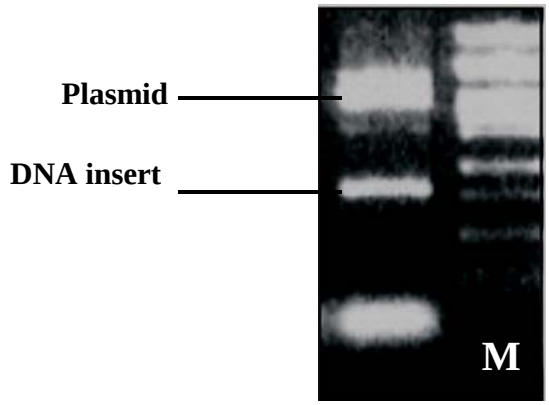

Figure 15: Agarose gel electrophoresis of pGEM-T Easy digested with *Eco*RI. A DNA fragment of approx. 900 bp was released from the vector. M = DNA size standard: Lambda-DNA digested with *Pst* I.

After further purification of one clone (2.2.10.8) the DNA was sequenced (2.2.10.9). Sequence analysis confirmed the presence of a BtMV-G specific insert. Additional primers designed 50-100 bases from the end of this insert were used to conduct the next cloning step. Another three overlapping clones were generated with the same approach, the lengths of four overlapping clones were, 916, 765, 544, 779 bp respectively. Fig. 16 shows the localization and length of the cDNA clones generated with the RH-PCR method and the poly-G tailing approach.

The adaptor ligation approach generated an additional DNA fragment of about 900 bp, which was transformed into *E .coli*. 10 colonies were screened by plasmid PCR and DNA inserts were found in three plasmids. The result was confirmed by sequencing and one of the plasmids revealed a DNA insert of 585 bp covering the extreme 5'-end except one nucleotide (adenosine residue; Fig. 17-A). This clone is the only overlapping sequence between the fourth clone produced by RH-PCR and another clone produced by using the poly-G tailing approach. Poly-G tailing was performed to determine the extreme 5'-end of BtMV-G. Four plasmids were selected after screening of 24 colonies and analyzed with plasmid-PCR. The DNA inserts of these four plasmids were sequenced to confirm the 5'-end sequence. The DNA inserts of these two independent clones (clone 1 and 3) showed that the BtMV-G sequence starts with four adenosine residues at the 5'– end as shown in Fig. 17-B.

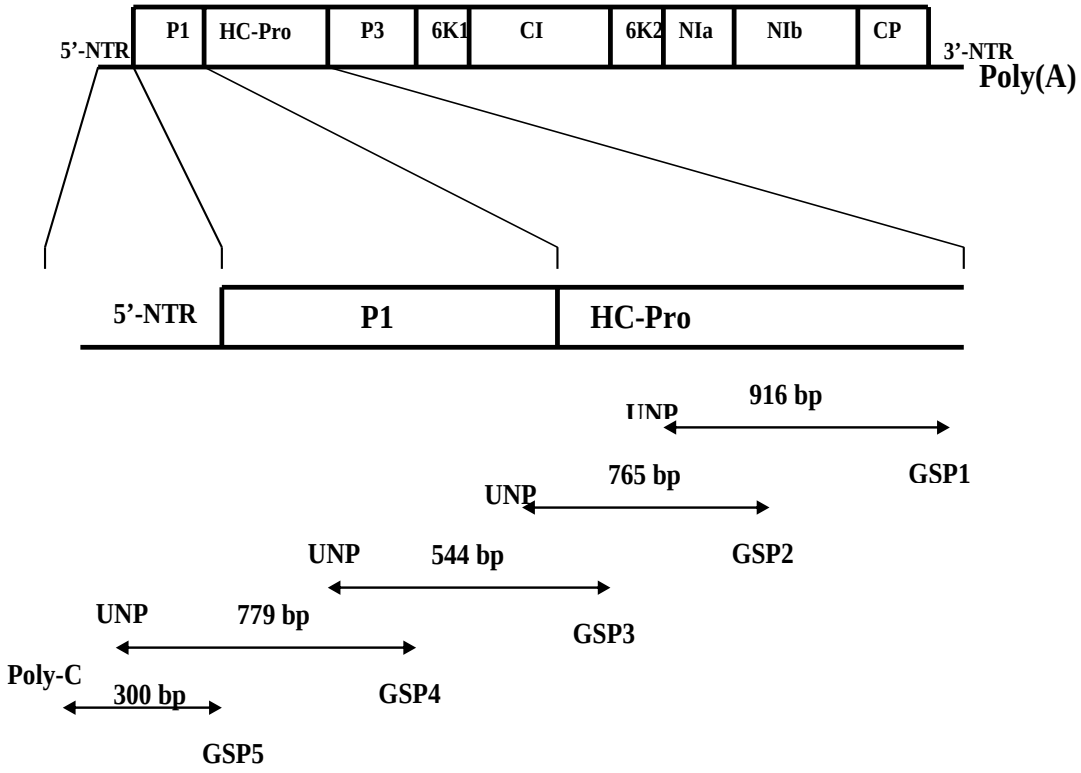

Figure 16: Localisation of cDNA clones covering the 5'-NTR, P1 and HC-Pro of BtMV-G.; GSP = Gene specific primer used for PCR amplification; UNP = Universal primer (see 2.2.7.1.4.)
Poly-C = Poly-C primer (see 2.2.7.2.2.2.)
GSP1: BtMV5 (Tab. 1).
GSP2: BtMV2 (Tab. 1).
GSP3: BtMV3 (Tab. 1).
GSP4: BtMV4 (Tab. 1).

```
A. Adaptor ligation:
Clone 1: 5'-.AAATTAAAACAT----3'

B. Poly G tailing:
Clone 1: 5'-AAAATTAAAACAT----3'
Clone 2: 5'-AAAACAAGCCAAC----3'
Clone 3: 5'-AAAATTAAAACAT----3'
Clone 4: 5'-TCCTTAAATTAAA----3'
```

Figure 17: Sequences of the ends of five clones produced by the adaptor ligation approach (A) and poly-G tailing (B) to elucidate the sequence of the 5'-end of the BtMV-G genome.

3.5. BtMV-G sequence analysis and comparison with other *Potyvirus* genomes

3.5.1. Primary structure of the BtMV-G genomic RNA

More than 70% of the BtMV-G genome has been cloned (Neeman, 1999; Berger, 2000) but BtMV-G has not been completely characterized at the molecular level and a complete genomic RNA sequence was not previously available. In this study the complete nucleotide sequence of BtMV-G was determined. For this purpose it was necessary to develop and apply a random hexamer PCR approach together with a poly-G tailing method to amplify and to clone the missing part of the BtMV-G genome. With this methods clones covering the 5'-non-translated region, P1 and HC-pro were established. Five overlapping clones were generated during this study to complete the BtMV-G sequence. A sequencing strategy view for determination of the complete BtMV-G sequence is presented in Fig. 18. In addition the sequence was confirmed by sequencing of the entire BtMV-G insert of a full-length clone (see 3.6.4).

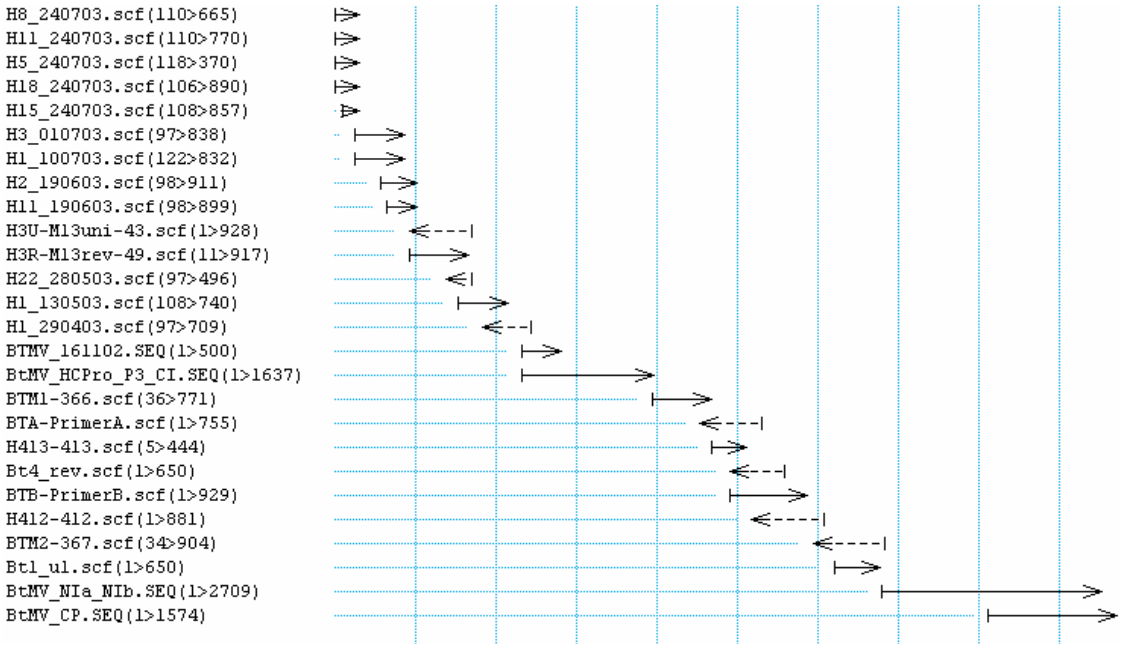

Figure 18: Sequencing strategy for the determination of the complete BtMV-G sequence. Arrows indicate the length and direction of compiled sequences. Sequences generated during this study are labelled with "H".

The BtMV-G genome comprises of 9592 nucleotides (nts) and contains one large open reading frame (Appendix, Fig. A1). The 5'- non-translated region (NTR) was determined with 166 nts starting with four adenosine residues. It contains 15 CAA repeats, which are also described for other potyviruses. In addition the so called potybox a' (ACAACAC) at nucleotide position 20-26 and potybox b' (TCAAATT) at nucleotide position 59-65 were identified. The single large open reading frame comprises of 9255 nucleotides, starts with an AUG codon at positions 167-169 and ends at nts 9422-9424 with a TAA termination codon, which is followed by a 3'- NTR of 171 nts and a terminal poly-A sequence.

3.5.2. Putative proteases cleavage sites identified in the BtMV-G polyprotein

The nucleotide sequence of the ORF is translated into a polyprotein of 3085 amino acid residues. Putative proteolytic cleavage sites can be identified in the polyprotein, typical for a member of the family *Potyviridae*. Based on the likely location of cleavage sites and based on alignments of amino acid sequences for each functional protein the polyprotein of BtMV-G is probably processed into ten mature proteins, which are typical for all members of the genus *Potyvirus*. The cistrons encoding the proteins are conserved in the same order in all studied potyviral genomes and include the first protein (P1) at the N-terminus of the polyprotein, followed by the helper component-proteinase (HC-Pro), the third protein (P3), the first 6 kDa protein (6K1), the cylindrical inclusion protein (CI), the second 6 kDa protein (6K2), the VPg domain of the nuclear inclusion protein a (NIa-VPg), the proteinase domain of the nuclear inclusion protein a (NIa-Pro), the nuclear inclusion protein b (NIb) and the coat protein (CP).

Three proteinases are involved in the complete processing of the potyviral polyprotein: P1 proteinase (Verchot *et al.*, 1991), HC-Pro proteinase (Carrington *et al.*, 1989) and NIa-Pro proteinase (Dougherty and Parks, 1991). Nine putative protease cleavage sites were predicted for BtMV-G by analogy with genome arrangements of other potyviruses (Tab. 4).

Table 4: Putative cleavage sites of the viral proteases in the polyprotein of BtMV-G and amino acid sequences adjacent to the cleavage sites.

Cleavage site	Amino acid position									Potyviral consensus
	P_6	P_5	P_4	P_3	P_2	P_1	$P_{1'}$	$P_{2'}$	$P_{3'}$	
P1/HC-Pro	G	R	T	V	H	**Y**	**S**	S	G	Y/S
HC-Pro/P3	K	W	Y	R	V	**G**	**G**	H	G	G/G
P3/6K1	A	E	V	E	Y	**Q**	**A**	N	K	Q/A
6K1/CI	E	E	V	K	Y	**Q**	**S**	L	D	Q/S
CI/6K2	D	T	V	Q	Y	**Q**	**S**	A	T	Q/S
6K2/VPg	E	I	V	Q	Y	**Q**	**G**	K	R	Q/G
VPg/NIa	E	V	V	R	E	**E**	**G**	K	S	E/S
NIa/NIb	D	E	V	V	E	**Q**	**G**	Y	S	Q/S
NIb/CP	D	I	V	T	Y	**Q**	**G**	D	E	Q/S

The cleavage site at the C-terminus of P1 occurs probably at the dipeptide Y/S (313-314 aa). The HC-Pro cleavage site at the C-terminal G/G dipeptide is part of the conserved motif of the potyvirus amino acid sequence YXVG/G (767-771). The remaining seven protease recognition sites are putatively cleaved by the NIa-Pro at dipeptides Q/A, Q/S, Q/G and E/G, which are also found in other potyvirus sequences. Most of these cleavage sites are very similar and quite common among some potyviruses, and some of them are different from the typical potyvirus cleavage sites, like the VPg/NIa and NIa/NIb junctions with glutamic acid/glycine (E/G) and glutamine/glycine (Q/G). However, glutamine/glycine residues at the NIb/CP junction, as found for BtMV-G, occur with a lower frequency. Nevertheless, an unusual cleavage site was not found. The putative proteolytic cleavage sites of BtMV-G are mostly related to those of TEV, PeMoV and SMV.

3.5.3. Functional conserved amino acid sequence motifs of BtMV-G shared with other potyviruses

Several conserved amino acid residues organized in functional motifs of potyviruses were detected in the BtMV-G polyprotein (Tab. 5), however also some differences were recognized. For example, amino acids probably involved in the viral long distance movement motif (SCCVT) and the cell-to-cell movement motif (LAIGN) were present in the HC-Pro of BtMV-G. The SCCVT motif contained a substitution, whereas cysteine was

replaced by serine. The conserved nucleotide-binding motif VGSGKST and the RNA helicase motif DESH were found in the CI. The RNA helicase motif of BtMV-G is slightly different from that of other potyviruses. The serine replaces a cysteine or phenylalanine, which were often found in other potyviruses. The conserved motifs of the RNA dependent RNA polymerases of positive-stranded RNA viruses were found in the NIb of BtMV-G.
Motifs known to be involved in the aphid transmissibility KITC, PTK, FRNK and DAG are found in the HC-Pro and the CP of potyviruses. However, instead of the KITC motif BtMV-G contained two amino acid substitutions and the motif KMAC was found. Despite of this mutations BtMV-G was transmissible with aphids to a frequency of 30% (see section 3.1.2.), indicating that the KMAC motif is also functional in aphid transmission.

Table 5: Conserved amino acid motifs of BtMV-G

Functional proteins	**Conserved motif**	**Suggested function**
P1	VIVRGR	Proteolytic domain
HC-Pro	SCCVT	Viral long distance movement
	KMAC	Aphid transmission
	LAIGN	Cell-to-cell movement
	PTK	Aphid transmission
	FRNK	Aphid transmission
CI	DESH	RNA helicase
	VGSGKST	Nucleotide binding site
NIb	CDADGS	RNA dependent RNA polymerase
	$SGX_3SX_3NTX_{30}GDD$	RNA dependent RNA polymerase
CP	DAG	Aphid transmission
	$RX_{43}D$	Viral long distance movement

3.5.4. Pairwise comparison, phylogenetic analyses and relationship of BtMV-G with other potyviruses

In order to establish evolutionary trees and a genetic relationship of BtMV-G with other potyviruses infecting other plants, phylogenetic analyses and amino acid sequence comparisons were carried out. In addition a multiple sequence alignment of the potyviruses complete protein sequences was done using ClustalX with default parameters. A phylogenetic tree was generated using the N-J tree option of ClustalX and visualisation of the multiple alignment of the complete amino acid sequences as a tree was done with the program Tree View.

The complete nucleotide sequence of BtMV-G showed high sequence identity with the BtMV-Wa (89%) at the nucleotide level followed with PeMoV (41%) and SMV (39%).

The 5'-NTR of BtMV-G shared also high nucleotide sequence identity with 5'-NTR of the same viruses BtMV-Wa (81%), PeMoV(42%) and SMV (44%). While the 3'-NTR sequence showed high sequence identity only with that of BtMV-Wa (99%) and revealed high level of discrepancies with 3'-NTRs of the other potyviruses.

BtMV-G shared the highest polyprotein amino acid sequence identity with the BtMV-Wa isolate (95%) followed by PeMoV (55%), SMV, BCMNV, ZYMV (50%) and CABMV (49%). The high percentage of identity among the aligned polyproteins of BtMV-G, BtMV-Wa and PeMoV places these viruses in one cluster within the branching order of the potyviral phylogenetic tree (Fig. 19). The next closest branch contains four potyviruses SMV, BCMNV, BCMV and CABMV, which naturally all infect hosts of the family *Leguminosae*.

The amino acid sequences of the ten mature functional proteins of the BtMV-G isolate were compared with the corresponding proteins of several viruses of the *Potyviridae*. Pairwise sequence comparison was made using the global alignment parameters (GAP: http://genome.cs.mtu.edu/align/align.html; AA-Matrix Blosum62; gap open penalty 2, gap extension penalty 2; gap size for const. penalty 2). The amino acid sequence identities for the different genomic regions between BtMV-G and other potyviruses are shown in Tab. 6.

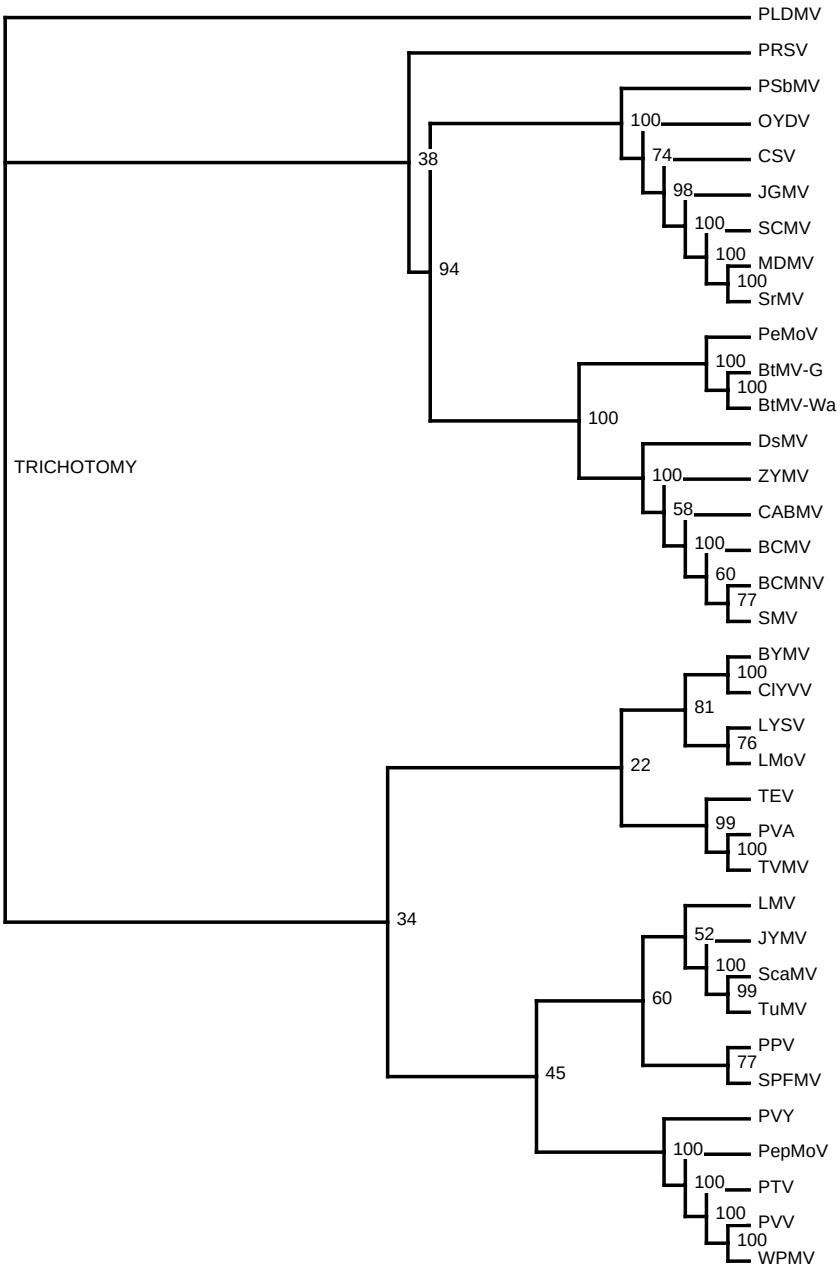

Figure 19: Phylogenetic relationship between BtMV-G and other viruses of the genus *Potyvirus*. The tree was generated using the N-J tree option of ClustalX and visualised with the program Tree View based on the multiple alignment of the complete amino acid sequences (Acronyms: see 2.1.11.) Bootstrap values indicate the number of times each branching was found in 100 bootstrap analyses with a random number generator seed of 111.

Table 6: Percentage (%) amino acid sequence identity of BtMV-G functional proteins compared to the corresponding proteins of other potyviruses.

Acronym	P1	HC-Pro	P3	6K1	CI	6K2	VPg	NIa	NIb	CP
BtMV-Wa	92	94	88	100	98	92	97	95	97	98
BCMV	19	57	25	50	60	37	56	52	58	60
BCMNV	24	55	26	53	59	33	53	51	60	59
BYMV	0	41	25	37	53	43	50	39	60	54
ClYVV	20	42	23	40	54	48	45	41	58	54
CSV	23	38	19	44	52	39	45	39	57	46
CABMV	20	55	26	51	59	39	58	50	57	57
DsMV	17	53	22	42	60	30	53	53	61	50
JYMV	17	45	24	38	53	35	53	43	61	57
JGMV	0	42	23	39	50	33	54	39	57	50
LYSV	6	42	22	37	54	37	44	42	58	55
LMV	13	46	19	36	55	43	51	46	61	55
LMoV	20	44	27	35	54	36	50	45	60	52
MDMV	16	42	25	26	55	37	42	40	57	53
OYDV	7	40	13	53	52	32	44	44	56	55
PLDMV	3	46	25	40	55	42	52	45	58	53
PRSV	6	45	24	36	56	33	50	39	56	52
PSbMV	18	43	20	44	59	30	43	45	59	53
PeMoV	30	53	35	55	67	58	61	55	67	56
PepMoV	20	42	25	36	54	36	45	44	57	59
PTV	19	46	22	38	53	40	41	44	58	57
PPV	19	46	27	32	58	35	56	46	62	49
PVA	20	51	24	32	55	39	48	44	56	56
PVV	20	40	21	40	53	38	45	46	58	53
PVY	17	43	22	44	53	44	48	44	58	58
ScaMV	13	45	19	30	53	39	54	47	59	55
SrMV	0	40	23	28	55	38	44	41	58	52
SMV	23	55	24	51	60	37	53	51	60	60
SCMV	14	41	26	26	55	36	46	39	59	51
SPFMV	10	50	23	32	55	41	57	46	61	46
TEV	20	46	24	33	53	45	45	46	61	56
TVMV	21	44	26	30	54	37	48	45	58	50
TuMV	20	45	21	36	53	37	55	49	60	55
WPMV	18	43	22	36	53	42	44	45	57	55
ZYMV	23	53	23	46	58	35	61	56	62	56

The degree of identity to the proteins of BtMV-G varied depending on the virus species and the functional protein. The highest level of identity was observed in comparison of BtMV-G to another BtMV isolate, BtMV-Wa, followed by PeMoV. This appeared for the highly

conserved functional proteins NIa, NIb and CI as well as for the less conserved HC-Pro, P3 and P1.

3.5.4.1. P1 sequence analysis

The BtMV-G P1 consists of 313 amino acids with a molecular weight of 35.3 kDa and is the most variable protein of the potyviruses. It is processed by itself as a trypsin-like serine proteinase from the HC-Pro at a Y/S cleavage site. Only few conserved amino acid residues are found in the C-terminal part, which represents the protease activity. The proteolytic domain (IVRGR) is located at amino acid position 285-290 in the BtMV-G isolate.
The multiple alignment of the deduced amino acid sequence showed that P1 is highly variable on the amino acid level. The highest identity was found with an isolate of BtMV, BtMV-Wa, with 92%, while the percentage identity with other potyviruses ranged from 30% with PeMoV to 13% with LMV and ScaMV. However, it has no significant identity with the P1 proteins of other potyviruses. With 3 potyviruses no identities were found, probably because of the selected program parameters. Only very low identities were recognized with PLDMV (3%), LYSV (6%), PRSV (6%) and OYDV (7%).
In the phylogenetic tree BtMV-G, BtMV-Wa and PeMoV are grouped in one cluster in the close neighbourhood to a cluster consisting of BCMV, BCMNV, SMV, CABMV and ZYMV with DsMV (Fig. 20). This cluster will be called the BCMV cluster later on.

3.5.4.2. HC-Pro sequence analysis

The BtMV-G HC-Pro belongs with 456 aa and a molecular weight of 51.4 kDa to a highly conserved region of the potyvirus genome. It is probably released by digestion from P1 at a Y/S junction (313-314 aa) and by self proteolysis due to a papain-like endopeptidase activity in its C-terminus at a G/G dipeptide. The alignment of BtMV-G HC-Pro with the HC-Pro of other potyviruses illustrated 6 conserved motifs. The HC-Pro includes three motifs involved in aphid transmission: KMAC, PTK and FRNK at amino acid position 365-368, amino acid position 622-624 and amino acid position 493-496, respectively. The cell-to-cell movement motif LAIGN at amino acid position 560-564, a viral long distance movement motif SCCVT at amino acid position 604-608 and another motif CDNQRD with

a still unknown function at amino acid position 509-514. This motif contains arginine in the BtMV-G HC-Pro instead of leucine in the HC-Pro of other potyviruses.

In comparison to other potyvirus HC-Pro amino acid sequences BtMV-G was found to have the highest percentage identity with BtMV-Wa (94%) followed by BCMV (57%) and BCMNV with 55%. The lowest identity of 38% was found with CSV.

The BtMV-G and BtMV-Wa were grouped in the same cluster which is located in the neighbourhood of the BCMV subgroup as shown in the phylogenetic tree (Fig. 21).

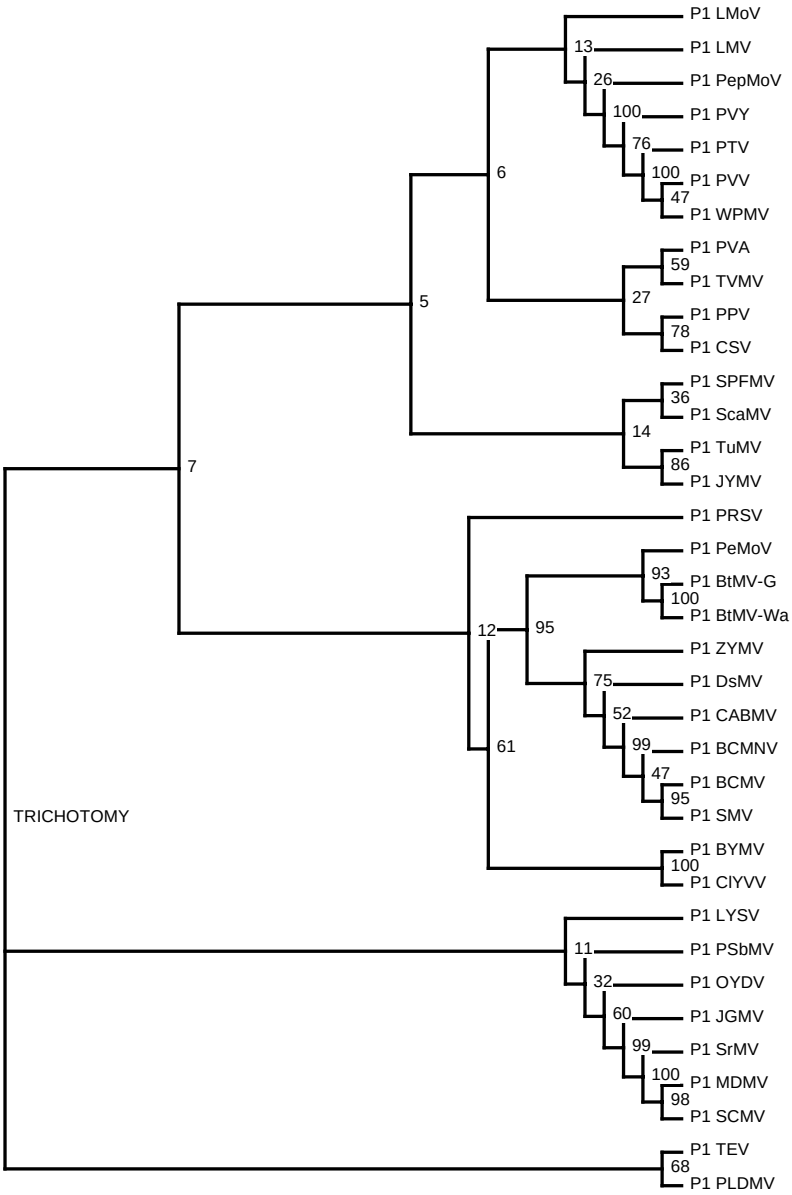

Figure 20: Phylogenetic tree illustrating the position of BtMV-G among members of the genus *Potyvirus* based on the P1 aa sequence.

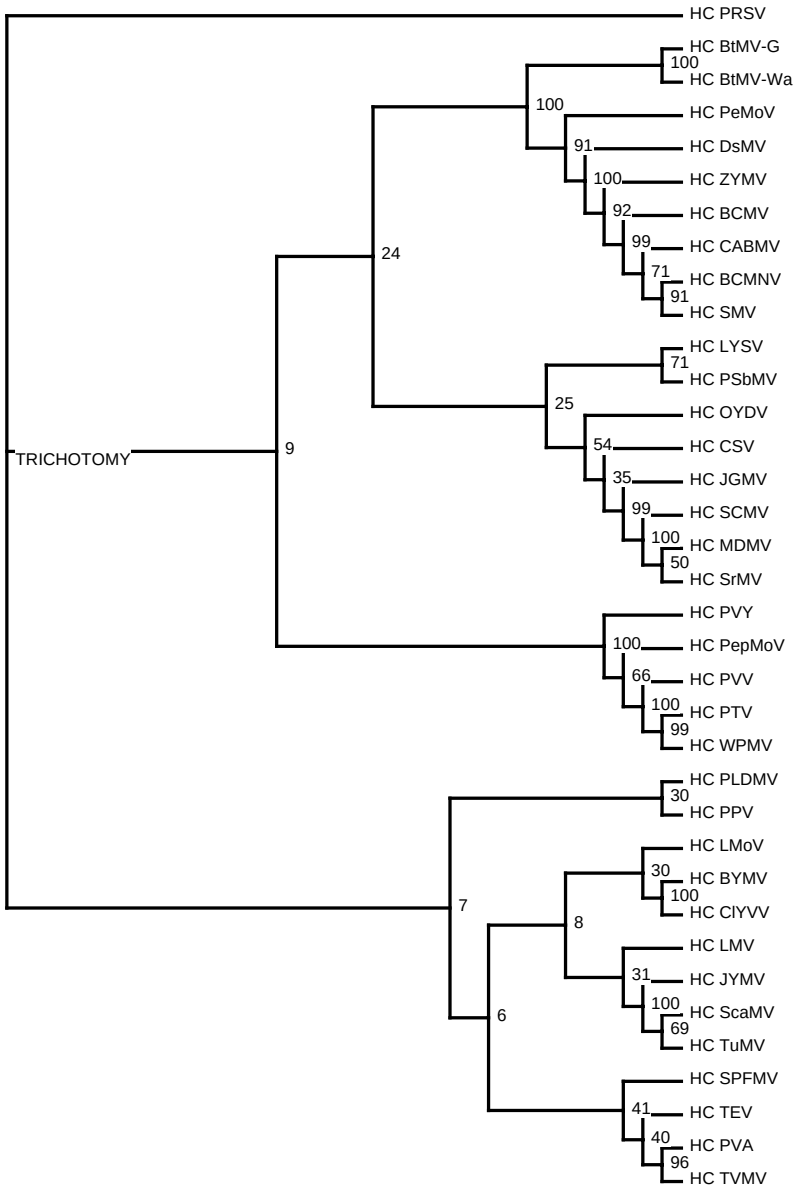

Figure 21: Phylogenetic tree illustrating the position of BtMV-G among members of the genus *Potyvirus* based based on the HC-Pro aa sequence.

3.5.4.3. P3 sequence analysis

The BtMV-G P3 consists of 347 amino acids and has a molecular weight of 40.31 kDa. The alignment of the amino acid sequence with other potyviruses revealed no significant

identity and showed that this protein belongs to a less conserved region in the potyviruses.

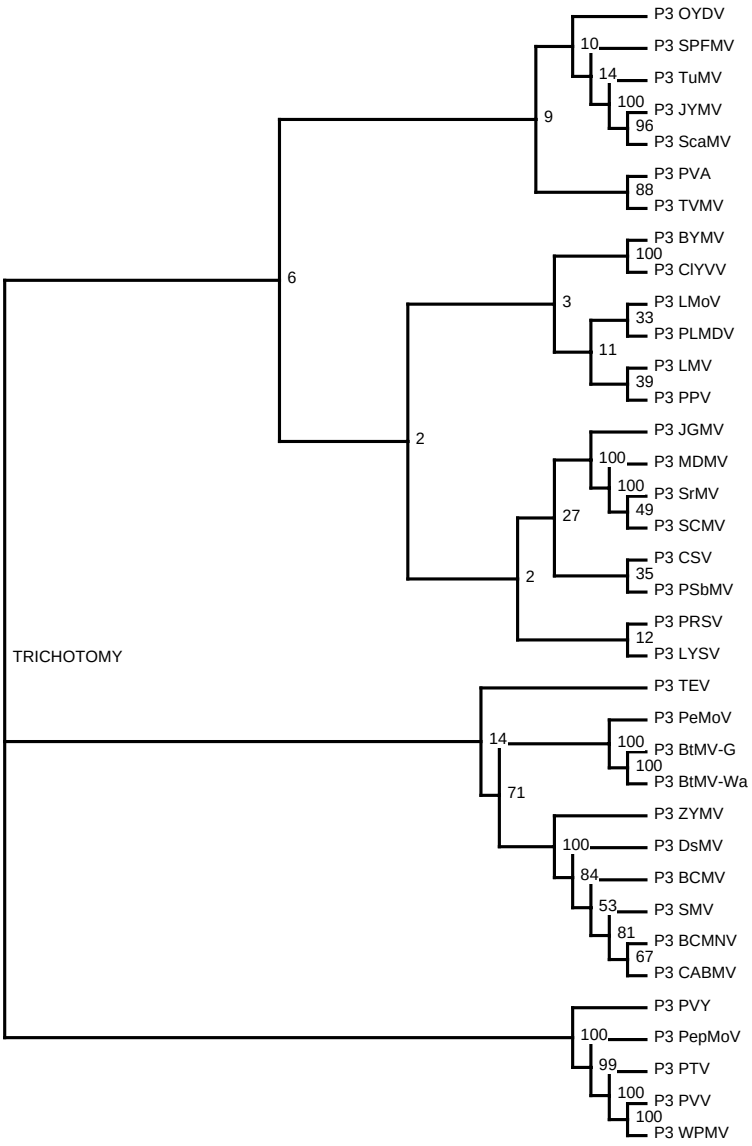

Figure 22: Phylogenetic tree illustrating the position of BtMV-G among members of the genus *Potyvirus* based on the P3 aa sequence.

No amino acid motifs were identified except a few conserved amino acids distributed through the sequence like proline (P_{812}), glutamine (Q_{861}), glycine (G_{910}) and lysine (K_{924}). The cleavage site G/G between P3 and HC-Pro is located in the conserved region KXYXV-G/G at the C-terminal end of HC-Pro amino acid sequence at amino acid position 765-771.

The BtMV-G P3 appeared to be highly variable on the amino acid level among potyviruses. The percentage identity of 88% with the BtMV-Wa isolate was lower compared to other proteins of BtMV-G and BtMV-Wa. The identity of BtMV-G P3 with the P3 of other potyviruses ranged from 35% (PeMoV) to 13% (OYDV). An identity of 27% of the BtMV-G P3 was found with LMoV and PPV. The phylogenetic tree of P3 places BtMV-G, BtMV-Wa and PeMoV in one cluster. The tree also showed a close relationship with the BCMV subgroup and TEV as shown in Fig. 22.

3.5.4.4. 6K1 sequence analysis

The predicted amino acid sequence of 6K1 is 52 amino acids with a molecular weight of 5.9 kDa. This is a usual length for a member of the potyviruses. The cleavage site between 6K1 and CI is located at amino acid position 1803-1804 and consists of Q/S dipeptide. Amino acid sequence alignments showed that BtMV-G shared a conserved amino acid motif at the N-terminus with the other potyviruses (E_{1127}-X_3-A_{1131}-X_6-M_{1138}-X6- S_{1145}-X_6-L_{1152}).

The small protein 6K1 appeared to be a slightly conserved region among potyviruses and the percentage of identity of the amino acid sequence rarely exceeded 50% in comparisons to most of the potyvirus members as PeMoV (55%), BCMNV and OYDV (53%). It was found that the 6K1 is the only protein with the same amino acid sequence as the BtMV-Wa isolate. The lowest identity was found with 26% to MDMV and SCMV.

The phylogenetic tree indicated that 6K1 of BtMV-G is closely related to the 6K1 of PeMoV and OYDV and grouped these viruses together with the BtMV-Wa isolate in one cluster, located beside the cluster of CSV and PSbMV and the cluster of the BCMV subgroup (Fig. 23).

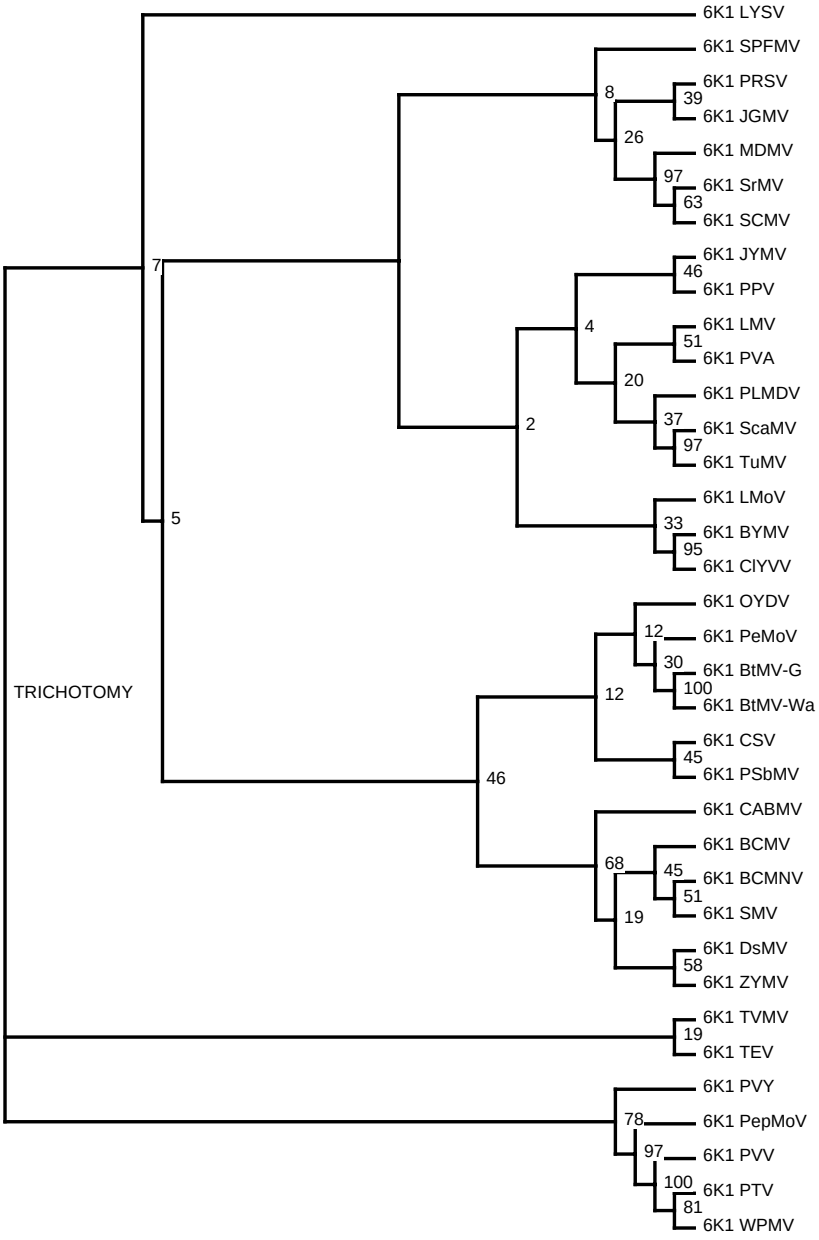

Figure 23: Phylogenetic tree illustrating the position of BtMV-G among members of the genus *Potyvirus* based on the 6K1 aa sequence.

3.5.4.5. CI sequence analysis

The CI consists of 634 amino acids and a molecular weight of 71.1 kDa. This protein is cleaved from the 6K1 at the cleavage site Q/S at amino acid position 1169-1170. Amino

acid sequence alignments indicated that the CI protein has 5 conserved motifs in its N-terminal half. These are proposed to have important functions in the viral replication process, including a RNA helicase activity motif (DESH) at amino acid position 1343-1346 and a RNA binding motif (LKVSATPP) at amino acid position 1369-1376. A motif involved in ATPase activity (GERLQRLGRVGR) is located at amino acid position 1516-1527, a motif with NTP binding activity (VGSGKST) at amino acid position 1256-1263 and a motif involved in NTP hydrolysis (VATNIIENGVTL) at amino acid position 1472-1484. However, the amino acid cysteine or phenylalanine of the RNA helicase motif of potyviruses ($DE_{C\&F}H$) is replaced with serine in BtMV-G.

Pairwise comparison on the amino acid level revealed the highest percentage identity with the BtMV-Wa isolate (98%) followed by PeMoV (67%), and BCMV, SMV and DsMV (60%). The lowest identity was observed with JGMV (50%).

The phylogenetic tree indicated that CI protein of BtMV-G is most closely related to CI of PeMoV. Together with the BtMV-Wa isolate BtMV-G and PeMoV are located in one cluster, which is closest to the BCMV cluster (Fig. 24).

3.5.4.6. 6K2 sequence analysis

The second small protein 6K2, with a number of 52 amino acids and a molecular weight of 5.8 kDa, is a slightly conserved region like the 6K1. In the aligned amino acid sequences of potyviral 6K2 proteins the 6K2 protein of BtMV-G shows no conserved amino acid motifs. In addition to the conserved cleavage site between P3 and 6K2 consisting of a Q/A junction at amino acid position 1117-1118 only two conserved amino acid residues with asparatic acid (D_{1826}) and glycine (G_{1836}) were found.

The BtMV-G 6K2 amino acid sequence has the highest identity with the BtMV-Wa isolate (92%) followed by PeMoV (58%). The lowest identity was recognized with PSbMV and DsMV (30%). Phylogenetic analysis revealed BtMV-G, BtMV-Wa, PeMoV and LMV in one cluster (Fig. 25).

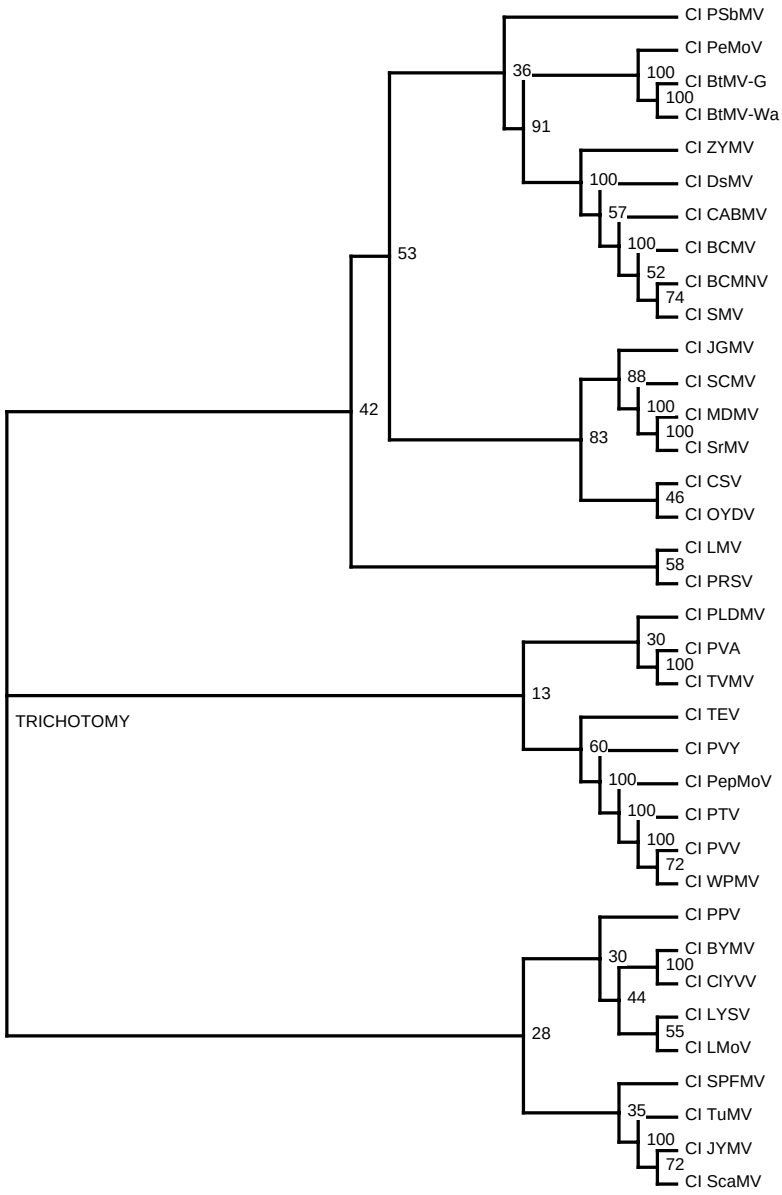

Figure 24: Phylogenetic tree illustrating the position of BtMV-G among members of the genus *Potyvirus* based on the CI aa sequence.

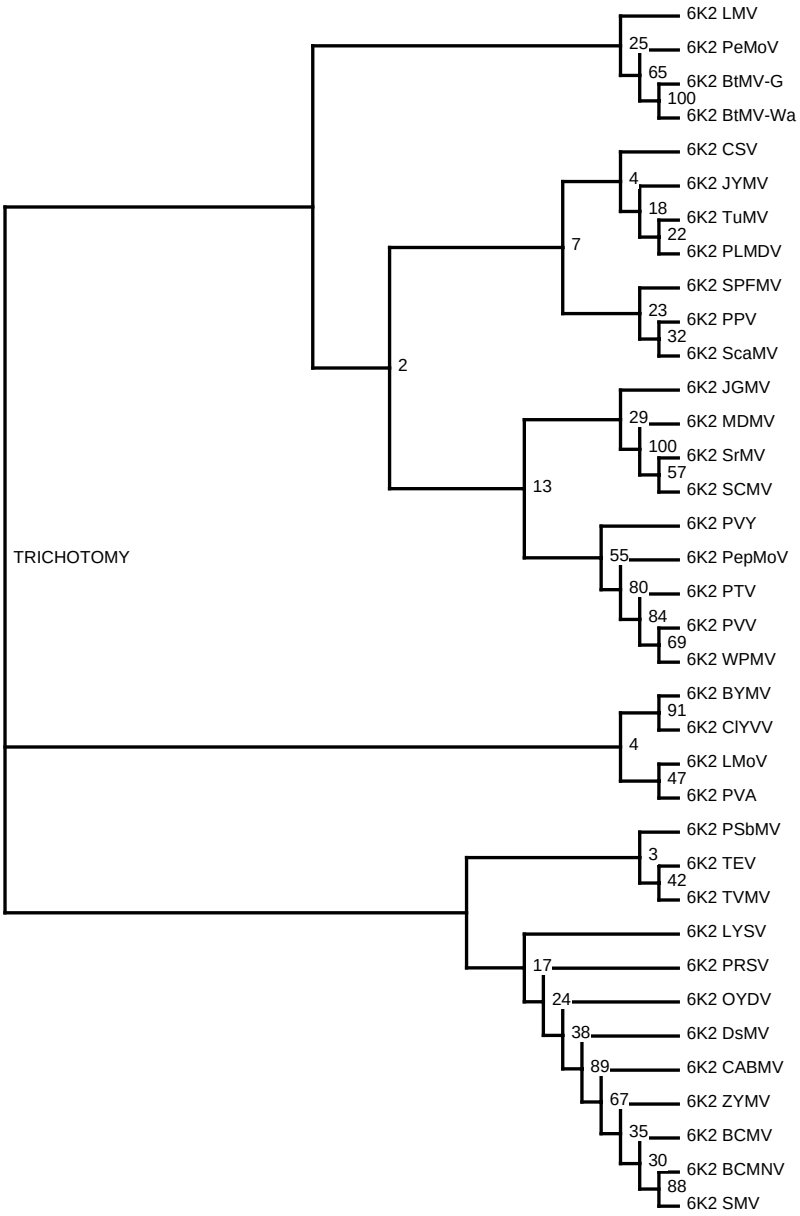

Figure 25: Phylogenetic tree illustrating the position of BtMV-G among members of the genus *Potyvirus* based on the 6K2 aa sequence.

3.5.4.7. VPg sequence analysis

The predicted VPg amino acid sequence length is 191 aa with a molecular weight of 21.7 kDa. It is released from the 6K2 at a Q/G cleavage site at amino acid position 1855-1856. This cleavage site was found to be identical to those of the other potyviruses.

Multiple alignment of the amino acid sequence of BtMV-G with other potyviruses showed a higher identity in the N-terminus region of the VPg than in the C-terminal region. The result from pairwise comparisons of the VPg amino acid sequences of BtMV-G with those of other potyviruses showed that the VPg is a variable region with identities ranging from 61% (ZYMV and PeMoV) to 41% (PTV). The BtMV-G shared the highest identity with BtMV-Wa isolate (97%) on the amino acid level. The phylogenetic tree of potyvirus-VPg amino acid sequences was constructed. It indicated that BtMV-G is closely related to BtMV-Wa and PeMoV. As a consequence the isolates of BtMV and PeMoV are grouped in one cluster, which is close to the BCMV cluster (Fig. 26).

3.5.4.8. NIa sequence analysis

The predicted NIa amino acid sequence length is 247 aa with a molecular weight of 27.8 kDa. It is cleaved from the VPg at E/G cleavage site at amino acid position 2046-2047. The same junction is found for DsMV, JGMV, LYSPV, PLDMV, PRSV and TEV. However this VPg/NIa junction is distinct from those of most other potyviruses, which have a E/S dipetide. The alignment showed that BtMV-G NIa contains the conserved amino acid sequence of the protease domain (H_{2092}-X_{34}-D_{2127}-X_{69}-C2198-G_{2199}) which is required for the proteolytic process of the potyvirus polyprotein.

Amino acid sequence comparison with other NIa proteins of potyviruses indicated that the NIa is a relatively conserved region at the amino acid level among potyviruses. The highest identity was found with the BtMV-Wa isolate (95%) and the lowest one with BYMV, CSV, JGMV, PRSV and SCMV (39%). The amino acid sequence of the predicted NIa region revealed a high identity with ZYMV (56%) and a slightly lower identity with PeMoV (55%), DsMV (53%) and SMV (51%).

The phylogenetic tree grouped BtMV-G, BtMV-Wa and PeMoV in one cluster, indicating that these viruses are closely related to each other. The tree also illustrated that there is a clear relationship between the NIa of the BtMV cluster and the BCMV cluster on the molecular level (Fig. 27).

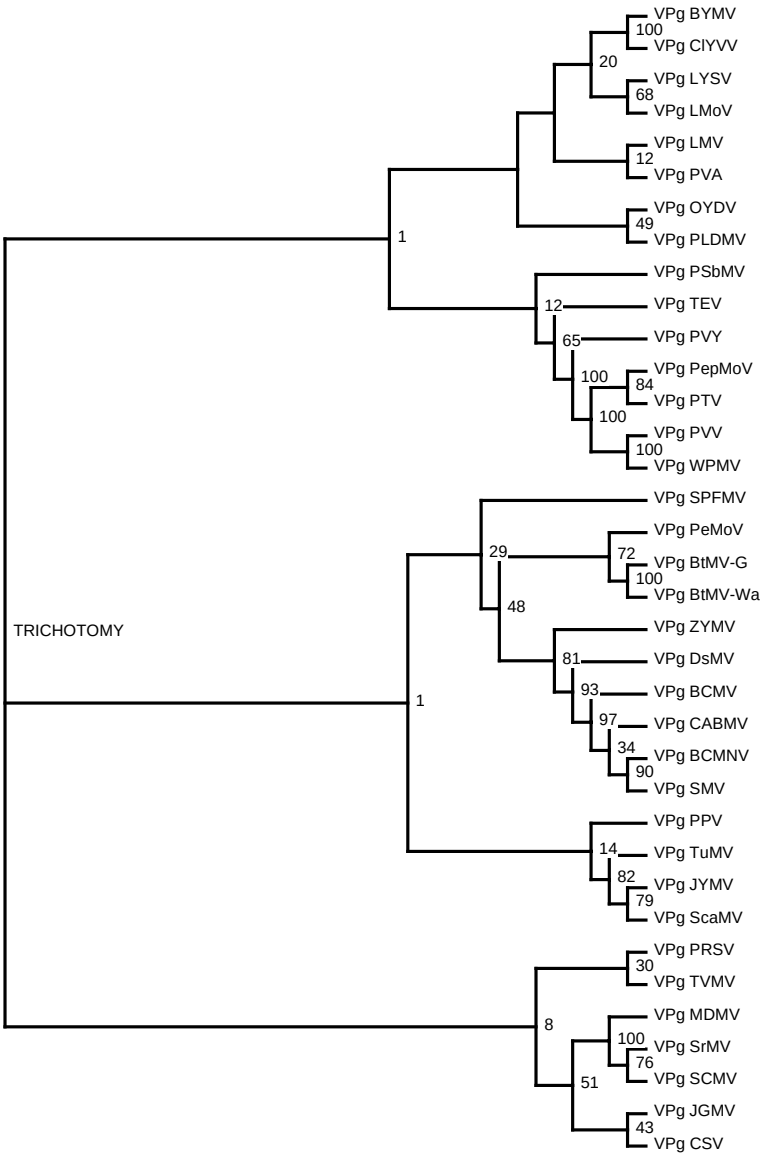

Figure 26: Phylogenetic tree illustrating the position of BtMV-G among members of the genus *Potyvirus* based on the VPg aa sequence.

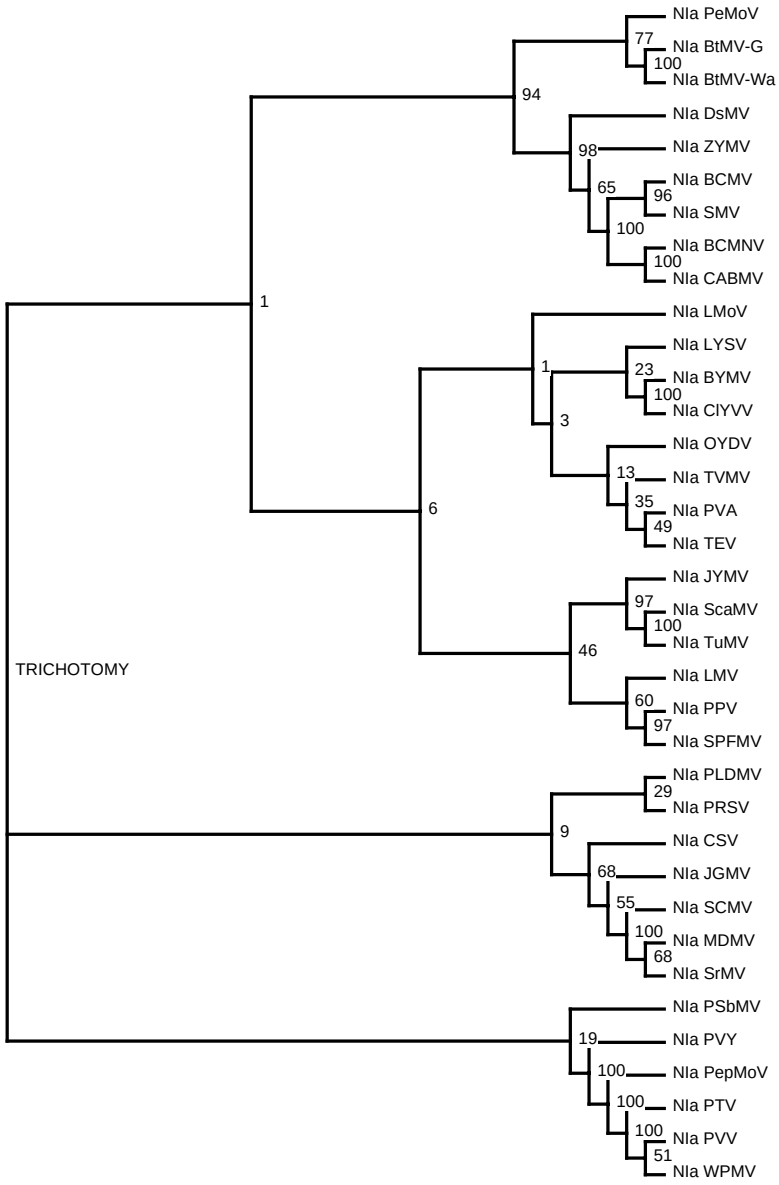

Figure 27: Phylogenetic tree illustrating the position of BtMV-G among members of the genus *Potyvirus* based on the NIa aa sequence.

3.5.4.9. NIb sequence analysis

The NIb consists of 516 amino acids and has a molecular weight of 59.5 kDa. The proteolytic cleavage site to the NIa is Q/G at amino acid position 2293-2294. This cleavage site dipeptide is also found in TEV, PeMoV, DsMV, JGMV, LYSPV and PLDMV. Sequence analysis showed that NIb protein is very highly conserved region among

potyviruses. The NIb contains the conserved amino acid sequence motifs CDADGS and SGX3SX3NTX30GDD, which are proposed to be responsible for the RNA dependent RNA polymerase activity. The motifs are located at amino acid position 2538-2544 and 2603-2646, respectively. However, the amino acid threonine found in most potyviruses was replaced by serine (amino acid position 2608). The GDD motif, probably involved in replication of the virus, is located near the NIb C-terminus at amino acid position 2644-2646.

The amino acid pairwise comparison with the NIb of other potyviruses showed that BtMV-G NIb is very closely related to BtMV-Wa (97%). The amino acid sequence percentage identity ranged from 56% with OYDV, PRSV and PVA to 67% with PeMoV. Identity to ZYMV and PPV was determined with 62%, to SPFMV, LMV, DsMV, JYMV and TEV with 651 followed by BCMNV, BYMV, LMoV, TuMV and SMV with 60% and to SCMV PSbMV and ScaMV with 59%.

The phylogenetic tree of potyviruses based on the NIb aa sequence indicated that BtMV-G is closely related to PeMoV. In the tree BtMV-G, BtMV-Wa and PeMoV were found in one cluster, which is near to the cluster of BCMV (Fig. 28).

3.5.4.10. CP sequence analysis

The CP of BtMV-G comprises of 276 amino acids. The molecular weight of the CP was calculated with 31.1 kDa, which is in accordance with the molecular weight determined by SDS-PAGE. The proteolytic cleavage site at the border to NIb consists of Q/G at amino acid position 2809-2810. Sequence comparisons revealed a DAG tripeptide close to the N-terminus of the BtMV-G CP at amino acid position 23-25. This motif is essential for the aphid transmissibility of potyviruses. Another highly conserved motif was found with RX43D at amino acid positions 2975-3019, which is essential for long distance movement in potyviruses. The BtMV-G CP revealed the highest identity on the amino acid level with the isolate BtMV-Wa (98%) followed by BCMV and SMV with 60%, BCMNV and PepMoV with 59%, PVY with 58% and CABMV, JYMV and PTV with 57%. The overall identity between BtMV-CP was lowest with the CPs of SPFMV and CSV (46%).

The phylogenetic tree of potyviruses based on the CP sequence (Fig. 29) mirrors BtMV-G as closely related to the BtMV-Wa isolate and PeMoV. This group is located in the neighbourhood of the BCMV cluster.

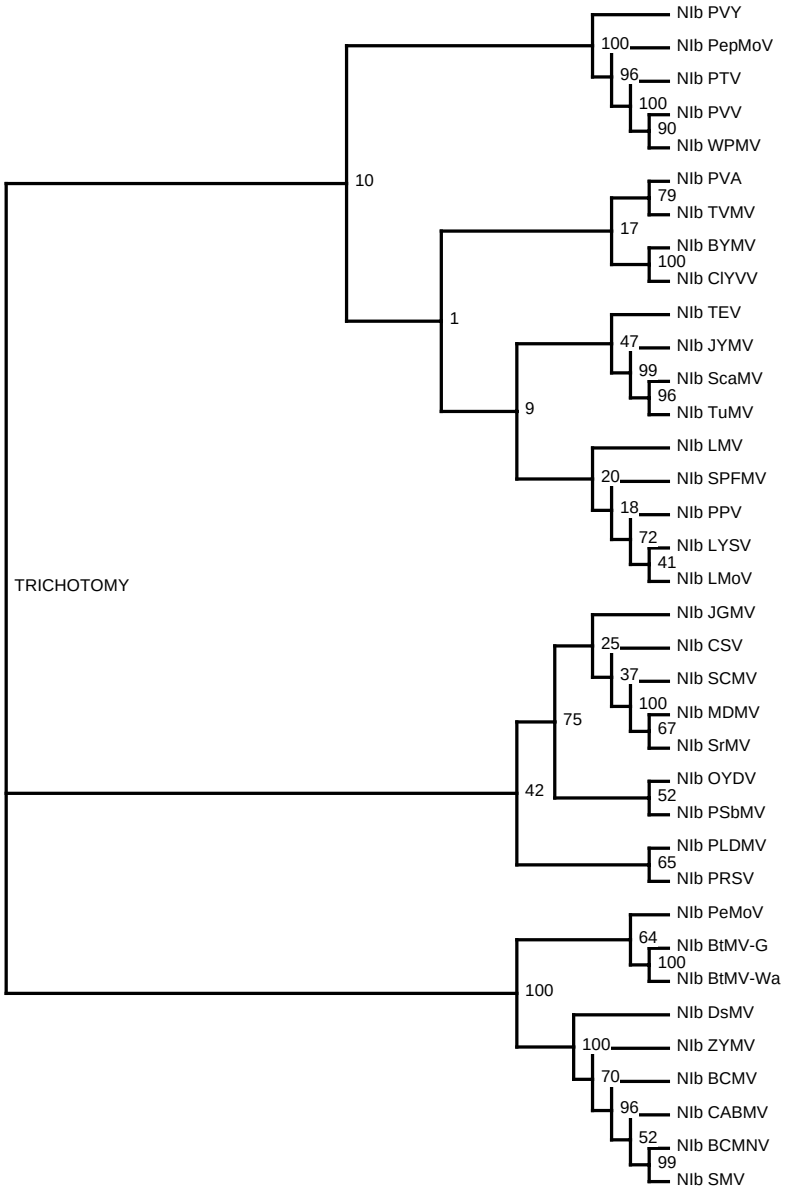

Figure 28: Phylogenetic tree illustrating the position of BtMV-G among members of the genus *Potyvirus* based on the NIb aa sequence.

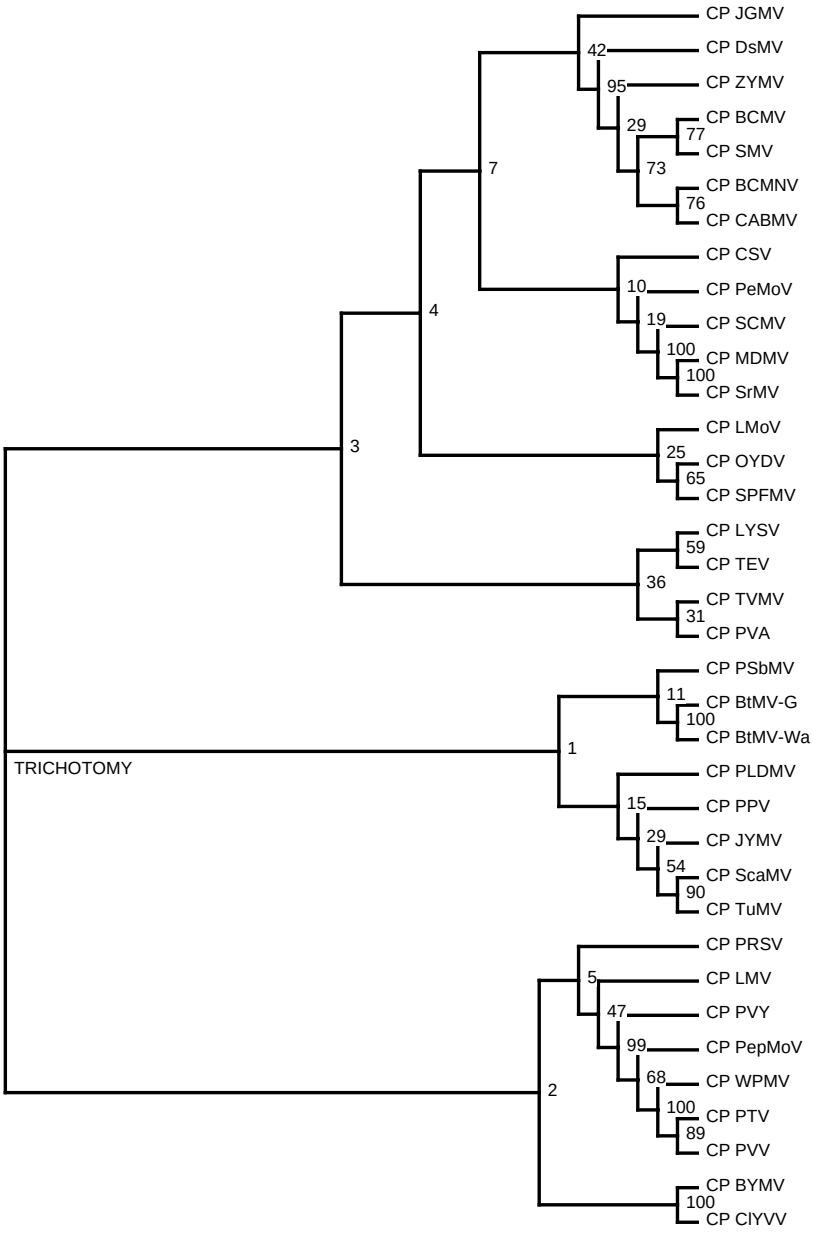

Figure 29: Phylogenetic tree illustrating the position of BtMV-G among members of the genus *Potyvirus* based on the CP aa sequence.

3.6. Construction of a full-length clone of the German isolate of BtMV

An efficient strategy to construct a full-length clone of BtMV-G involves the use of the CaMV 35S promoter and terminator to generate infectious in vivo transcripts of BtMV. The strategy was designed to clone the BtMV-G genome as four overlapping fragments that could be joined at unique endonuclease restriction enzyme sites. The strategy is outlined in Fig. 30.

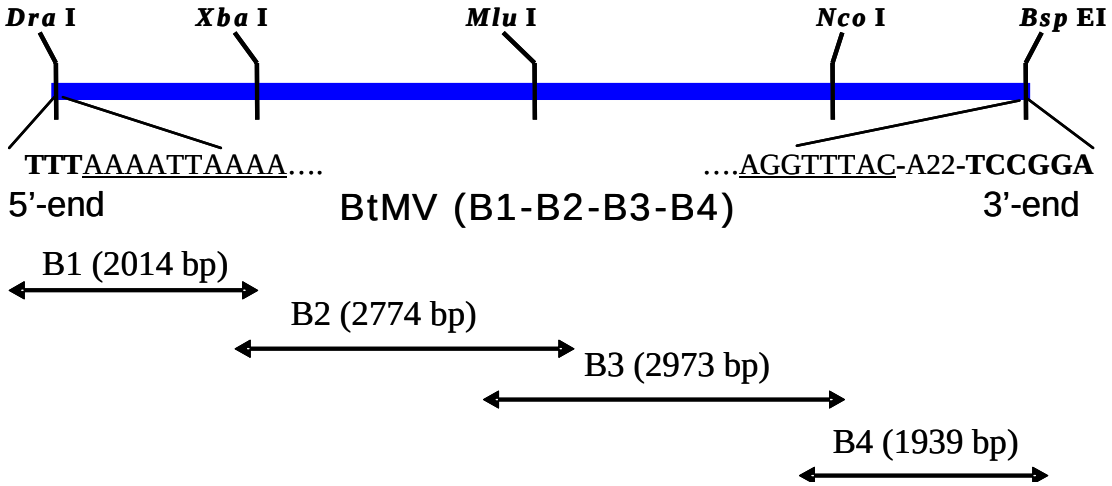

Figure 30: Generation of four BtMV PCR fragments (B1 to B4 indicated by lines with double arrows and fragment sizes in bp) covering the complete BtMV-G genome, represented by the blue bar on the top. Endonuclease restriction sites used for later cloning are indicated. Underlined sequences represent the 5'-end and the 3'-end of BtMV-G genome, respectively. Bold letters indicate additional nucleotides to generate recognition sequences of restriction enzymes *Dra* I and *Bsp*E I. A22 = 22 Adenosine residues.

3.6.1. Generation of four overlapping clones

The complete genome sequence of BtMV-G (9592 nt) was divided into four overlapping fragments according to the three single restriction sites (*Xba* I, *Mlu* I, *Nco* I), which are present in the genomic sequence. The first fragment (B1) contains a *Dra* I restriction site which was created by using the oligonucleotide B1s-*Dra* I (Tab. 3.) at the 5'end of the genome to add three additional T-residues to obtain the *Dra* I site. B1 started with three T-nucleotides added at the exact 5'-end and it ended at nucleotide 2008 (numbering according

to Fig. A1) including a *Xba* I restriction site (2014 nt; see Fig. 30). The second fragment (B2) started from nucleotide 1984 and ended at 4752 including the *Xba* I and *Mlu* I single restriction sites (2774 nt), while the third fragment (B3) started from nucleotide 4731 and ended at 7711 including *Mlu* I and *Nco* I single restriction sites (2973 nt). The fourth fragment (B4) started from nucleotide 7687 and ended with a poly-A tail (22A) including *Nco* I and *Bsp*E I restriction sites (1939 nt). These four fragments covered the complete genomic RNA sequence of BtMV-G with three T-residues at the 5'end and 22 A-residues as a poly-A tail followed by the restriction site *Bsp*E I at the 3'end (Fig. 30).

Four specific antisense primers (C1, C2, C3, C4) were designed based on the BtMV-G complete sequence for cDNA synthesis (Tab. 3.). These four primers were used separately to synthesise the cDNA starting from the purified RNA by using the Expand Reverse Transcriptase system as described in section (2.2.11.1.).

The overlapping products representing the BtMV-G genome were amplified from the first strand cDNA with primer pairs containing the suitable restriction sites, B1(B1s-*Dra* I & B1as-*Xba* I), B2 (B2s-*Xba* I & B2as-*Mlu* I), B3 (B3s-*Mlu* I & B3as-*Nco* I), B4 (B4s-*Nco* I & B4as-*Bsp*E I) (see Tab. 2.). PCR products were successfully generated and analyzed in agarose gels (Fig. 31).

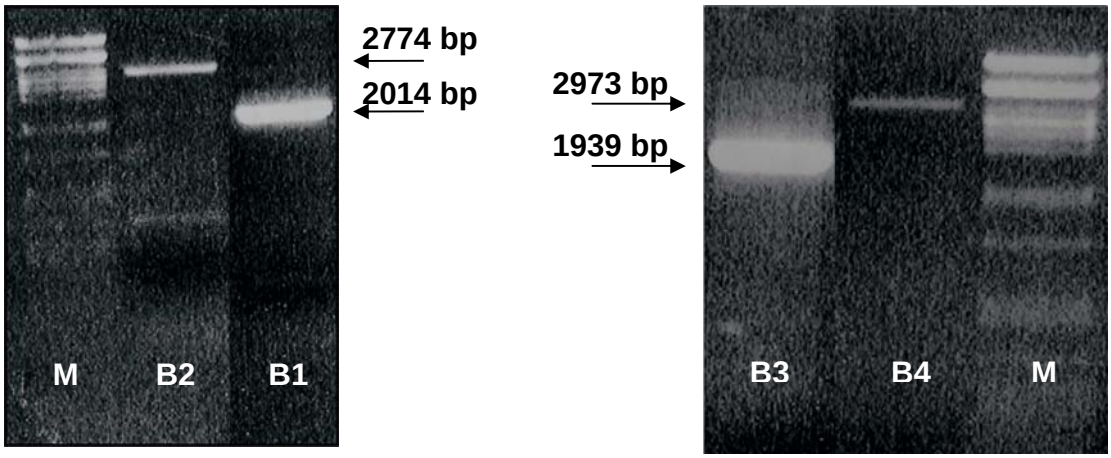

Figure 31: Four overlapping fragments (B1, B2, B3, B4) produced by PCR representing the entire BtMV genome. M = DNA size standard: Lambda-DNA digested with *Pst* I.

3.6.2. Cloning of BtMV-G overlapping fragments

The gel purified PCR products were ligated into the pGEM-T Easy vector as described in section (2.2.10.1) and transformed into competent *E.coli* NM522 cells (2.2.10.3.). The resulting plasmids were checked for the presence of the insert by colony PCR and digestion with restriction enzymes. For assembly of the full-length clone subcloning steps described below were carried out. In each cloning step 7 to 8 clones containing the fragments B1, B2, B3 or B4, respectively or combinations thereof were mixed before digestion and ligation.

3.6.3. Cloning steps for the construction of the BtMV-G full-length clone

The pGEM-T vectors containing B1 fragments were mixed and subsequently digested with *Dra* I and *Xba* I. The B1 fragment was extracted from the agarose gel and ligated into the *Stu* I / *Xba* I cut plasmid V209-pe35StupAII containing the *Cauliflower mosaic virus* (CaMV) 35S promoter, creating plasmid pe35B1 (Fig. 32)

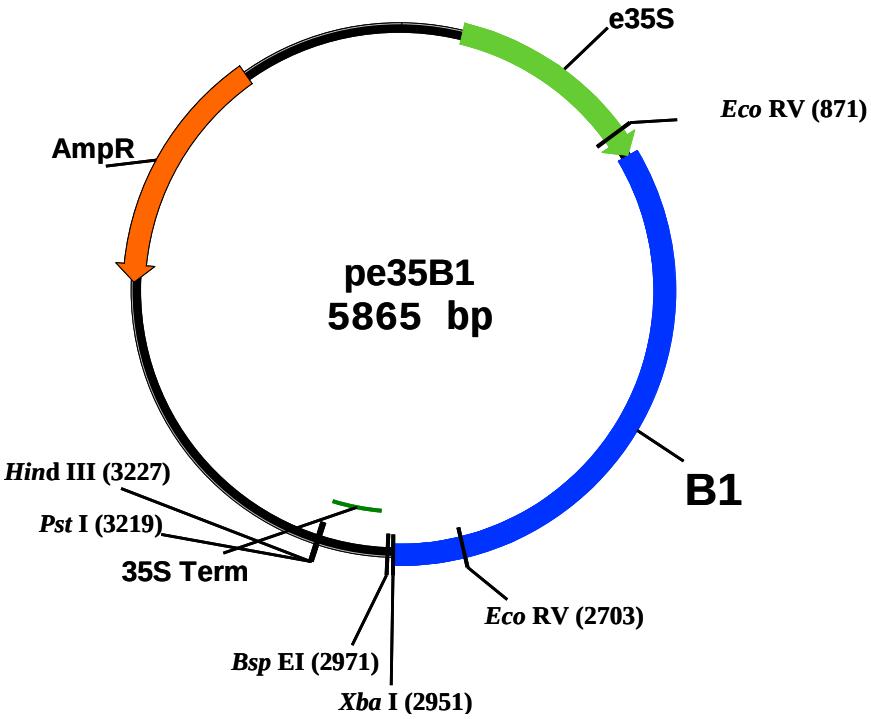

Figure 32: Plasmid pe35B1 with selected restriction endonucleases sites.
e35S = enhanced 35S CaMV promoter; B1 = BtMV-G (1995 bp);
35S Term = CaMV Termination signal; AmpR = ß-lactamase gene.

The pGEM-T vectors containing B4 were digested with *Nco* I and *Bsp*E I. The B4 fragments were ligated into the *Nco* I and *Bsp*E I digested plasmid V217-pBlueKSP and named V217-pBlueKSPB4. Subsequently V217-pBlueKSPB4 was digested with *Xba* I and *Bsp*E I to deliver B4 with an additional non-viral sequence (15 nts) and ligated into the plasmid pe35B1 digested with the same enzymes resulting in pe35B1B4 (Fig. 33).

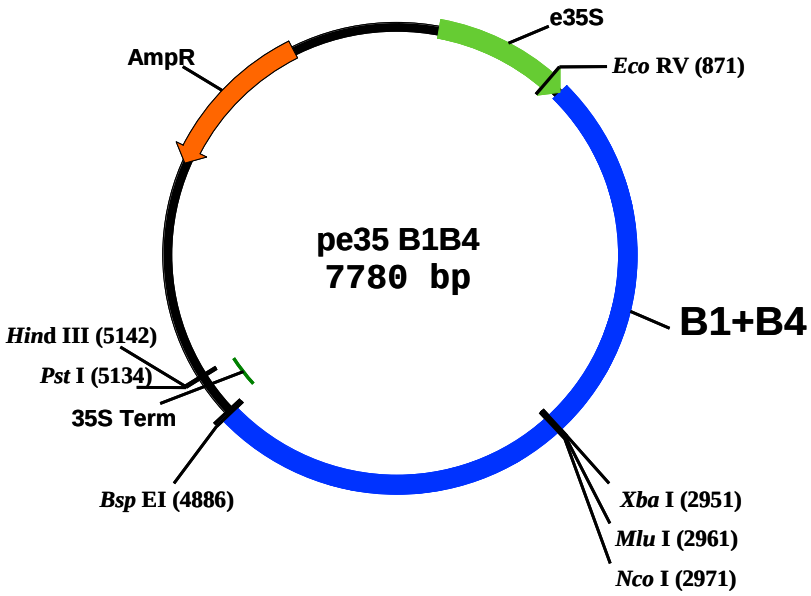

Figure 33: Plasmid pe35B1B4 with selected restriction endonucleases sites.
e35S = enhanced 35S CaMV promoter; B1+B4 = BtMV-G (3902 bp);
35S Term = CaMV Termination signal; AmpR = ß-lactamase gene.

The pGEM-T vectors containing B2 were digested with *Xba* I and *Mlu* I to retrieve the B2 fragments, which were ligated to the plasmid V223-pDRIVE2, cut with the same enzymes to create pDRIVE2-B2 (Fig. 34).

The pGEM-T vectors containing B3 were digested with *Mlu* I and *Nco* I to release B3 fragments, which were ligated to plasmid pDRIVE2-B2 cut with the same enzymes to create plasmids containing B2 and B3. This plasmid was named pDRIVE2-B2B3 (Fig. 35).

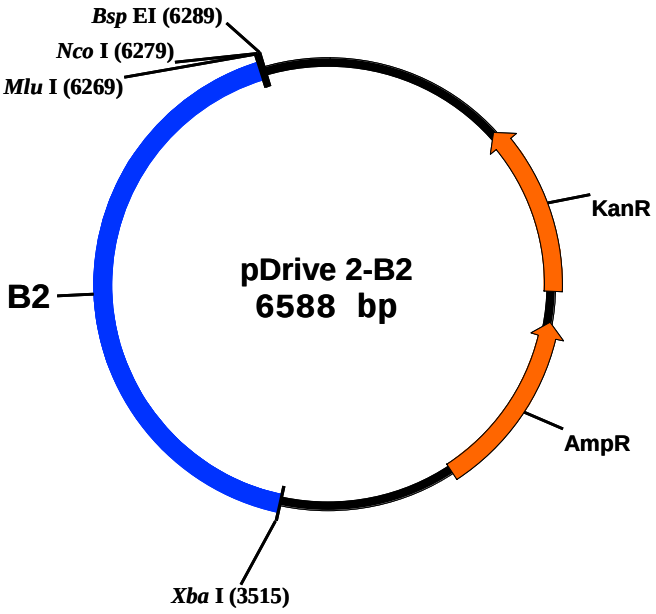

Figure 34: Plasmid pDRIVE2-B2 with selected restriction endonuclease sites.
B2 = BtMV-G (2760 bp); AmpR = ß-lactamase gene;
KanR = neomycin phosphotransferase.

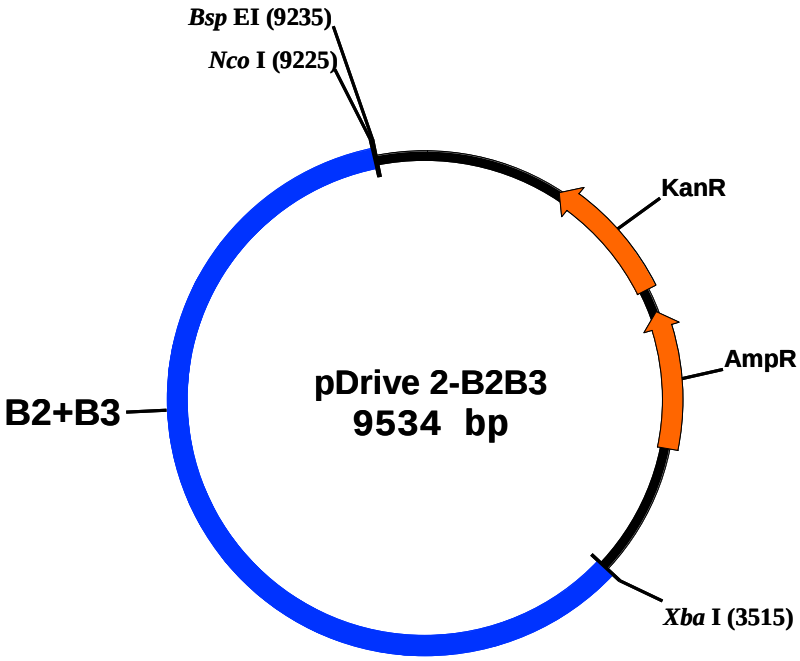

Figure 35: Plasmid pDRIVE2-B2B3 with selected restriction endonuclease sites.
B2B3 = BtMV-G (5716 bp); AmpR = ß-lactamase gene;
KanR = neomycin phosphotransferase.

In addition B3 fragments generated with *Mlu* I and *Nco* I (see above) were ligated to V223-pDRIVE2 cut with the same enzymes to create plasmids containing B3 named pDRIVE2-B3 (Fig. 36).

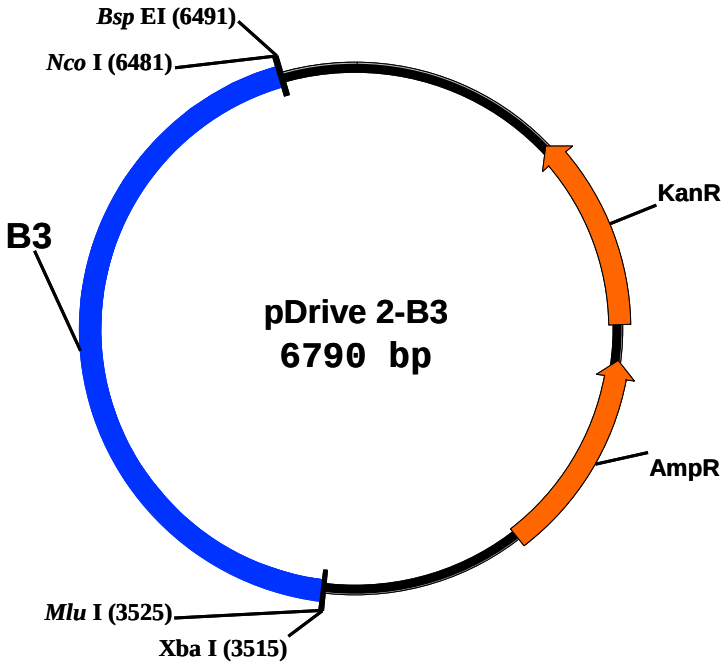

Figure 36: Plasmid pDRIVE2-B3 with selected restriction endonuclease sites.
B3 = BtMV-G (2962 bp); AmpR = ß-lactamase gene;
KanR = neomycin phosphotransferase.

The plasmids pDRIVE-B3 were digested with *Mlu* I and *Nco* I. The released fragment B3 was ligated to pe35-B1B4, digested with the same enzymes to create a plasmid containing B1, B3 and B4. This plasmid was named pe35-B1B3B4 (Fig. 37).

The plasmids pe35-B1B3B4 were digested with *Xba* I and *Nco* I. The released fragment B3 was replaced by fragments B2-B3 which were delivered from plasmids pDRIVE2-B2B3 digested with *Xba* I and *Nco* I. The ligation resulted in plasmids containing the entire BtMV-G sequence. These plasmids were named pe35BtMV (Fig. 38).

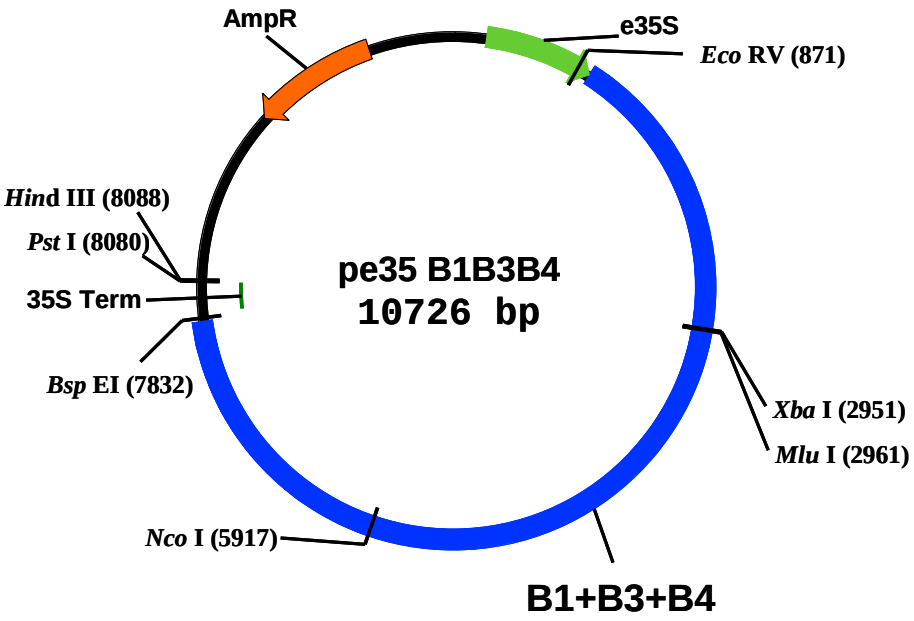

Figure 37: Plasmid pe35-B1B3B4 with selected restriction endonucleases sites. e35S = enhanced 35S CaMV promoter; B1+B3+B4 = BtMV-G (6848 bp); 35S Term = CaMV Termination signal; AmpR = ß-lactamase gene.

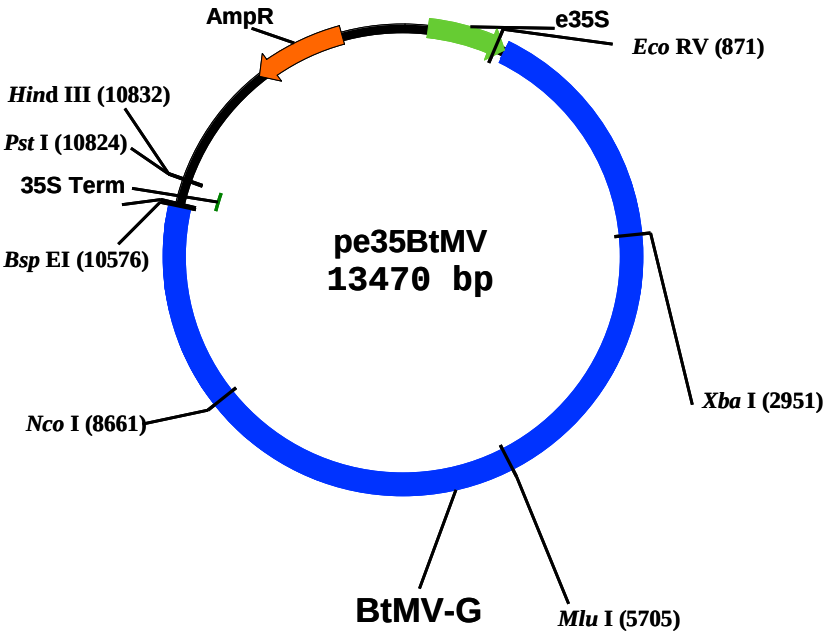

Figure 38: Plasmid pe35BtMV with selected restriction endonucleases sites. e35S = enhanced 35S CaMV promoter; B1+B2+B3+B4 = BtMV-G (9592 bp); 35S Term = CaMV Termination signal; AmpR = ß-lactamase gene.

3.6.4. Analysis of the BtMV-G full-length clone

After the last cloning step 72 colonies were screened by restriction digestion and colony PCR. Seven plasmids were selected depending on the colony PCR results. The insert size in the plasmids was characterized by restriction enzyme analysis to confirm the integrity of the full-length clone containing the four fragments. The selected constructs were digested with *Eco*R I (Fig. 39-A). Clone 2 appeared to be different in comparison to the other six clones (Fig. 39-A). However, when clone 2 was digested with the restriction enzymes *Xho* I, *Pst* I, *Eco*R V, *Sty* I, *Dra* I and *Pvu* II fragments of the expected sizes appeared (Fig. 39-B). The fragment patterns were in accordance with the determined BtMV-G sequence.

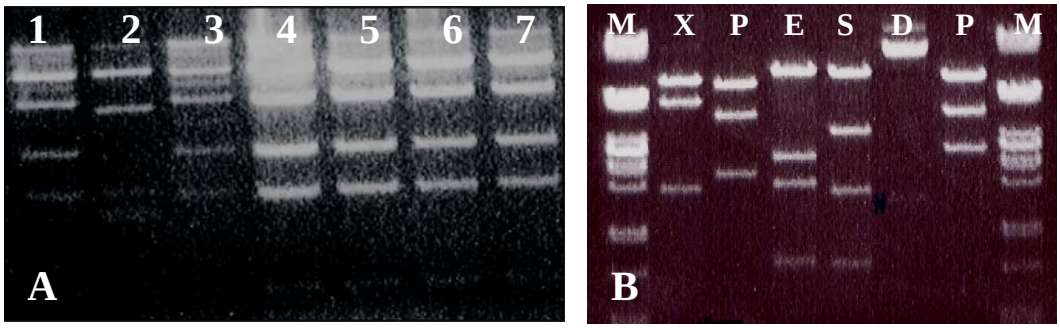

Figure 39: (A) Restriction analysis with *Eco*R I of seven individual BtMV-G full-length clones (lanes 1 to 7) used for inoculation. (B) Restriction analysis of full-length clone 2. From left to right: M = DNA size standard: Lambda-DNA digested with *Pst* I. X = *Xho* I, P = *Pst* I, E = *Eco*R V, S = *Sty* I, D = *Dra* I and P = *Pvu* II.

The nucleotide sequence of one full-length clone was determined in order to verify the construction of the BtMV-G full-length clone. Sequence analysis of the complete full-length clone revealed 27 nucleotide and 19 amino acid changes compared to the originally determined sequence of BtMV-G (see Fig. A1). Therefore, the sequence alignment showed a high sequence identity (99%) at the nucleotide and amino acid sequence level. As expected from the restriction enzyme digest with *Eco*R I an additional *Eco*R I restriction site was found in the full-length clone 2 compared to the original sequence.

3.6.5. Infectivity test and symptoms of the BtMV-G full-length constructs on *N. benthamiana* plants

The BtMV-G full-length clones were tested for infectivity and stability in the plant. The seven clones described above were selected, depending on the results of the restriction enzyme analysis and used to inoculate plants by particle bombardment. Particle bombardment delivers the DNA to the nucleus for initial transcription from the CaMV 35S promoter. The in vivo generated transcript of the cloned BtMV-G genome will be translated and infection will be established in the plant cells. The constructs were bombarded into young *N. benthamiana* plants, whereas four plants were inoculated with each clone. Only clone 2 was infectious and caused very clear mosaic symptoms on four plants (100%: 4/4) approx. 3 weeks after inoculation. The infection was confirmed by ELISA. The particle bombardment with the same clone was repeated with 20 plants. On the other hand, BtMV-G wild type virus was mechanically inoculated to serve as a control. 14 plants inculcated with the BtMV-G clone 2 were successfully infected (70%: 14/20) and developed symptoms, indistinguishable from wild type virus. However, the symptom development was slower than with the wild type virus and appeared after three weeks. Therefore the incubation time of the full-length clone infected plants is significantly delayed compared to that of plants infected with the native virus (Fig. 40-A, 40-B, 40-C).

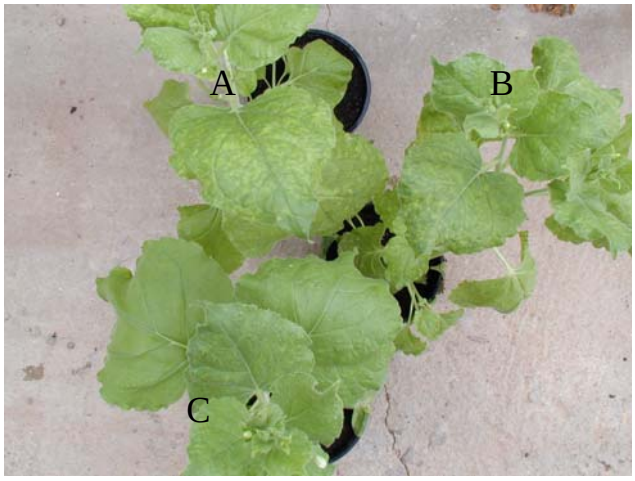

Figure 40: Symptoms in *Nicotiana benthamiana* plants (A,B) after particle bombardment with the full-length clone 2 of BtMV-G (pe35BtMV). The plant depicted with (C) is a healthy control. The picture was taken 21 days after particle bombardment.

The BtMV-G derived from the infectious full-length clone was tested for its stability in a transmission experiment. For that virus was mechanically transmitted from the primary infected plants to *N. benthamiana* plants. The plants showed typical systemic symptoms after one week like plants inoculated with the wild type virus (Fig. 41).

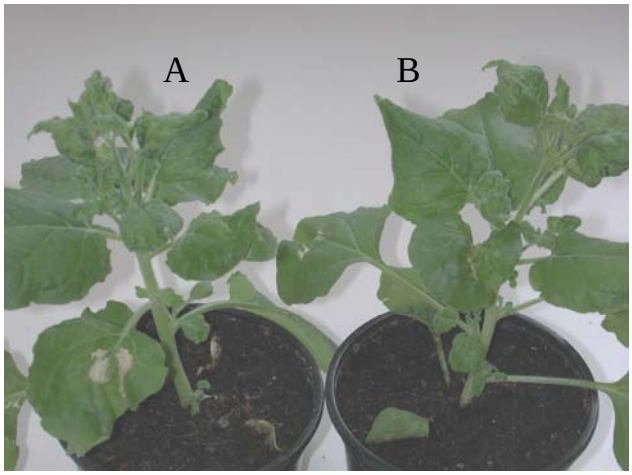

Figure 41: Symptoms in *Nicotiana benthamiana* plants caused by wild-type BtMV-G (A) and BtMV-G derived from the full-length clone 2 of BtMV-G after mechanical inoculation. The picture was taken 21 days after mechanical inoculation.

3.6.6. Host range and symptomatology of BtMV-G wild type and BtMV-G full-length clone.

Different plant species were used to test the infectivity and symptomatology of either wild-type or full-length clone derived BtMV-G (Tab. 7.). Five plants of each species were mechanically inoculated with the viruses. Symptoms appeared after 10 days on the susceptible plant species. Mosaic symptoms associated with leaf deformation, curled leaves and stunted plants were the most predominant symptoms. These symptoms appeared most obviously on *N. benthamiana* plants with both viruses (Fig. 42). Virus derived from the full-length clone was able to infect *Beta vulgaris* ssp. *vulgaris* var. *conditiva* cv. Rote Kugel 2 (Fig. 43) and sugar beet cultivar 6B22840 (Fig. 44), but failed to infect sugar beet cultivar 8T0015. Symptoms of the virus derived from the BtMV-G full-length clone were not as severe on *Beta vulgaris* ssp. *vulgaris* var. *cicla* cv. Lukullus plants compared with

the very clear symptoms caused by wild type virus (Fig. 45). In addition the virus from the infectious full length clone of BtMV-G was not able to infect *Spinacia oleracea* plants (Fig. 46) like the wild type virus.

Figure 42: Symptoms of wild-type BtMV-G (left) and BtMV-G derived from the full-length clone pe35BtMV in *Nicotiana benthamiana* (right). The picture was taken 14 dpi.

Figure 43: Symptoms of wild-type BtMV-G (left) and BtMV-G derived from the full-length clone pe35BtMV in *Beta vulgaris* ssp. *vulgaris* var. *conditiva* cv. Rote Kugel 2 (right). The picture was taken 14 dpi.

Figure 44: Symptoms of wild-type BtMV-G (left) and BtMV-G derived from the full-length clone pe35BtMV in *Beta vulgaris* ssp. *vulgaris* 6B22840 (right). The picture was taken 14 dpi.

Figure 45: Symptoms of wild-type BtMV-G (left) and BtMV-G derived from the full-length clone pe35BtMV in *Beta vulgaris* ssp. *vulgaris* var. *cicla* cv. Lukullus (right). The picture was taken 14 dpi.

Figure 46: Symptoms of wild-type BtMV-G (left) and BtMV-G derived from the full-length clone pe35BtMV in *Spinacia oleracea* cv. Matador (right). The picture was taken 14 dpi.

In *Atriplex hortensis* cv. Rheinische symptoms were hardly visible in case of virus derived from the BtMV-G full-length clone. Only small blots on the leaves occurred, whereas plants inoculated with wild type BtMV-G developed a severe infection (Fig. 47). This indicates that the virus derived from the BtMV-G full-length clone was able to propagate in this plant but failed to cause severe symptoms.

Figure 47: Symptoms of wild-type BtMV-G (left) and BtMV-G derived from the full-length clone pe35BtMV in *Atriplex hortensis* cv. Rheinische (right). The picture was taken 14 dpi.

The plants were checked by ELISA for the presence of the virus (Tab. 7.).

Table 7: Results of infectivity assay (ELISA) on plants inoculated with BtMV-G wild type virus and with BtMV-G derived from the pe35BtMV full-length clone.

Plant species	ELISA value (BtMV-G wild type) (Infect./inoculated plants)	ELISA value[1)] (BtMV-G from pe35BtMV) (Infect./inoculated plants)
Nicotiana benthamiana	2.760 ±0.157 (5/5)	1.952±0.982 (5/5)
Beta vulgaris ssp. *vulgaris* var. *conditiva* cv. Rote Kugel 2	2.449±0.628 (5/5)	1.446 (1/5)
Beta vulgaris (6B22840)	1.241±0.424 (5/5)	1.212±0.685 (2/5)
Beta vulgaris (8T0015)	0.782±0.828 (3/5)	---[2)] (0/5)
Beta vulgaris ssp. *vulgaris* var. *cicla* cv. Lukullus	1.644±0.888 (5/5)	0.195±0.775 (3/5)
Spinacia oleracea	2.62±0.979 (5/5)	--- (0/5)
Atriplex hortensis cv. Rheinische	2.979±0.473 (5/5)	2.803±0.373[3)] 2/5

[1)] ELISA value: the average of infected plants ± standard deviation. [2)] --- no ELISA value determined. [3)] Samples were taken from leaves showing small blots.

4. Discussion

4.1. Overview

The aim of this study was to obtain the complete nucleotide sequence of BtMV-G as a member of the genus *Potyvirus* and to compare the sequence with other potyvirus sequences, which are available in the GenBank. In addition an infectious full-length clone was constructed in order to provide a possibility to study the virus multiplication cycle and to obtain thorough understanding of the molecular biology of BtMV.

The points studied included the symptomatology of BtMV-G, its aphid transmissibility, viral nucleic acid extraction and cloning. RH-PCR and poly-G tailing techniques were adopted to determine the nucleotide sequence of HC-Pro, P1 and the extreme 5'-end of BtMV-G. Sequence analyses of BtMV-G were performed and comparisons were done with other *Potyvirus* genomes. In addition a BtMV-G infectious full-length clone was generated and the host range and symptomatology of the BtMV-G derived from the full-length clone was compared with the wild type BtMV.

4.2. Symptoms of BtMV-G on *B. vulgaris*, *N. benthamiana* and *C. quinoa* plants.

Clearly visible systemic symptoms were observed on leaves of *B. vulgaris* and ***N.*** *benthamiana* plants after inoculation with BtMV-G. In case of *B. vulgaris* the leaves showed green mottle and mosaic symptoms and a characteristic light vein clearing, while normal growth of the plant was not severely affected. Russell (1971) stated that BtMV infected plants are sometimes stunted but severe distortion of the leaves is uncommon. Dusi (1999) indicated that the BtMV affected the overall growth of infected sugar beet. Other studies indicated that the growth of potyviruses-infected plants is often retarded. Nemchinov *et al*. (2004) demonstrated that another isolate of BtMV caused nearly the same symptoms on *B. vulgaris*. However, virus symptoms differ depending on the virus strain, virus source, host plant and environmental conditions. Russell (1971) reported that BtMV strains differ in virulence towards sugar beet. Gibbs and Harrison (1964) stated that BtMV symptoms in sugar beet might be confused with yellow mosaic and blotching caused by *Tomato black ring virus* and *Tobacco rattle virus*.

BtMV caused very clear mosaic symptoms on *N. benthamiana* with distortion of leaves. Many studies indicated that BtMV can infect plants of the *Solanaceae,* where it causes clear symptoms. *N. Benthamiana* turned out to be a susceptible and suitable host for BtMV.
The symptoms after aphid transmission were typical and indistinguishable from the symptoms delivered after mechanical inoculation.
BtMV-G caused chlorotic local lesions in *C. quinoa* as demonstrated by Russell (1971). There are only few species, which develop local lesion symptoms after BtMV infection like *Beta patellaris, Amaranthus retroflexus, A. caudatus, C. quinoa* and *Gomphrena globosa* (Russel, 1971).

4.3. RNA extraction from BtMV-G infected *N. benthamiana*

Preparation of viral RNA is a fundamental step before Reverse Transcription-PCR (RT-PCR). In order to find the most rapid and efficient procedure, three different protocols of viral RNA extraction were compared. The RNA was subsequently used for RT-PCR with a random hexamer PCR.
Morris and Dodds (1979) described a dsRNA extraction method as an efficient procedure to isolate dsRNA from virus infected plants. Extraction of dsRNA is a suitable method to obtain a highly pure viral RNA, since dsRNA is usually detected only in virus infected plants. The amount of dsRNA depends on the type of plant tissue used for extraction. A very limited amount of BtMV-G dsRNA was extracted from infected *N. benthamiana* plants and amplification of a BtMV-G specific DNA-fragment by RT-PCR confirmed the viral origin of the dsRNA. However, the dsRNA extraction method was found to be laborious and only dsRNA of poor quality and quantity was retrieved.
The hot phenol extraction method was used as described by Verwoerd *et al*. (1989). This method was efficient to extract total RNA from the plant and to give result with RT-PCR and RH-PCR, but sometimes the presence of left over phenol or chloroform can cause inhibition of enzyme activity in following procedures. In addition the presence of plant RNA can lead to amplification of non viral fragments with random or universal primers.
Virus purification followed by genomic viral RNA extraction was found to be the best in terms of giving the highest amount of viral RNA. Virus aggregation during purification did not occur and consequently a high virus yield was obtained. In our preparation method the

sucrose cushion centrifugation step was omitted in order to obtain a higher amount of the virus. As reported in many studies virus purification and RNA extraction was used successfully for many potyviruses. Hammond and Lawson (1988) used an efficient purification protocol to purify twelve different potyviruses utilizing a single purification protocol. Wang *et al*. (1992) purified PSbMV and the virus preparation was stable. Nemchinov *et al*. (2004) obtained a high concentration of the purified American isolate of BtMV (BtMV-Wa) from infected sugar beet leaves, but the preparation was not stable *in vitro*, which has been also reported by Russel (1971). In this study the purified BtMV-G preparation from *N. benthamiana* plants contained a single protein with a molecular weight of approximately 29 kDa. This result is different from the results of other studies, where the sizes of the capsid proteins of potyviruses were determined with approximately 35 kDa (Hollings and Brunt, 1981; Milne, 1988; Rogov *et al*., 1991). The molecular weight of the coat protein determined from the BtMV-G sequence is 31 kDa. The molecular weight of the purified BtMV-G is lower, presumably due to PAGE conditions, e.g. overload of the gel, or a partial degradation of the coat protein subunits by contaminant plant proteases.

The purified virus was suitable for preparing viral genomic RNA. A conventional method for RNA extraction resulted in pure RNA from the partially purified virus preparations by SDS-proteinase K digestion followed by phenol / chloroform extraction. The BtMV-G RNA preparation was stable at –20 ^{o}C.

4.4. Cloning and sequencing of the BtMV-G genome

Several methods failed to amplify the unknown region of the BtMV-G genome. Construction of primers according to homologous potyvirus genomic sequences also failed to give any new genomic RNA region, despite of 25 nucleotides in one approach.

Random Hexamer-PCR was successfully used to construct most of the unknown region of the genomic RNA. This quick method allows construction of the cDNA and dsDNA populations from small amounts of RNA and involves two PCR reactions: the first one to generate dscDNA from cDNA by using Klenow DNA-polymerase and a modified random hexamer primer, the second one to amplify the dscDNA by using a gene specific antisense primer and an universal primer complementary to the conserved part of the modified random hexamer primer.

PCR products were obtained in the range from 0.2 - 1.2 kb which were detectable on an agarose gel and suitable for subsequent cloning. This result is similar to that of Froussard (1992) with the exception that an additional gene specific antisense primer was used in the last amplification step to increase amplification efficiency and specificity for the viral dscDNA.
Two antisense primers at the end of each clone were designed to amplify additional parts of the RNA. Four clones with different lengths were generated by this method covering most of the RNA sequence. In addition the RH-PCR method worked also with total RNA extracted from infected plants by the method of Verwoerd *et al*. (1989). This result indicates that the modified method is able to generate fragments from unknown RNA sequences from pure RNA as well as from total RNA. A similar method, called tagged random hexamer amplification method, was applied by Wong *et al*. (1996) for generating fragments of a 180 kb plasmid isolated from *Sphingomonas*.
Two additional methods were used to amplify the extreme 5'-end of BtMV-G. Firstly, ligation of an adaptor was used as described by Potgieter *et al.* (2002) with some modifications. This method was not efficient to give the complete sequence of the 5'-end. One reason could be a ligation of the adaptor to an incomplete RNA molecule because of fragmentation of the RNA or the presence of the small genome-linked protein (VPg). However, the method generated a clone at the extreme 5'-end containing 585 bp except the first adenosine residue of the BtMV-G sequence. The method was successful to determine the complete nucleotide sequences of dsRNA viruses as described by Potgieter *et al*. (2002) and Attoui *et al*. (2000a; 2000b).

An efficient Poly-G tailing approach was used to generate the complete sequence of the BtMV-G 5'-end as described by Götz and Maiss (2002). Tailing of cDNA with dG by a terminal transferase was followed by PCR amplification with a gene-specific antisense primer and an oligo-dC15-primer. This approach produced two clones containing the extreme 5'-end sequence, which starts with four-adenosine residues (5'-AAAATTAAACAT…-3'). The same method was used by Kusov *et al*. (2001) to determine the poly-A tail length at the 3'-end of hepatitis A virus. All together the findings support the fact that the protocol can be used to obtain sequence information of the 5'- and 3'-ends of unknown RNA sequences.
Five overlapping clones resulted were sequenced and the aligned sequences cover the so far unknown region of BtMV-G. The sequence of BtMV-G confirmed findings of Nemchinov *et*

al. (2004), who determined the genome of the isolate BtMV-Wa, except that BtMV-G contains an additional adenosine residue at its 5'-end.

4.5. BtMV-G sequence analysis and comparison with other viruses of the genus *Potyvirus*

In this study the genome of BtMV-G isolate was sequenced and the primary structure of the BtMV-G genomic RNA was determined. The genomic sequence of BtMV was assembled. The sequence of BtMV-G revealed a single strand positive RNA of 9592 nucleotides in length excluding the 3'- terminal poly (A) tail. The BtMV-G genome sequence length is typical for most of the potyviruses sequences like *Cowpea aphid-borne mosaic virus* with 9465 nt (Mlotshwa *et al*., 2002), a prevalent strain of a potyvirus infecting maize in China with 9595 nt (Fan *et al*., 2003), *Watermelon mosaic virus* with 10021 nt (Desbiez and Lecoq, 2004), *Japanese yam mosaic virus* with 9757 nt (Fuji and Nakamae, 1999), *Cocksfoot streak virus* with 9663 nt (Götz and Maiss, 2002), *Wild potato mosaic virus* with 9853 nt (Spetz and Valkonen, 2003), *Lily mottle virus* with 9644 nt (Zheng *et al*., 2003), *Plum pox virus* with 9741 nt (Maiss *et al*., 1989) and *Pea seed-borne mosaic potyvirus* with 9895 nt (Olsen and Johansen, 2001). The data showed that BtMV-G genome contains a single large predicted open reading frame (ORF) starting at an AUG initiation codon located within the sequence 5'-…TAAAAUGGC…-3', which is reasonable agreement with a consensus sequence 5'-…AACA**AUG**GC…-3' proposed by Lütcke *et al*. (1987) for optimal initiation of translation in plants. In addition similarities exist also with the consensus sequence 5'-…A/GCCATGG…-3' proposed by Kozak (1986) for mammalian translation initiation. The ORF of BtMV-G is terminated by a TAA stop codon. The predicted translation product of the ORF contains 3085 amino acid residues with a calculated molecular weight of 349.69 kDa. Furthermore, the 5'- and 3'-nontranslated regions were determined with 166 and 171 nt, respectively.

The 5'-NTR sequence length of BtMV-G is typical for potyviruses. It contains a high content of adenine residues and 15 CAA nucleotide motifs described for the TMV 5'-leader sequence associated with translation enhancement (Gallie and Walbot, 1992), suggesting an influence on the efficiency of translation. In addition very few guanosine residues were found in the 5'-NTR sequence of BtMV-G. This seems to be a common feature of plant viral 5'-leader

sequences (Gallie et al., 1987). Two highly conserved regions called potyboxes ‘a’ (5’-ACAACAC-3’) and ‘b’ (5’-TCAAATT-3’) (Turpen, 1989) exist also within the 5’-NTR of BtMV-G. These nucleotide sequences were also found by Nemchinov *et al*. (2004) in the BtMV-Wa sequence. Furthermore, the 5’- NTR of BtMV-G showed a high sequence identity with the BtMV-Wa strain (83%), which is expected for isolates of one virus species.

According to the known data, the length of 3’-NTR of potyviruses varies from 161 nt for *Pea seedborne mosaic virus* to 596 nt for *Leek yellow stripe virus* and *Garlic virus* 2, while it is 545 nt for *Garlic potyvirus* and 475 nt for *Johnsongrass mosaic virus* (Berger *et al*., 1997). Thus the 3’-NTR of BtMV-G belongs to the shortest ones within potyviruses like BYMV (167 nt) and CYVV (176 nt) (Berger *et al*., 1997). The 3’-NTR sequence revealed high-level discrepancies in comparison with other potyviruses, but showed a high sequence identity (99%) with BtMV-Wa. The variability of the 3’-NTRs of potyviruses supported the recommendation to use the 3’-NTR to distinguish among different potyviruses. The almost complete identity of the 3’-NTRs of BtMV-G and BtMV-Wa clearly indicates two strains of the same virus (Berger *et al*., 1997).

Alignments of the polyprotein sequence of BtMV-G with those of other potyviruses revealed nine putative protease cleavage sites. These were identified by comparing the putative coding region of BtMV-G with the consensus protease recognition motifs of other potyviruses. The sequences of almost all putative cleavage sites are in agreement with the NIa consensus cleavage sites described by Riechmann *et al*. (1992). The cleavage site at the C-terminus of P1 is in a consensus reported for most of the potyviruses like *Cowpea aphid-borne mosaic virus* (Mlotshwa *et al*., 2002) and *Cocksfoot streak virus* (Götz and Maiss, 2002). The HC-Pro cleavage site at the C-terminal G/G dipeptide is part of the conserved motif of the potyvirus amino acid sequence YXVG/G (Fuji and Nakanae, 1999; Carrington and Herndon, 1992). Seven sites putatively cleaved by NIa at dipeptides Q/A, Q/S, Q/G, E/G, Q/G were also found in other published potyvirus sequences. However, also some differences exist. For example at the VPg/NIa and NIa/NIb junction of BtMV-G, respectively, glutamic acid/glycine (E/G) and glutamine/glycine (Q/G) are quite common among some potyviruses, while (Q/G) at the NIb/CP junction occurs with a lesser frequency. These differences in comparison to other potyviruses cleavage sites may cause a polyprotein cleavage of BTMV-G not to be as efficient as found for TEV (Carrington *et al*., 1993). Nevertheless, a new or unusual cleavage site was

not found. Most of them are typical for at least some potyviruses. The putative proteolytic cleavage sites of BtMV-G were most close to those of TEV, PeMoV and SMV.

Based on alignment of amino acid sequences for each protein and based on the likely location of cleavage sites, the polyprotein of BtMV-G is probably processed into ten smaller putative proteins by the three viral-encoded proteases (Dougherty and Semler, 1993; Carrington *et al*., 1990) P1 (Verchot *et al*, 1991), HC-Pro (Carrington *et al*., 1989a) and NIa (Dougherty and Parks, 1991; Dougherty *et al*., 1988). The proteolytic process of BtMV-G was supported by the presence of conserved motifs, such as the serine-type protease domain (H_{2092}-X_{43}-D_{2167}-X_{69}-C_{2198}-G_{2199}) in NIa and (I_{286}VRGR$_{290}$) in P1 (Dougherty and Semler, 1993; Fan *et al*., 2003). The ten mature proteins were found to occur in the same order in all identified potyviral genomes: P1, HC-Pro, P3, 6K1, CI, 6K2, NIa, VPg, NIb, and CP.

P1: As shown in table (4), the P1 amino acid sequence displayed high sequence variability in comparison with other proteins. No significant sequence identity was found with the P1 proteins of other potyvirus. Additionally, there are no functional conserved amino acid motifs in this protein sequence, except only some conserved amino acid residues, which are found in the C-terminal part and the proteolytic domain (IVRGR), necessary for polyprotein processing as found in other potyviruses (Fan *et al*., 2003). The amino acid sequence identity between P1 of BtMV-G and BtMV-Wa was found to be 92%, while it ranged from 3% to 30% with other potyviruses. This result confirmed that both isolates belong to one virus species. The phylogenetic tree based on the amino acid sequence of P1 indicated PeMoV as the closest related virus to BtMV-G.

HC-Pro: Six conserved motifs of other potyviruses were found in the HC-Pro. Three are involved in aphid transmission: KMAC, PTK and FRNK (Atreya *et al*., 1992; Fan *et al*, 2003; Hammond and Hammond, 2003). However, the amino acid sequence motif KMAC in BtMV-G is different in comparison to the amino acid sequence motif of other potyviruses, because of the substitution in two amino acid residues MA instead of IT, which were found in most of the HC-Pro of potyvirus amino acid sequences, or LT, VS or LS, respectively, which were found in a few potyvirus amino acid sequences (PVV, PeMoV, ZYMV). Nemchinov *et al*. (2004) also found the KMAC motif in the BtMV-Wa isolate. In spite of this replacement, the BtMV-G isolate is still aphid transmissible, but the substitution in this essential conserved motif for aphid transmission could be responsible for a decreased aphid transmissibility of BtMV-G of

only 30%. This is supported by a previous report, where the importance of amino acid residues K and I within the KITC motif in aphid transmission (Llave *et al*., 2002) was shown. Additionally, this result does not agree exactly with other suggestions that certain mutations affecting this amino acid position prevent aphid transmissibility of several potyviruses (Atreya *et al.*, 1992; Granier *et al.*, 1993; Huet *et al*., 1994; Canto *et al*., 1995).

The conserved amino acid motifs PTK and FRNK have been shown to be involved in aphid transmission in ZYMV (Granier *et al*., 1993; Huet *et al*., 1994; Peng *et al*., 1998) and were detected in most of the potyvirus amino acid sequences like CABMV (Mlotshwa *et al*., 2002), WPMV (Spetz and Valkonen, 2003) and the BtMV-Wa isolate (Nemchinov *et al*., 2004).

Huet et al. (1994) reported that a change from R to I within the conserved FRNK motif of ZYMV-HC reduced aphid transmission (20,6% for the mutated virus compared to 57,4% in the wild-type ZYMV). A mutation within the PTK resulted in total loss of HC-Pro activity. The SCC motif was found in BtMV-G HC-Pro instead of the conserved CCC box of other potyviruses HC-Pro. The motif is involved in viral long distance movement of TEV (Cronin *et al*., 1995) and the substitution of C to S could affect long distance movement of BtMV-G. The conserved motif LAIGN involved in cell-to-cell movement was present in HC-Pro of BtMV-G as well as in other potyvirus sequences (Spetz and Valkonen, 2003; Cronin *et al*., 1995). Another conserved motif CDNQLD of so far unknown function has a replacement of leucine by arginine in BtMV-G. Detailed analyses of the motif by mutation or deletion of amino acids are required to determine its function and responsibility.

In the N-terminal part of BtMV-G HC-Pro a zinc finger metal binding motif (C_{340}-X_8-C_{349}-X_{13}-C_{363}-X_4-C_{368}-X_2-C_{371}) was found, which was first identified in PVY HC-Pro (Robaglia *et al*., 1989). It was also reported for JYMV (Fuji and Nakamae, 1999), CSV (Götz and Maiss, 2002) and CABMV (Mlotshwa *et al*., 2002) with a V instead of C at position C_{363}.

The BtMV-G HC-Pro amino acid sequence showed the highest sequence identity with BtMV-Wa (94%) while it ranged from 57% with BCMV to 38% with CSV. The result indicates that BtMV-G HC-Pro has a relatively low sequence similarity with BtMV-Wa in comparison to other proteins of these both isolates.

CI: In the CI protein of BtMV-G important motifs for the viral replication process were found, like a RNA helicase motif (D_{1343}-E-X-H_{1346}), which was also identified in a prevalent strain of a potyvirus infecting maize in China (Fan *et al*., 2003). In addition a conserved amino acid

motif (LKVSATPP), which functions for RNA binding, is also present in the CI of BtMV-G and a motif with NTP binding activity (VGSKST) (Spetz and Valkonen, 2003; Lain *et al.*, 1990). These amino acid motifs are not perfectly conserved in all potyviruses amino acid sequences. Substitutions have been reported in the motif LKVSATPP which is not completely conserved in OYDV (LKTSATIP) and PVY (LKVSATPV).

NIa: The protease domain H-X_{43}-D-X_{69}-CG in the NIa protein (Fuji and Nakamae, 1999) was also found in BtMV-G.

NIb: The NIb protein sequence contains conserved amino acid motifs (Domier *et al*., 1987; Spetz and Valkonen, 2003) for a RNA-dependent RNA polymerase SG-X_3-T-X_3-NT-X_{30}-GDD and CDADGS (Glasa *et al*., 2003: Fan *et al.*, 2003).

CP: The essential amino acid motif for aphid transmissibility DAG was found in the BtMV-G CP as well as in other potyviruses (Dolja *et al*., 1994; Fan *et al.*, 2003). The position of this motif is not exactly fixed among potyviruses, but it is always located in the CP near the N-terminus (Glasa *et al*., 2003). However, it is not completely conserved in all potyviruses CP sequences, for example WPMV-CP has a substitution in this motif leading to DTG (Spetz and Valkonen, 2003), which could affect aphid transmissibility as reported by Flasinski and Cassidy (1998) for *Peanut stripe virus* (PStV). Another motif RX_{43}D was found in the BtMV CP sequence, which was reported as an important motif for viral long distance movement (Dolja *et al*., 1994; Fan *et al*., 2003).

4.6. Pairwise comparison, phylogenetic analyses and relationship of BtMV-G with other potyviruses.

Phylogenetic analyses clearly showed BtMV-G as a distinct member of the genus Potyvirus. The phylogenetic trees grouped BtMV-G, BtMV-Wa and PeMoV in one cluster located in the neighbourhood of the BCMV cluster, which might indicate that they have evolved from the same ancestor. This result is similar to the result of a phylogenetic analysis of the 3'-terminal region of BtMV (Glasa *et al*., 2003).

According to the obtained sequence, the isolate BtMV-G has a significant relationship with the American isolate BtMV-Wa with the highest nucleotide and amino acid sequence identity of 89% and 95%, respectively. The isolates shared the same amino acid sequence of 6K1 (100%) and the sequence identity with other proteins ranged from 98% for CI and CP to 88% for P3,

respectively. According to Shukla and Ward (1988), a sequence identity of 38% to 71% between two viruses indicates two distinct species, while 90% to 100% sequence identity indicates two strains of one virus species. Therefore BtMV-G and BtMV-Wa are two strains of the species BtMV.

BtMV-G shared the highest amino acid sequence identity (55%) with PeMoV, and it is closely related to BCMNV, SMV, ZYMV (50%) and CABMV (49%), which were called *Bean common mosaic virus* subgroup (Berger *et al*., 1997; Mink *et al.*, 1994). This result is in agreement with findings reported by Glasa *et al.* (2003) and Nemchinov *et al*. (2004) for other BtMV isolates.

The amino acid sequence identities for the different genomic regions of BtMV-G compared with other potyviruses showed that P1 is the least similar region. The amino acid sequence identity of P1 does not exceed 24% for any potyvirus comparison, while 6K1, 6K2, NIa, NIb and CP are more similar to other potyviruses. It is suggested to conduct more studies with P1 to elucidate the functions and influences of amino acid changes within this protein.

Taken together, the nucleotide and deduced amino acid sequences of BtMV-G, the high similarity of the proteolytic cleavage sites and the highest phylogenetic relatedness shared by BtMV-G, PeMoV and the BCMV subgroup, which was previously also reported according to CP amino acid sequence comparisons, confirmed the fact that BtMV-G belongs on the molecular level clearly to the genus *Potyvirus*.

4.7. Construction of the BtMV-G full- length cDNA clone.

One important part of this study was the construction of an infectious BtMV-G full-length clone. So far there has been no report of an infectious cDNA clone of BtMV. However, infectious RNA transcripts or cDNA clones of some potyviruses were described during the last years including several potyviruses such as ZYMV (Gal-On *et al.*, 1991), Taiwan strain of ZYMV (Lin *et al*., 2002), PPV (Maiss *et al*., 1992; Riechmann *et al*., 1990; Lopez-Moya and Garcia, 2000), PVA (Puurand *et al.*, 1996), TVMV (Nicolas *et al*., 1996), PSbMV (Johansen *et al*., 1996; Olsen and Johansen, 2001), PVY-N605 and a cell-free infectious PVY clone (Jakab *et al*., 1997; Fakhfakh *et al*., 1996) and TuMV (Sanchez *et al.*, 1998).

An efficient strategy was designed in this study to construct the full-length cDNA clone. Four cDNA fragments were amplified separately by RT-PCR using gene specific primers

containing suitable restriction sites. The overlapping cDNA fragments were cloned between the transcription initiation site of the 35S promoter and terminator sequences of *Cauliflower mosaic virus* (CaMV) (Mori *et al*., 1991; Maiss *et al*., 1992; Gal-On et al., 1995; Fakhfakh *et al*., 1996; Johansen *et al*., 1996; Jakab *et al*., 1997; Yang *et al*., 1998). The cDNA clones were joined by ligating fragments from overlapping clones sharing common restriction endonuclease sites. One of the primers, which were used to amplify a 5'-terminal cDNA fragment, was designed to create a *Dra* I site (TTTAAA) at the exact 5'-end of the BtMV-G cDNA (Maiss *et al*., 1992), to obtain the insert cDNA without extra nucleotides between the transcription initiation site of the 35S promoter and the first nucleotide of BtMV-G genome. The presence of more than five non-viral nucleotides at the 5'-end may interfere with infectivity of the transcript (Janda *et al*., 1987; Boyer *et al*., 1993). However, several studies showed that extra non-viral nucleotides had no effect on infectivity (Holy and Aboul-Haidar, 1993; Viry *et al*., 1993). Additional cloning steps were conducted to assemble the cDNA fragments as an entire full-length clone. BtMV-G full-length clones were maintained in *E. coli.* However, only one clone turned out to deliver infectious virus in *N. benthamiana*. This could be due to expression of putative "toxic sequences" of BtMV-G in *E. coli* or because of amino acid changes in the full-length clone. Also Chiang and Yeh (1997) reported infectious PRSV clones, without instability or toxicity problems in *E. coli*. However, cloning difficulties were faced during construction of several infectious potyvirus clones due to instability or toxicity in bacteria (Maiss *et al.*, 1992; Jakab *et al*., 1997). To avoid these problems alternative approaches were conducted, such as insertion of intron sequences into full-length clones to prevent production of putative toxic proteins from PSbMV (Johansen, 1996), PPV (Lopez-Moya *et al.*, 2000) and PSbMV-L1 (Olsen and Johansen, 2001). An alternative method was also used. In this cell-free method Fakhfakh *et al*. (1996) firstly cloned separate fragments of PVY. The complete cDNA of PVY was obtained by ligation of these fragments immediately before inoculation. The "toxic sequences" of the potyvirus genomes are not exactly determined, and there are some suggestions that the toxicity of the potyvirus genome in bacteria resulted from the simultaneous presence of cryptic promoter elements, translational initiation sites and the expression of cytotoxic gene products (Fakhfakh *et al*., 1996). Jakab *et al*. (1997) suggested that the toxicity is probably associated with the CI gene, which codes for a RNA helicase. It is suggested that the toxicity could be more understood and solved, if the

toxicity genes or products could be identified. However, the positions of sequences leading to toxicity could be different in various potyvirus species, strains and isolates.

The nucleotide sequence of the full length clone was determined and compared with the originally determined sequence. 27 nucleotide and 19 amino acid changes were found in comparison to the firstly determined sequence. The nucleotide changes could be due to errors at various stages in the cloning and sequencing process. For example, incorporation of false nucleotides can occur with Reverse Transcription, in the PCR or sequencing reaction, respectively. These changes may lead to new characters of gene expression, pathogenesis and replication. It is suggested that mutations of these amino acids are conducted to determine the role and influence in virus replication and symptomatology.

Infectivity rates of infectious full-length clones depend on the inoculation method and the cDNA concentration used for inoculation. To transcribe infectious RNA the cDNA clone must enter the cell nucleus (Lopez-Moya and Garcia, 2000). An efficient method is required to deliver the cDNA especially for *in vivo* transcribed clones. Particle bombardment was reported as an effective method to introduce DNA directly into the plant cell nucleus (Gal-On *et al*., 1995), Lopez–Moya and Garcia (2000) and Fakhfakh *et al*. (1996). The authors reported that biolistic inoculation was more efficient compared to the mechanical inoculation of plants and increasing the amount of bombarded DNA increased the efficiency of infection. In this study the biolistic method was used to bombard the BtMV-G cDNA clones into the plant cells of *N. benthamiana*. Three weeks post-inoculation symptoms of BtMV-G, delivered from the full-length clone, appeared as a systemic mosaic. Symptoms were delayed approximately two weeks but were indistinguishable from symptoms caused by the wild-type virus. Delays in symptom development using infectious clones were reported in some previous studies for PSbMV and ZYMV.

The BtMV-G derived from the full-length clone was stable and mechanical inoculation was highly efficient resulting in 100% infection of *N. benthamiana* after one week with similar symptoms of the wild-type virus. However, BtMV-G derived from the full-length clone showed reduced infectivity and different symptoms in other plant species as presented in table (7). In addition BtMV-G from the infectious clone was not able to infect some susceptible cultivars to the wild-type virus like *Spinacia oleracea* and *Beta vulgaris* (8T0015). Completely different symptoms were observed on *Atriplex hortensis* leaves, where only small

yellow blots on leaves appeared compared to a severe infection of the wild-type virus. These differences between the wild-type virus and the virus from the full-length clone in pathogenicity and symptom severity could be due to amino acid changes of proteins, which are involved in the symptom expression and interaction with the resistance genes of the host plant. Saenz *et al.* (2002) suggested that HC-Pro might be a factor for controlling the host range of PPV. In case of BtMV-G amino acid changes in HC-Pro, P1 and P3 could influence the symptom severity and the virulence of the virus as reported in some previous studies. Jenner *et al.* (2003) suggested that P3 protein of TuMV plays an important role as a symptom and virulence determinant. Suehiro *et al.* (2004) indicated that P3 coding region is responsible for systemic infection in TuMV and suggested it as an important factor in the infection cycle of TuMV and determining the host range of TuMV. Stenger and French (2004) suggested that HC-Pro affects the virulence of *Wheat streak mosaic virus (*WSMV) and is required for virus-host interaction and systemic infection. In addition, Rodriguez-Cerezo *et al.* (1991) indicated that a noncoding region of the genome might involve in the direct effect in the induction of the disease symptoms by an RNA virus.

There is only one amino acid change in the CP of the full-length clone compared with the firstly determined sequence. This may not influence the symptom phenotype as reported before. Mutation in the coat protein gene of a full length clone of PStV produced symptoms indistinguishable from the native virus (Flasinski *et al.*, 1996). However, it has still to be determined, which protein(s) of the BtMV-G could be involved in symptom phenotypes and virulence. To address this question exchanges and mutations in the 5'-part of the BtMV-G region could be made. In any event, the infectious clone confirmed the genome sequence and indicated that in the determination of the genome of BtMV-Wa an adenosine residue at the extreme 5'-end is missing. The construction of the BtMV-G full-length clone is an important step forward to study functions of BtMV-G genes and the role of specific sequences for virus encapsidation, pathogenicity, transmission, replication and translation. It could provide a powerful tool for elucidating the relationship between gene function and the biological properties of the virus.

4.8. Conclusion

In this study the complete nucleotide sequence of the German isolate of *Beet mosaic virus* (BtMV-G) was determined (9592 nt) and compared with other published potyvirus sequences. The result indicated that BtMV-G is a distinct member of the genus *Potyvirus* in the family *Potyviridae*. Nine putative proteolytic cleavage sites were determined resulting in ten mature proteins. Most of the cleavage sites are similar to those of other members of genus *Potyvirus*. The sequence analysis revealed important specific conserved amino acid motifs described for all potyviruses. Phylogenetic analysis based on alignment of the polyprotein demonstrated that BtMV-G, BtMV-Wa and PeMoV belong in one cluster. BtMV-G may have evolved from the same ancestor as other viruses that infect plants of the family *Leguminosae*. Nucleotide sequence and amino acid sequence comparison indicated that BtMV-G and BtMV-Wa are two isolates of the same virus.

A full-length clone of BtMV-G was successfully constructed revealing infectious *in vivo* transcripts. The symptom development of BtMV-G transcript infected plants was slower than with wild-type virus. However, the symptoms were indistinguishable from those of the wild-type virus. Host range experiments showed some virulence differences between the virus delivered by the full-length clone and wild type virus. The infectious BtMV-G full-length clone can serve as a tool to study the biological properties of BtMV and should contribute towards a more understanding of the molecular biology of the genus *Potyvirus*.

5. References

Aleman-Verdaguer, M.E., Goudou-Urbino, C., Dubern, J., Beachy, R.N., Fauquet, C. 1997. Analysis of the sequence diversity of the P1, HC, P3, NIb and CP genomic regions of several Yam mosaic potyvirus isolates: implications for the intraspecies molecular diversity of potyviruses. *J. Gen. Virol*. 78: 1253-1264.

Allison, R., Johnston, R.E., Dougherty, W.G. 1986. The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polyprotein. *Virology* 154:9-20.

Andrews, J.H., Shalla, T.A. 1974. The origin, development and conformation of amorphous inclusion body components in tobacco etch virus-infected cells. *Phytopathology* 64: 1234-1243.

Arbatova, J., Lehto, K., Pehu, E., Pehu,T. 1998. Localization of the P1 protein of Potato Y potyvirus in association with cytoplasmic inclusion bodies and in the cytoplasm of infected cells. *J. Gen. Virol*. 79: 2319-2323.

Atabekov, J.G., Rodionova, N.P., Karpova, O.V., Kozlovsky, S.V., Poljakov, V.Yu. 2000. The movement protein-triggered in situ conversion of potato virus X virion RNA from nontranslatable into a translation form. *Virology* 271: 259-263.

Atreya, C.D., Raccah, B., Pirone, T.P. 1990. A point mutation in the coat protein abolishes aphid transmissibility of a potyviruses. *Virology*178: 161-165.

Atreya, C.D., Atreya, P.L., Thornbury, D.W., Pirone, T.P. 1992. Site-directed mutations in the potyvirus HC-Pro gene affect helper component activity, virus accumulation, and symptom expression in infected tobacco plants. *Virology* 191: 106-111.

Atreya, P.L., Lopez-Moya, J.J., Chu, M.H., Atreya, C.D., Pirone, T.P. 1995. Mutation analysis of the coat protein N-terminal amino acids involved in potyvirus transmission by aphids. . *J. Gen. Virol.* 76: 265-270.

Attoui, H., Billoir, F., Cantaloube, J.F., Biagini, P., de Micco, P., de Lamballerie, X. 2000a. Strategies for the sequence determination of viral dsRNA genomes. *J. Virol. Methods* 89: 147-58.

Attoui, H., Billoir, F., Cantaloube, J.F., Biagini, P., de Micco, P., de Lamballerie, X. 2000b. Complete sequence determination and genetic analysis of Banna virus and Kadipiro virus: proposal for assignment to a new genus (*Seadornavirus*) within the family *Reoviridae*. *J. Gen. Virol*. 81: 1507-1515.

Barnett, O.W. 1991. *Potyviridae*, a proposed family of plant viruses. *Arch. Virol*. 18: 139-141.

Barnett, O.W. 1992. A summary of potyvirus taxonomy and definitions, In: Barnet O.W. (ed) Potyvirus Taxonomy. *Arch. Virol.* 5: 435-444.

Bennet, C.W, 1944. Latent virus of dodder and its effect on sugar beet plants. *Phytopathology* 34: 77-91.

Bennet, C.W, 1949. Comparison of methods for artificially inoculating sugar-cane seedlings with the mosaic virus. *Phytopathology* 39: 669

Bercks, R. 1960. Serological relationships between beet mosaic virus, potato virus Y and bean yellow mosaic virus. *Virology* 12: 311-313.

Berg, J.M. 1990. Zinc fingers and other metal binding domains. *J. Bio. Chem*. 265: 6513-6516.

Berger, C. 2000. Molekulare Charakterisierung der Proteine NIb, P3 und 6K1 des *Beet mosaic virus*. Diplomarbeit. Institute of Plant Diseases and Plant Protection, Hannover University, Germany.

Berger, P.H., Wyatt, S.D., Shiel, P.J., Silbernagel, M.J., Druffel, K., Mink., G.I. 1997. Phylogenetic analysis of the potyviridae with emphasis on legume-infecting potyviruses. *Arch. Virol*. 142: 1979-1999.

Birnboim, H.C., Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res*. 7: 1513-1523.

Blanc, S., Lopez-Moya, J.J., Wang, R., Garcia-Lampasona, S., Thornbury, D.W., Pirone, T.P. 1997. Specific interaction between coat protein and helper component correlates with aphid transmission of a potyvirus. *Virology* 231: 141-147.

Blanc, S., Ammar, E.D., Garcia-Lampasona, S., Dolja, V.V., Llave, C., Baker, J., Pirone, T.P., 1998. Mutations in the potyvirus helper component protein: effects on interaction with virions and aphid stylets. *J. Gen.Virol.* 79: 3119-3122.

Blanko-Urgoiti, B., Sanchez, F., Dopazo, J., Ponze, F. 1996. A strain type clustering of potato virus Y based on genetic distance between isolates calculated by RFLP analysis of the amplified coat protein gene. *Arch. Virol*. 141: 2425-2442.

Borgstrom B., Johansen, I.E. 2001. Mutations in pea seedborne mosaic virus genome-linked protein VPg after pathotype-specific virulence in *Pisum sativum*. *Mol. Plant Microbe Interact*. 14: 707-14.

Bos, L. 1969. Inclusion bodies of bean yellow mosaic virus, some less known closely related viruses and beet mosaic virus. *Neth. J. Plant Pathol*. 75: 137-143.

Bos, L. 1970. The identification of three new viruses isolated from wisteria and pisum in the Netherlands, and the problem of variation within the potato virus x *group. Neth. J. Plant Pathol.* 76: 8-46.

Boyer, J.C., Drugeon,G., Seron, K., Morch-Devignes, M.D., Agnes, F., Haenni, A.L. 1993. *In vitro* transcripts of Turnip yellow mosaic virus encompassing a long 3' extension or produced from a full-length cDNA clone harbouring a 2kb-long PCR amplified segment are infectious. *Res. Virol*. 144. 393-348.

Brakke, M, K., Van Pelt, N. 1970. Properties of infectious ribonucleic acid from wheat streak mosaic virus. *Virology* 42: 699-706.

Brantley, J.D., Hunt, A. G. 1993. The N-terminal protein of the polyprotein encoded by the potyvirus tobacco vein mottle virus is an RNA-binding protein. *J. Gen. Virol*. 74: 1157-1162.

Canto, T., Lopez-Moya, J.J., Serra-Yoldi, M.T., Diaz-Ruiz, J.R., Lopez-Abella, D. 1995. Different helper component mutations associated with lack of aphid transmissibility in two isolates of potato virus Y. *Phytopathology* 85: 1519-1524.

Carrington, J.C., Dougherty, W.G. 1987a. Small nuclear inclusion protein encoded by a plant potyvirus genome is a protease. *J. Virol*. 61: 2540-2548.

Carrington, J.C., Dougherty, W.G. 1987b. Processing of the tobacco etch virus 49K protease requires auto-proteolysis. *Virology* 160: 355-362

Carrington, J.C., Dougherty, W.G. 1988. A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proc. Natl. Acad. Sci. U S A.* 85: 3391-5.

Carrington, J.C., Freed, D.D. 1990. Cap-independent enhancement of translation by a plant potyvirus 5'-nontranslated region. *J. Virol*. 64: 1590-1597.

Carrington, J.C., Herdnon, K.L. 1992. Characterization of the potyviral HC-Pro autoproteolytic cleavage site. *Virology* 187: 308-315.

Carrington, J.C., Cary, S.M., Dougherty, W.G. 1988. Mutational analysis of tobacco etch virus polyprotein processing: cis and trans proteolytic activities of polyproteins containing the 49-kilodalton proteinase. *J. Virol*. 62: 2313-2320.

Carrington, J.C., Cary, S.M., Parks, T.D., Dougherty, W.G. 1989a. A second proteinase encoded by a plant potyvirus genome. *EMBO. J*. 8: 365-370.

Carrington, J.C., Freed, D.D., Sanders T.C. 1989b. Autocatalytic processing of the potyvirus helper component proteinase in *Escherichia coli* and *in vitro*. *J. Virol.* 63: 4459-4463.

Carrington, J.C., Freed, D.D., Oh, C.H. 1990. Expression of potyviral polyproteins in transgenic plants reveals three proteolytic activities required for complete processing. *EMBO J*. 9: 1347-1353.

Carrington, J.C., Haldeman, R., Dolja, V.V., Restrepo-Hartwig, M.A. 1993. Internal cleavage and trans-proteolytic activities of the VPg-proteinase (NIa) of tobacco etch potyvirus in vivo. *J. Virol.* 67: 6995-700.

Chang, C.A., Hiebert, T.E., Purcifull, D.E. 1988. Analysis of in vitro translation of bean yellow mosaic virus RNA: inhibition of proteolytic processing by an antiserum to the 49K nuclear inclusion protein. *J. Gen. Virol.* 69: 1117-1122.

Chiang, C.H., Yeh, S.D. 1997. Infectivity assays of *in vitro* and *in vivo* transcripts of Papaya ringspot potyvirus. *Botanical Bulletin of Acadimia Sinica* 38: 153-163.

Chu, M., Lopez-Moya, J.J., Llave-Correas, C., Pirone, T.P. 1997. Two separate regions in the genome of tobacco etch virus contain the determinants of the wilting response of tabasco pepper. *Mol. Plant-Microbe Interact*. 10: 472-480.

Clark, M.F., Adam, N. 1977. Characteristics of the microtiter plate method of enzyme-linked immunoassay (ELISA) for the detection of plant viruses. *J. Gen. Virol*. 34: 475-483.

Cronin, S., Verchot, J., Haldeman-Cahill, R., Schaad, M.C., Carrington, J.C. 1995. Long-distance movement factor: a transport function of the potyvirus helper component proteinase. *Plant Cell* 7: 549-559.

Daros, J.A., Schaad, M.C., Carrington, J.C. 1999. Functional analysis of the interaction between VPg-Proteinase (NIa) and RNA polymerase (NIb) of tobacco etch potyvirus, using conditional and suppressor mutants. *J. Virol*. 73: 8732-8740.

David A.B., Pritam, D., Mary, E.B. 1991. A Temporal study of the expression of the capsid, cytoplasmic inclusion and nuclear inclusion proteins of tobacco etch potyvirus in infected plants. *J. Gen. Virol*. 72: 487-492.

Desbiez, C., Lecoq, H. 2004. The nucleotide sequence of Watermelon mosaic virus (WMV, Potyvirus) reveals interspecific recombination between two related potyviruses in the 5' part of the genome. *Arch. Virol.* 149: 1619-32.

Dolja, V.V., Boyko, V.P., Agranovsky, A.A., Koonin, E.V. 1991. Phylogeny of capsid proteins of rod-shaped and filamentous RNA plant viruses: Two families with distinct patterns of sequence and probably structure conservation. *Virology* 184: 79-86.

Dolja, V.V., Herndon, K.L., Pirone, T.P., Carrington, J.C. 1993. Spontaneous mutagenesis of a plant potyvirus genome after insertion of a foreign gene. *J. Virol*. 67: 5968-5975.

Dolja, V.V., Haldeman, R., Robertson NL, Dougherty, W.G., Carrington, J.C. 1994. Distinct functions of capsid protein in assembly and movement of tobacco etch potyvirus in plants. *EMBO J*. 13: 1482-1491.

Domier, L.L., Franklin, K.M., Shahabuddin, M., Hellmann, G.M., Overmeyer, J.H. 1986. The nucleotide sequence of tobacco vein mottling virus RNA. *Nucl. Acids Res*. 14: 5417-5430.

Domier, L.L., Shaw, J.G., Rhoads, R.E. 1987. Potyviral proteins share amino acid sequence homology with picorna- and caulimoviral proteins. *Virology*. 158: 20-27.

Dougherty, W.G., Hiebert, E. 1980. Translation of potyviral RNA in a rabbit reticulocyte lysate: identification of nuclear inclusion proteins as products of tobacco etch virus RNA translation and cylindrical inclusion protein as a product of the potyvirus genome. *Virology* 104: 174-182.

Dougherty, W.G., Carrington, J.C. 1988. Expression and function of potyviral gene products. *Ann. Rev. Phytopathology* 26: 123-143.

Dougherty, W.G., Parks, T.D. 1991. Post-translational processing of the tobacco etch virus 49-kDa small nuclear inclusion polyprotein: identification of an internal cleavage site and delimitation of VPg and proteinase domains. *Virology*. 183: 449-456.

Dougherty, W.G., Semler, B.L. 1993. Expression of virus-encoded proteinase: functional and structural similarities with cellular enzyme. *Microbiol*. *Rev*. 57: 781-822.

Dougherty, W.G., Carrington, J.C., Cary, S.M., Parks, T.D. 1988. Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *EMBO J*. 7: 1281-1278.

Dusi, A.N. 1999. Beet Mosaic virus: epidemiology and damage. PhD thesis Wageningen, Netherlands.

Dusi, A.N., Peters, D. 1999. Beet Mosaic virus: Its vector and host relationships. *Phytopathology*. 147: 293-298.

Dusi, A.N., Peters, D., Vanderwerf, W. 2000. Measuring and modeling the effects of inoculation date and aphid flights on the secondary spread of Beet Mosaic virus in sugar beet. *Ann. appl. Biol*. 136: 131-146.

Eckel, R.V.W., Lampert, E.B. 1993. Spatial and temporal analysis of tobacco etch virus distribution and its relationship to aphid (Homoptera: aphididae) vectors in flue-cured tobacco. *J. Econ. Entomol*. 86: 1534-1545.

Edwardson, J.R. 1974. Some properties of the potato virus Y-group. *Fla. Agric. Exp. Stn. Monogr. Ser. V. 255pp.*

Edwardson, J.R., Christie, R.G., Ko, N.J. 1984. Potyvirus cylindrical inclusions subdivision IV. *Phytopathology* 74: 1111-114.

Fakhfakh, H., Vilain, F., Makni, M., Robaglia, C. 1996. Cell-free cloning and biolistic inoculation of an infectious cDNA of potato virus Y. *J. Gen. Virol*. 77: 519-523.

Fan, Z. F., Chen, H.Y., Liang, X.M., Li, H.F. 2003. Complete sequence of the genomic RNA of the prevalent strain of potyvirus infecting Maize in China. *Arch. Virol*. 148: 773-782.

Fedorkin, O.N., Solovyev, A.G., Yelina, N.E., Zamyatnin, A.A., Zinovkin, R.A., Makinen, K., Schiemann, J., Morozov, S.Yu. 2001. Cell-to-cell movement of potato virus X involves distinct functions of the coat protein. *J. Gen. Virol.* 82: 449-458.

Fellers, J., Wan, J., Hong, Y., Collins, G.B., Hunt, A.G. 1998. In vitro interactions between a potyvirus-encoded, genome-linked protein and RNA-dependent RNA polymerase. *J. Gen. Virol.* 79: 2043-2049.

Fernandez, A., Garcia, J.A. 1996. The RNA helicase CI from plum pox potyvirus has two regions involved in binding to RNA. *FEBS Lett.* 388: 206-210.

Fernandez, A., Lain, S., Garcia, J.A. 1995. RNA helicase activity of the plum pox potyvirus CI protein expressed in *Escherichia coli*. Mapping of an RNA binding domain. *Nucl. Acids Res.* 23:1327-1332.

Fernandez, A., Guo, H.S., Saenz, P., Simon-Buela, L., Gomez de Cedron, M., Garcia, J.A. 1997. The motif V of plum pox potyvirus CI RNA helicase is involved in NTP hydrolysis and is essential for virus RNA replication. *Nucl. Acids Res*. 25: 4474-4480.

Flasinski, S., Cassidy, B.G. 1998. Potyvirus aphid transmission requires helper component and homologous coat protein for maximal efficiency. *Arch. Virol*. 143: 2159-2172.

Flasinski, S., Gunasinghe, U.B., Gonzales, R.A., Cassidy, B.G. 1996. The cDNA sequence and infectious transcripts of peanut stripe virus. *Gene*. 171: 299-300.

Frenkel, M.J., Ward, C.W., Shukla, D.D. 1989. The use of 3' noncoding nucleotide sequences in the taxonomy of potyviruses: application to watermelon mosaic virus 2 and soybean mosaic virus-N. *J. Gen. Virol.* 70: 2775-2783.

Froussard, P. 1992. A random-PCR method (rPCR) to construct whole cDNA library from low amount of RNA. *Nucl. Acids Res*. 20: 2900.

Fuji, S., Nakamae, H. 1999. Complete nucleotide sequence of the genomic RNA of a Japanese yam mosaic virus, a new potyvirus in Japan. *Arch. Virol*. 144: 231-240.

Fuji, S., Inoue, M., Yamamoto, H. Furuya, H., Naito, H., Matsumoto. 2003. Nucleotide sequences of the coat protein gene of potyviruses infecting *Ornithogalum thyrsoides*. *Arch. Virol*. 148: 613-621.

Fujisawa, I., Matsui, C., Yamaguchi, A. 1967. Inclusion Bodies Associated with Sugar Beet Mosaic. *Phytopathology* 57: 210-213.

Fujisawa, J.R., Tsuchizaki, T., Lizuka, N. 1983. Purification and serology of beet mosaic virus. *Annals of Phytopathology Society of Japan*. 49: 22-31.

Gallie, D.R., Walbot, V. 1992. Identification of the motifs within the tobacco mosaic virus 5'-leader responsible for enhancing translation. *Nucl. Acids Res*. 11: 4631-38.

Gallie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C., Wilson, T.M.A. 1987. A comparison of eukaryotic viral 5'-leader sequences as enhancers of mRNA expression *in vivo*. *Nucl. Acids Res*. 15: 8693-8711.

Gal-On, A., Antignus, Y., Rosner, A., Raccah, B. 1991. Infectious *in vitro* RNA transcripts derived from cloned cDNA of the cucurbit potyviruses, Zucchini yellow mosaic virus. *J. Gen. Virol*. 72: 2639-2643.

Gal-On, A., Meiri, E., Huet, H., Hau, J.W., Raccah, B., Gaba, V. 1995. Particle bombardment drastically increases the infectivity of cloned DNA of Zucchini yellow mosaic potyvirus. *J. Gen. Virol*. 76: 3223-3227.

Garcia, J.A., Lain S. 1991. Proteolytic activity of plum pox virus - tobacco etch virus chimeric NIa proteases. *FEBS Lett*. 281: 67-72.

Garcia, J.A., Riechmann, J.L., Lain, S. 1989. Proteolytic activity of the plum pox potyvirus NIa-like protein in *Escherichia coli*. *Virology*.170: 362-369.

Garcia, J.A., Lain, S., Cervera, M.T., Riechmann, J.L. 1990. Mutational analysis of plum pox potyvirus polyprotein processing by the NIa protease in *Escherichia coli*. *J. Gen. Virol*. 71: 2773-2779.

Ghabrial, S.A., Smith, H.A., Parks, T.D., Dougherty, W.G. 1990. Molecular genetic analyses of the soybean mosaic virus NIa proteinase. *J. Gen. Virol*. 71: 1921-1927.

Gibbs, J., Harrison, B.D. 1964. Nematode-transmitted viruses in sugar beet in East Anglia. *Plant Pathol.* 13: 144-150.

Glasa, M., Kudela, O., Subr, Z. 2003. Molecular analysis of the 3' terminal region of the genome of *Beet mosaic virus* and its relation with other potyviruses. *Arch.Virol*. 148: 1863-1871.

Goldbach, R. 1986. Molecular evolution of plant RNA viruses. *Ann. Rev. Phytopathology* 24: 289-310.

Goldbach, R. 1992. The recombinative nature of potyviruses: implications for setting up true phylogenetic taxonomy. *Arch. Virol*. 5: 299-304.

Gorbalenya, A.E., Donchenko, A.P., Kooning, E.V. 1989. Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases: a distinct protein superfamily with a common structural fold. *FEBS Lett*. 243: 103-114.

Götz, R., Maiss, E. 2002. The complete sequence of the genome of Cocksfoot streak virus (CSV), a grass infecting Potyvirus. *Arch. Virol.* 147:1573-1583.

Granier, F., Durant-Tardif, M., Casse-Delbart, E., Lecoq, H., Robaglia, C. 1993. Mutations in Zucchini yellow mosaic virus helper component protein associated with loss of aphid transmissibility. *J. Gen. Virol*. 74: 2737-2742.

Grüntzing, M., Fuchs, E. 1979. Untersuchungen zur serologischen Verwandschaft von beet mosaic virus, potato virus Y, bean yellow mosaic virus und plum pox virus. *Archiv für Phytopath. und Pflanzenschutz* 15: 135-139.

Guo, H.S., Lopes-Moya, J.J., Garcia, J.A. 1998. Susceptibility to recombination rearrangements of a chimeric plum pox potyvirus genome after insertion of a foreign gene. *Virus Res*. 57: 183-195.

Habera, L.F., Berger, P.H., Reddick, B.B. 1994. Molecular evidence from 3'-terminus sequence analysis that tobacco vein-banding mosaic virus is a distinct member of the potyvirus group. *Arch. Virol*. 138: 27-38.

Hammond, J., Lawson, R.H. 1988. An improved purification procedure for preparing potyviruses and cytoplasmic inclusions from the same tissue. *J. Virol. Methods*. 20: 23-27.

Hammond, J., Hammond, R.W. 2003. The complete nucleotide sequence of isolate BYMV-GDD of *Bean yellow mosaic virus*, and comparison to other potyviruses. *Arch. Virol*. 148: 2461-2470.

Hampton, R., Ball, E., De Boer, S. (*Editors*) 1990. Serological methods for detection and identification of viral and bacterial plant pathogens: a laboratory manual. *American Phytopathological Society. St. Paul, Minn*.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166: 557.

Hari, V. 1981. The RNA of tobacco etch virus: further characterization and detection of protein linked to RNA. *Virology* 112: 391-399.

Hari, V., Siegel, A., Rozek, D., Temberlake, W.E. 1979. The RNA of tobacco etch virus contains poly (A). *Virology* 92: 568-571.

Heathcote, G. D. 1973. Beet Mosaic - a Declining Disease in England. *Pl. Path*. 22: 42-45.

Hellmann, G.M., Shaw, J.G., Rhoads, R.E. 1988. In vitro analysis of tobacco mottling virus NIa Cistron: evidence for a virus encoded protease. *Virology* 163: 554-562.

Hiebert, E., McDonald, J.G. 1973. Characterization of some proteins associated with viruses in the potato virus Y group. *Virology* 56: 349-361.

Higgins, C.M., Cassidy, B.G, Teycheney, P.Y, Wongkaew, S. Dietzgen, R.G. 1998. Sequences of the coat protein gene of five peanut stripe virus (PStV) strains from Thailand and their evolutionary relationship with other bean common mosaic virus sequences. *Arch. Virol*. 143:1655-1667.

Hoefert, L.L. 1969. Proteinaceous and Virus-like inclusions in Cell infected with Beet mosaic Virus. *Virology* 37: 498.

Hollings, M., Brunt, A.A. 1981. Potyviruses. In: Handbook of plant virus infections and comparative diagnosis. *Ed. E. Kurstak, pp*. 732-777. *New York: Elsevier/North Holland Biomedical*. 943 *pp*.

Hollings, M., Brunt, A.A. 1981a. Potyvirus group. *CMI/AAB Description of plant viruses* 245, Kew, England. 7pp.

Holy, S., Aboul-Haidar, M.G. 1993. Production of infectious *in vitro* transcripts from a full-length clover yellow mosaic virus cDNA clone. *J. Gen. Virol.* 74: 718-784.

Hong, Y., Hunt, A.G. 1996. RNA polymerase activity catalyzed by a potyviruses-encoded RNA-dependent RNA polymerase. *Virology* 226: 146-151.

Hong, Y., Levay, K., Murphy, J.F., Klein, P.G., Shaw, J.G., Hunt, A.G. 1995. A potyvirus polymerase interacts with the viral coat protein and VPg in yeast cells. *Virology* 214: 159-166.

Huet, H., Gal-On, A., Meir, E., Lecoq, H., Raccah, B. 1994. Mutations in the Helper component protease gene of Zucchini yellow mosaic virus affect its ability to mediate aphid transmissibility. *J. Gen. Virol*. 75: 1407-1414.

Hull, R. 2002. Mattews 'Plant virology' (4th Ed), Academic press. London. pp. 33.

Jafri, A.M. 1972. Localization of virus and ultrastructural changes in sugar beet leaf cells associated with infection by a necrotic strain of beet mosaic virus. *Phytopathology* 62: 1103.

Jagadish, M.N., Huang, D., Ward, C.W. 1993. Site-directed mutagenesis of a potyvirus coat protein and its assembly in *Escherichia co*li. *J. Gen. Virol*. 74: 893-896.

Jagadish, M.N., Ward, C.W., Gough, K.H., Tulloch, P.A., Whittaker, L.A., Shukla, D.D. 1991. Expression of potyvirus coat protein in *Escherichia coli* and yeast and its assembly into virus-like particles. *J. Gen. Virol*. 72: 1543-1550.

Jakab, G., Droz, E., Brigneti, G., Baulcombe, D., Malnoe, P. 1997. Infectious *in vivo* and in vitro transcripts from a full-length cDNA clone of PVY-N605, a Swiss necrotic isolate of potato virus Y. *J. Gen. Virol.* 78: 3141-3145.

Janda, M., French, R., Ahlquist. P. 1987. High efficiency T7 polymerase synthesis of infectious RNA from cloned brome mosaic virus cDNA and effects of 5' extensions on transcript infectivity. *Virology* 158: 259-262.

Johansen, I.E., Keller, K.E., Dougherty, W.G., Hampton, R.O. 1996. Biological and molecular properties of a pathotype P-1 and pathotype P-4 isolate of pea seed- borne mosaic virus. *J. Gen. Virol*. 77: 1329-1333.

Kasschau, K.D., Carrington, J.C. 2001. Long-distance movement and replication maintenance function correlated with silencing suppression activity of potyviral HC-Pro. *Virology* 285: 71-81.

Kasschau, K.D., Cronin, S., Carrington, J.C. 1997. Genome amplification and long-distance movement functions associated with the central domain of tobacco etch potyvirus helper component-proteinase. *Virology* 228: 251-262.

Kennedy, J.S., Day, M.F., Eastop, V.F. 1962. A conspectus of aphids as vectors of plant viruses. Common W. Inst. Ent. London.

Klein, P.G., Klein, R.R., Rodriguez-Cerezo, E., Hunt, A.G., Shaw, J.G. 1994. Mutational analysis of the tobacco vein mottling virus genome. *Virology* 204: 759-769.

Knorr D.A., Dawson, W.O. 1988. A point mutation in the tobacco mosaic virus capsid protein gene induces hyposensitivity in *Nicotiana sylvestris*. *Proc. Natl. Acad. Sci. USA.* 85: 170-174.

Knuhtsen, H., Hiebert, E., Purciful, D.E. 1974. Partial purification and some properties of tobacco etch virus induced intranuclear inclusions. *Virology* 61: 200-209.

Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44: 283-92.

Kusov, Y.Y., Shatirishvili, G., Dzagurov, G., Gauss-Muller, V. 2001. A new G-tailing method for the determination of the poly(A) tail length applied to hepatitis A virus RNA. *Nucl. Acids Res.* 15: 57-7.

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

Lain, S., Riechmann, J.L., Garcia, J.A. 1989. The complete nucleotide sequence of plum pox potyvirus RNA. *Virus Res*. 13: 157-172

Lain, S., Riechmann, J.L. Garcia, J.A. 1990. RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucl. Acids Res*. 18: 7003-7007.

Lain, S., Martin, M.T., Riechmann, J.L., Garcia, J.A. 1991. Novel catalytic activity associated with positive-strand RNA virus infection: nucleic acid-stimulated ATPase activity of the plum pox potyvirus helicaselike protein. *J. Virol*. 65: 1-6.

Langenberg, W.G. 1986. Virus protein associated with cylindrical inclusions of two viruses that infect wheat. *J. Gen. Virol*. 67: 1161-1168.

Langenberg, W.G., L. Zhang. 1997. Immunocytology shows the presence of tobacco etch virus P3 protein in nuclear inclusions. *J. Struc. Biol.* 118: 243-247.

Lanneau, M., Candresse, T. 1993. Simplified Immunocapture-PCR protocol. Cost 88. Plum Pox Virus Workshop, Valencia, Spain.

Lawson, R.H., Hearon, S.S. 1971. Association of pinwheel inclusions with plasmodesmata. *Virology* 44: 454-456.

Lecoq, H., Pitrat, M. 1985. Specificity of the helper component–mediated aphid transmission of three potyviruses infecting muskmelon. *Phytopathology* 75: 890-893.

Lesemann, D.E. 1988. Cytopathology; the plant viruses. New York, Plenum Press. 4: 179-235.

Li, X.H., Valdez, P., Olvera, R.E., Carrington, J.C. 1997. Function of the tobacco etch virus RNA polymerase (NIb): subcellular transport and protein-protein interaction with VPg proteinase (NIa). *J. Virol.* 71: 1598-1607.

Lin, S.S., Roger, F.H., Yeh, S.D. 2002. Construction of *in vitro* and *in vivo* infectious transcripts of a Taiwan strain of *Zucchini yellow mosaic virus*. *Bot. Bull. Acad. Sin* 43: 261-269.

Lisa, V., Lecoq. H. 1984. Zucchini Yellow Mosaic Virus. Kew: *CMI- AAB, Description of plant viruses* No 282.4p.

Llave, C., Martinez, B., J.R. Diaz-Ruiz, D. Lopez-Abella. 2002. Amino acid substitutions within the Cys-Rich domain of the tobacco etch virus HC-Pro result in loss of transmissibility by aphids. *Arch. Virol*.147: 2365-2375.

Loebenstein, G., Raccah, B. 1980. Control of non-persistently transmitted aphid borne viruses. *Phytoparasitica* 8: 221-235.

Lopez-Moya, J.J., Pirone, T.P. 1998. Charge changes near the N-terminus of the coat protein of two potyviruses affect virus movement. *J. Gen. Virol*. 79: 161-165.

Lopez-Moya, J.J., Garcia, A.J. 2000. Construction of a stable and highly infectious intron-containing cDNA clone of *Plum pox potyvirus* and its use to infect plants by particle bombardment. *Virus Res*. 68: 99-107.

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

Maia, I.G., Haenni, A.L., Bernardi, F. 1996. Potyviral HC-Pro: A multifunctional protein. *J. Gen. Virol.* 77: 1335-1341.

Maiss, E., Timpe, U., Brisske-Rode, A., Lesemann, D.E., Casper, R. 1992. Infectious in vivo transcripts of a plum pox potyvirus full-length cDNA clone containing the cauliflower mosaic virus 35S RNA promoter. *J. Gen. Virol.* 73: 709-713.

Maiss, E., Timpe, U., Brisske, A., Jelkmann, W., Casper, R., Himmler, G., Mattanovich, D., Katinger, H.W.D. 1989. The complete nucleotide sequence of Plum Pox Virus RNA. *J. Gen. Virol*. 70: 513-524.

Martelli, G.P., Russo, M. 1969. Nuclear Changes in Mesophyll Cell of Gomphrena globosa L. Associated with Infection by Beet Mosaic Virus. *Virology* 38: 297-308.

Martin, M.T., Gelie, B. 1997. Non-structural plum pox potyviruses proteins detected by immunogold labelling. *Eur. J. Plant Pathol*.103: 427-431.

Masuta, C., Nishimura, M., Morishita, H., Hataya, T. 1999. A single amino acid change in viral genome-associated protein of potato virus Y correlates with resistance breaking in 'Virgin A mutant' tobacco. *Phytopathology* 89: 118-123.

Merits, A., Guo, D., Saarma, M. 1998. VPg, coat protein and five non-structural proteins of potato A potyvirus bind RNA in a sequence-unspecific manner. *J. Gen. Virol*. 79: 3123-3127.

Merits, A., Rajamaki, M.L., Lindholm, P., Roos, P.R., Kekarainen, T., Puustinen, P., Makelainen, K., Volkonen, J.P.T., Saarma, M. 2002. Proteolytic processing of potyviral proteins and polyprotein processing intermediates in insect and plant cells. *J. Gen. Virol*. 83: 1211-1221.

Milne, R.G. 1988. Taxonomy of the rod-shaped filamentous viruses. The plant viruses. 4. pp. 3-50.

Mink, G.I., Vetten, H.J., Ward, C.W., Berger, P.H., Morales, F.J., Myers, J.M., Silbernagel, M.J., Barnett, O.W. 1994. Taxonomy and classification of legume-infecting potyviruses. *Arch. Virol.* 139: 231-235.

Mlotshwa, S., Verver, J., Sithole-Niang, I., Van Kampen, T., Van Kammen, A., Wellink, J. 2002. The genomic sequence of cowpea aphid-borne mosaic virus and its similarities with other potyviruses. *Arch. Virol.* 147: 1043-1052.

Mori, M., Mise. K., Kobayashi, K., Okuno, T., Furusawa, I. 1991. Infectivity of plasmides containing brome mosaic cDNA linked to the cauliflower mosaic virus 35S promoter. *J. Gen. Virol*. 72: 243-246.

Morris, J.A., Dodds, T.J. 1979. Isolation and analysis of double–stranded RNA from virus-infected plant and fungal tissue. *Phytopathology* 69: 854-858.

Murphy, J.F., Rhoads, R.E., Hunt, A.G., Shaw, J.G. 1990. The VPg of tobacco etch virus RNA is the 49-kDa proteinase or the N-terminal 24 kDa part of the proteinase. *Virology* 178: 285-288.

Mushegian, A., Shepherd, R. 1995. Genetic elements of plant viruses as tools for genetic engineering. *Microbiol. Rev*. 59: 548-578.

Nakashima, H., Sako, N., Hori, K. 1993. Nucleotide sequence of the helper component-proteinase genes of aphid transmissible and non-transmissible isolates of turnip mosaic virus. *Arch. Virol*. 131: 17-27.

Neemann, M. 1999. Klonierung und Expression des Hüllproteingens des Beet mosaic virus (BtMV). Diplomarbeit. Institute of plant diseases and plant protection, Hannover University, Germany.

Nemchinov, L.G., Hammond, J., Jordan, R., Hammond, R.W. 2004. The complete nucleotide sequence, genome organization, and specific detection of Beet Mosaic Virus. *Arch. Virol*. 149: 1201-14.

Nicolaisen, M., Johansen, E., Poulsen, G.B., Borkhardt, B. 1992. The 5' untranslated region from pea seedborne mosaic potyvirus RNA as a translational enhancer in pea and tobacco protoplasts. *FEBS Lett.* 303: 169-72.

Nicolas, O., Laliberte, J.F. 1992. The complete nucleotide sequence of turnip mosaic potyvirus RNA. *J. Gen. Virol* . 73: 2785-2793.

Nicolas, O., Pirone, T.P., Hellmann, G.M. 1996. Construction and analysis of infectious transcripts from a resistance-breaking strain of tobacco vein mottling potyvirus. *Arch. Virol*. 141: 1535-1552.

Nicolas, O., Dunnington, S.W., Gotow, L.F., Piron, T.P., Hellmann, G.M. 1997. Variations in the VPg protein allow a potyvirus to overcome Va gene resistance in tobacco. *Virology* 237: 452-459.

Nikolic, V. 1958. Ispitivanja stentnosti mozaika na industrijskoj secernojrepi. Zastita bilja 46: 69-73.

Olsen, B.S., Johansen, I.E. 2001. Nucleotide sequence and infectious cDNA clone of the L1 isolate of *Pea seed-borne mosaic potyvirus*. *Arch. Virol*. 146: 15-25.

Parks, T.D., Dougherty, W.G. 1991. Substrate recognition by the NIa proteinase of two potyviruses involves multiple domains: characterization using genetically engineered hybrid proteinase molecules. *Virology* 182: 17-27.

Paulsen, A.R., Sill, S.W. 1970. Absence of cross-protection between maize dwarf mosaic virus strains A and B in grain sorghum. *Plant Dis. Rep*.54: 627-629.

Peng, Y., Kadoury, D., Gal-On A., Huet, H., Wang, Y., Raccah, B. 1998. Mutations in the HC-Pro gene of Zucchini yellow mosaic potyvirus: effects on aphid transmission and binding to purified virions. *J. Gen. Virol*. 79: 897-904.

Pirone, T.P. 1972. Sugar cane mosaic virus. *CMI/ABB Descriptions of plant viruses. No* 88.

Pirone, T.P. 1981. Efficiency and selectivity of the helper-component-mediated aphid transmission of purified potyvirus. *Phytopathology* 71: 922-924.

Pirone, T.P., Thornbury, D.W. 1983. Role of virion and helper component in regulating aphid transmission of tobacco etch virus. *Phytopathology* 73: 872-875.

Pirone, T.P., Blanc, S. 1996. Helper –dependent vector transmission of plant viruses. *Ann. Rev. Phytopathology* 34: 227-247.

Pirone, T.P., Raccah, B., Madden, L.V. 1988. Suppression of aphid colonization by insecticides: effect on the incidence of potyviruses in tobacco. *Plant Dis*. 72: 350-353.

Plisson, C., Drucker, M., Blanc, S., German-Retana, S., Le Gall, O., Thomas D., Bron, P. 2003. Structural characterisation of HC-Pro, a plant virus multifunctional protein. *J. Biol Chem.* 278: 23753-61.

Potgieter, A.D., Steele, A. D., van Dijk, A.A. 2002. Cloning of complete genome sets of six dsRNA viruses using an improved cloning method for large dsRNA genes. *J. Gen. Virol* 83: 2215-2223.

Puurand, U., Valkonen, J.P.T., Makinen, K., Rabenstein, F., Sarma, M. 1996. Infectious *in vitro* transcripts from cloned cDNA of the potato A potyvirus. *Virus Res*. 40: 135-140.

Puustinen, P., Rajamaki, M.L., Ivannov, K.I., Valkonen, J.P.T., Makinen, K. 2002. Detection of the Potyviral Genome-Linked Protein VPg in Virions and Its Phosphorylation by Host Kinases. *J. Virol.* 76:12703-12711.

Raccah, B. 1986. Nonpersistent viruses: epidemiology and control. *Adv. Virus Res.* 31: 387-429.

Rajamaki, M.L., Valkonen, J.P.T. 1999. The 6K2 protein and the VPg of potato Virus A are Determinants of systemic infection in Nicandra physaloides. *Mol. Plant Microbe Interact*. 12: 1074-1081.

Rajamaki, M.L., Valkonen, J.P. 2002. Viral genome-linked protein (VPg) controls accumulation and phloem-loading of a potyvirus in inoculated potato leaves. *Mol. Plant Microbe Interact*. 15: 138-49.

Rasco-Gaunt, S., Riley, A., Barcelo, P., Lazzeri, P. 1999. Analysis of particle bombardment parameters to optimise DNA delivery into wheat tissues. *Plant Cell Reports* 19: 118–127.

Restrepo, M.A., Freed, D.D., Carrington, J.C. 1990. Nuclear transport of plant potyviral proteins. *Plant Cell* 10: 987-98.

Restrepo-Hartwig, M.A., Carrington, J.C. 1992. Regulation of nuclear transport of plant potyvirus protein by autoproteolysis. *J. Virol.* 66: 5662-5666.

Restrepo-Hartwig, M.A., Carrington, J.C. 1994. The tobacco etch potyvirus 6-kilodalton protein is membrane associated and involved in virus replication. *J. Virol*. 68: 2388-2397.

Riechmann, J.L., Cervera, M.T., Garcia, J.A. 1995. Processing of the plum pox virus polyprotein at the P3-6K1 junction is not required for virus viability. *J. Gen. Virol*. 76: 951-956.

Riechmann, J.L., Lain, S., Garcia, J.A. 1989. The genome-linked protein and 5' end RNA sequence of plum pox virus. *J. Gen. Virol*. 70: 2785-2789.

Riechmann, J.L., Lain, S., Garcia, J.A. 1990. Infectious *in vitro* transcripts from a plum pox virus cDNA clone. *Virology* 177: 710-716.

Riechmann, J.L., Lain, S., Garcia, J.A. 1992. Highlights and prospects of potyvirus molecular biology. *J. Gen. Virol*. 73: 1-16.

Riedel, D., Lesemann, D.E., Maiss, E. 1998. Ultrastructural localisation of non-structural and coat proteins of 19 potyviruses using antisera to bacterially expressed protein of plum pox potyvirus. *Arch. Virol*. 143: 2133-2158.

Robaglia, C., Durand-Tardif, M., Tronchet, M., Boudazin, G., Astier-Manifacier, S., Casse-Delbart, F. 1989. Nucleotide sequence of potato virus Y (N strain) genomic RNA. *J. Gen. Virol.* 70: 935-947.

Roberts, I.M., Wang, D., Findlay, K., Maule, A.J. 1998. Ultrastructural and temporal observations of the potyvirus cylindrical inclusions (CIs) show that the CI protein acts transiently in aiding virus movement. *Virology* 245: 173-181.

Roberts, J.M.F., Hodgson, C.J., Jackai, L.E.N., Thottappilly, G., Singh, S.R. 1993. Interaction between two synthetic pyrethroids and the spread of two nonpersistent viruses in cowpea. *Ann. Appl. Biol*. 122: 57-67.

Rodriguez-Cerezo, E., Shaw, J.G. 1991. Two newly detected nonstructural viral proteins in potyvirus infected cells. *Virology* 185: 572-579.

Rodriguez-Cerezo, E., Klein, P.G., Shaw, J.G. 1991. A determinant of disease symptom severity is located in the 3'-terminal noncoding region of the RNA of a plant virus. *Proc. Natl. Acad. Sci. U S A*. 88: 9863-7.

Rodriguez-Cerezo, E., Findlay, K., **Shaw, J.G. 1997.** The coat protein and cylindrical inclusion protein of potyvirus are associated with connections between plant cells. *Virology* 236: 296-306

Rodriguez-Cerezo, E., Ammar, E.D., Pirone, T.P., Shaw, J.G. 1993. Association of the non-structural P3 viral protein with cylindrical inclusions in potyvirus-infected cells. *J. Gen. Virol*. 74: 1945-1949.

Rogov, V.V., Karasev, A.V., Agranovsky, A.A., Gorbunova, N.I. 1991. Characterization of an isolate of beet mosaic virus from south Kazakhstan. *J. Plant Pathol*. 40: 515-523.

Rojas, M.R., Zerbini, F.M., Allison, R.F., Gilbertson, R.L., Lucas, W.J. 1997. Capsid protein and helper component-proteinase function as potyvirus cell-to-cell movement proteins. *Virology* 237: 283-295.

Russel, G.E. 1971. Beet mosaic virus. *CMI/AAB Description of plant viruses, No,* 53.

Saenz, P., Salvador, B., Simon-Mateo, C., Kasschau, K.D., Carrington, J.C., Garcia, J.A. 2002. Host specific involvement of the HC protein in the long-distance movement of potyviruses. *J. Virol*. 76: 1922-1931.

Sako, N., Ogata, K. 1981. Different helper factors associated with aphid transmission of some potyviruses. *Virology* 144: 260-267.

Salmon, R. 1989. A possible mechanism for exclusion of plant viruses from embryonic and meristem culture. *Res. Virol*. 140: 453-460.

Salmon, R., Raccah, B. 1990. The role of the N-terminus of potyvirus coat in aphid transmission. Abstract. P 83-7, *VIII Int. Cong. Virol*. Berlin, Aug. 26-31.

Sambrook, J., Fritish, E.F., Maniatis, T. 1989. Molecular Cloning: A laboratory manual, 2th ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sanchez, F., Martinez-Herrera, D., Aguilar, I., Ponz, F. 1998. Infectivity of turnip mosaic potyvirus cDNA clones and transcripts on the systemic host *Arabidopsis thaliana* and local lesion hosts. *Virus Res*. 55: 207-219.

Sasaya, T., Torrance, L., Cowan, G., Ziegler, A. 2000. Aphid transmission studies using helper component proteins of potato virus Y expressed from a vector derived from potato virus X . *J. Gen. Virol*. 81:.1115-1119.

Schaad, M.C., Haldeman-Cahill, R., Cronin, S., Carrington, J.C. 1996. Analysis of the VPg-proteinase (NIa) encoded by tobacco etches potyvirus: effects of mutations on subcellular transport, proteolytic processing and genome amplification. *J. Virol.* 70: 7039-7048.

Schmelzer, K., Hartleb, H. 1977. Beta-und Brassica-Rüben. In. Klinkowski, M. *et al*., Eds. Pflanzliche Virologie 3, 3rd Edition, Akademie-Verlag Berlin, p. 114-143.

Schneider, F., Mundry, K.W. 1956. Beet mosaic virus. Die elektronenmikroskopische Darstellung des Mosaikvirus der Zuckerrübe. *Naturforschung* 11: 393-394.

Shinlaku, M., Palukaitis, P.F. 1989. Genetic mapping of cucumber mosaic virus. *In Symp on Viral Genes in Plant Pathogenesis,* Lexington.

Shukla, D.D., Ward, C.W. 1988. Amino acid sequence homology of coat proteins as a basis for identification and classification of the potyvirus group. *J. Gen. Virol*. 69: 2703-2710.

Shukla, D.D., Ward, C.W. 1989a. Identification and classification of potyviruses on the basis of coat protein sequence data and serology. *Arch. Virol*. 106: 171-200.

Shukla, D.D., Ward, C.W. 1989b. Structure of potyvirus coat proteins and its application in the taxonomy of the potyvirus group. *Adv. Virus Res*. 36:273-314.

Shukla, D.D., Ward, C.W. 1989c. Amino acid sequence homology of coat proteins as a basis for identification and classification of the poty virus group. *J. Gen. Virol*. 69: 2703-2710.

Shukla, D.D., Frenkel, M.J., Ward, C.W. 1991. Structure and function of the potyvirus genome with special reference to the coat protein-coding region. Can. J. Plant Pathol. 13: 178-191.

Shukla, D.D., Jilka, J., Tosic, M., Ford, R.E. 1989. A novel approach to the serology of potyviruses involving affinity- purified polyclonal antibodies directed towards virus-specific N-termini of coat protein. *J. Gen. Virol*. 70: 13-23.

Shukla, D.D., Ford, R.E., Tosic, M., Jilka, J. Ward, C.W. 1989a. Possible members of the potyvirus group transmitted by mites or whiteflies share epitopes with aphid-transmitted definitive members of the group. *Arch. Virol*. 105:.143-150.

Shukla, D.D., Tosic, M., Jilka, E., Ford, R.E, Toler R.W., Langham M.A.C. 1989b. Taxonomy of potyviruses infected maize sorghum and sugarcane in Australia and the United States as determined by reactivities of polyclonal antibodies directed towards virus-specific N-termini of coat proteins. *Phytopathology* 79: 223-229.

Shukla, D.D., Lauricella, R., Ward, C.W. 1992a. Serology of potyviruses: current problems and some solutions. *Arch. Virol*. 5: 57-69.

Shukla, D.D., Frenkel, M.J., McKern, N.M., Ward, C.M., Jilka, J., Tosic, M., Ford, R.E. 1992b. Present status of the sugarcane mosaic subgroup of potyviruses. *Arch. Virol*. 5: 363-373.

Shukla, D.D., Ward. C.W., Brunt, A.A. 1994. The Potyviridae. CAB. International Wallingford, UK, p. 516.

Simon-Buela L., Guo, H.S., Garcia, J.A. 1997. Long sequences in the 5' noncoding region of plum pox virus are not necessary for viral infectivity but contribute to viral competitiveness and pathogenesis. *Virology* 23: 157-62.

Soumounou, Y., Laliberte, J.F. 1994. Nucleic-acid binding properties of the P1 protein of turnip mosaic potyvirus produced in *Escherichia coli*. *J. Gen. Virol* 75:2567-2573.

Spetz, C., Valkonen, J.P.T. 2003. Genomic sequence of *Wild potato mosaic virus* as compared to the genomes of other potyviruses. *Arch. Virol*. 148: 373-380.

Staki, C., Jasnic, S. 1985. Effect of beet mosaic virus on germination of sugar beet seeds and on the length of primary roots seedlings. Zastita-Bilja. 36: 167-170.

Stenger, D.C., French, R. 2004. Functional replacement of Wheat streak mosaic virus HC-Pro with the corresponding cistron from a diverse array of viruses in the family Potyviridae. *Virology* 323: 257-67.

Suehiro, N., Natsuaki, T., Watanabe, T., Okuda, S. 2004. An important determinant of the ability of Turnip mosaic virus infect Brassica spp. and/or Raphanus sativus is in its P3 protein. *J. Gen. Virol*. 85: 2087-98.

Sylvester, E.S. 1949. Beet mosaic virus-green peach aphid relationships. *Phytopathology* 39: 417-424.

Sylvester, E.S. 1952. Comparative transmission of beet-mosaic virus by four aphid species. *Phytopathology* 42: 252, 1952.

Tacahashi, Y., Uyeda, I. 1999. Restoration of the 3_ end of potyvirus RNA derived from poly(A)-deficient infectious cDNA clones. *Virology* 265: 147–152

Takeda, N., Kuhn, R.J., Yang, C.F., Takegami, T., Wimmer, E. 1986. Initiation of poliovirus plus-strand RNA synthesis in a membrane complex of infected HeLa cells. *J. Virol*. 60: 43-53.

Thompson, J.D., Gibson, T.J., Plewniak, F., Higgins, G.D. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res*. 24: 4876-4882.

Thornbury, D.W., Hellman, G.M., Rhoads, R.E., Pirone, T.P. 1985. Purification and characterization of potyvirus helper component. *Virology* 144: 260-267.

Timpe, U., Maiss, E., J. Landsmann., Casper, R. 1992. Coat protein mediated “cross protection" gegen das Scharka Virus. *Mitteilungen der Biologischen Bundesanstalt fuer Land- und Forstwirtschaft*. 238: 212.

Tobin, G.J., Young, D.C., Flanegan, J.B. 1989. Self-catalyzed linkage of poliovirus terminal protein VPg to poliovirus RNA. *Cell* 59: 511-19.

Turpen, T. 1989. Molecular cloning of a potato virus Y genome: nucleotide sequence homology in non-coding regions of potyviruses. *J. Gen. Virol*. 70: 1951-60.

Urcuqui-Inchima, S., Haenni, L,A., Bernardi, F. 2001. Potyvirus proteins: a wealth of functions. *Virus Res*. 74:157-175.

Urcuqui-Inchima, S., Maia, I.G. Drugeon, G., Haenni, A.L., Bernardi, F. 1999. Effect of mutations within the Cys-rich region of potyvirus helper component-proteinase on self-interaction. *J. Gen. Virol*. 80: 2809-2812.

Urcuqui-Inchima, S., Maia, I.G., Arruda, P., Haenni, A.L., Bernardi, F. 2000. Deletion mapping of the potyviral helper component-proteinase reveals two regions involved in RNA binding. *Virology* 268: 104-11.

Van Regenmortel, M.H.V., Bishop, D.H.L., Fauquet, C.M., Mayo, M.A., Maniloff, J., Calisher, C.H. 1997. Guidelines to the demarcation of virus species. *Arch. Virol*. 142: 1505-1518.

Vance, V.B., Moore, D., Turben, T.H., Bracker, A., Hollowell, V.C. 1992. The complete nucleotide sequence of pepper mottle virus genomic RNA: comparison of the encoded polyprotein with those other sequenced potyviruses. *Virology* 191: 19-30.

Varrelmann, M., Maiss, E. 2000. Mutations in the coat protein gene of plum pox virus suppress particle assembly, heterologous encapsidation and complementation in transgenic plants of Nicotiana Benthamiana. *J. Gen. Virol*. 81: 567-576.

Verchot, J., Carrington, J.C. 1995. Debilitation of plant potyvirus infectivity by P1 proteinase-inactivating mutation and restoration by second-site modifications. *Virology* 69: 1582-1590.

Verchot, J., Carrington, J.C. 1995a. Evidence that the potyvirus P1 Proteinase functions in trans as an accessory factor for genome amplification. *J. Virol.* 69: 3668-3674.

Verchot, J., Koonin, E.V., Carrington, J.C. 1991. The 35-kDa protein from the N-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. *Virology* 185: 527- 535.

Viry, M., Serghini, M.A., Hans. F., Ritzenthaler, C., Pinck, M., Pinck, L. 1993. Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. *J. Gen. Virol.* 74: 169-174.

Voloudakis, A.E., Malpica, C.A., Aleman-Verdaguer, M.E., Stark, D.M., Fauquet, C.M., Beachy, R.N. 2004. Structure characterization of Tobacco etch virus coat protein mutants. *Arch. Virol*. 149:699-712.

Vorwoerd, T.C. Dekker, B.M. Hoekema, A. 1989. A small-scale procedure for the rapid isolation of plant RNAs. *Nucl. Acids Res*. 25: 2362.

Wang, D, Hayes, I.M., Maule, A.J. 1992. Procedures for the efficient purification of pea seed-borne mosaic virus and its genomic RNA. *J. Virol. Methods*. 36: 223-30.

Ward, C. J., Shukla, D. D. 1991. Taxonomy of potyviruses: current problems and some solutions. *Intervirology* 32: 269-296.

Watson, D.J., Watson, M.A. 1953. Comparative physiological studies on the growth of field crops. III. The effect of infection with beet yellow and beet mosaic viruses on the growth and yield of the sugar beet crop. *Ann. Appl. Biol*. 40. 1-37.

Wong, K.K., Lisa, C.S., Christopher, A.D., Jeffrey, D.S. 1996. Use of tagged random hexamer amplification (TRHA) to clone and sequence minute quantities of DNA-application to a 180 Kb plasmid isolated from *Sphingomonas* F199. *Nucl. Acids Res.* 24: 3778-3783.

Xiao, X.W., Frenkel, M.J., Ward, C.W., Shukla, D.D. 1994. Nucleotide sequence of the 3'-terminal region of the genome confirms that pea mosaic virus is a strain of bean yellow mosaic potyvirus. *Arch. Virol*. 136: 381-7.

Yang, S.G., Revers, F., Souche, S., Lot, H., Gall, O.L., Candresse, T., Dunez, J. 1998. Construction of full-length cDNA clones of lettuce mosaic virus (LMV) and the effects on their viability in Escherichia coli and on their infectivity to plants. *Arch. Virol.* 143: 2443-2451.

Zeyen, R.J., Berger, P.H. 1990. Is the concept of short retention times for aphid-borne nonpersistent plant viruses sound? *Phytopathology* 80: 769-771.

Zheng, H.Y., Chen, J., Zhao, M.F., Lin, L., Chen, J.-P., Antoniw, J.F. Adams, M.J. 2003. Occurrence and sequences of *Lily mottle virus* and *Lily symptomless virus* in plants grown from imported bulbs in Zhejiang province, China. *Arch. Virol*. 148: 2419-2428.

A. Appendix

```
aaaattaaaacatctcaatacaacacatgcaaaaacatacaaacaagccaacaaactttcaaattcaaaaca    72

tttcacaacaatccacccgatttacttgaagcccaaatttcaagcaaaatcatcagatatttaccatagtcc   144

                          M   A   A   M   M   H   F   G   Q   F   P   S       12
aggaagttgcaatttttctaaa ATG GCA GCA ATG ATG CAC TTT GGT CAA TTC CCA AGC     202

 D   I   P   L   R   A   A   T   C   C   T   K   K   H   S   P   L   V      30
GAC ATT CCA CTC AGG GCA GCT ACA TGC TGC ACA AAG AAA CAC AGT CCT TTG GTT    256

 T   K   E   M   M   A   S   F   A   M   P   T   E   S   S   S   V   A      48
ACG AAA GAA ATG ATG GCC TCC TTC GCG ATG CCA ACT GAG AGC AGT TCT GTT GCT    310

 R   P   I   I   Y   S   S   A   A   T   D   S   Y   E   K   A   Q   R      66
CGA CCT ATT ATA TAC TCT TCT GCA GCC ACT GAT AGC TAT GAG AAA GCT CAG CGA    364

 A   F   E   A   S   L   R   E   K   Y   S   E   K   L   E   A   M   K      84
GCT TTT GAA GCA TCT CTC AGA GAG AAA TAC TCT GAA AAA TTA GAG GCC ATG AAA    418

 Y   G   K   V   V   K   K   G   G   L   T   Y   V   K   R   A   G   P     102
TAT GGC AAA GTG GTC AAG AAG GGC GGT TTG ACG TAT GTC AAA CGT GCT GGA CCC    472

         M
 Q   A   I   A   R   G   V   E   M   D   A   A   I   E   K   F   N   T     120
CAA GCA ATT GCG AGG GGC GTT GAG ATG GAT GCT GCG ATT GAG AAA TTC AAT ACA    526
          G

 A   F   N   A   G   E   L   E   D   V   T   L   E   G   D   I   T   A     138
GCA TTC AAT GCT GGA GAA TTA GAA GAT GTC ACA TTG GAA GGT GAC ATC ACA GCG    580

 G   I   S   V   T   R   G   E   S   V   W   L   R   S   V   F   W   S     156
GGA ATT TCT GTT ACC CGT GGT GAA TCT GTG TGG CTC AGA AGC GTA TTC TGG AGC    634

 R   S   L   K   K   Q   A   R   K   K   T   P   K   L   V   A   K   S     174
CGT TCT CTA AAG AAA CAG GCG AGG AAG AAA ACC CCG AAG CTT GTC GCA AAA TCT    688
          G

 D   F   D   D   L   F   N   K   V   L   K   V   A   S   L   G   N   I     192
GAC TTT GAT GAT TTA TTC AAC AAA GTT CTG AAA GTA GCC TCA CTC GGG AAT ATC    742

 P   V   E   I   V   G   K   K   A   N   K   T   L   R   C   G   Y   K     210
CCT GTT GAA ATT GTT GGA AAG AAG GCC AAC AAA ACT CTC AGA TGC GGA TAC AAA    796

 R   V   N   T   S   T   I   P   Y   F   H   L   P   H   H   N   S   N     228
AGA GTC AAC ACC TCA ACT ATC CCA TAC TTC CAC TTA CCC CAC CAT AAT AGT AAC    850

 Y   I   C   R   E   L   H   P   Q   R   V   R   W   L   V   P   L   L     246
TAC ATC TGC AGG GAA CTA CAT CCA CAA CGT GTT CGT TGG TTA GTT CCT CTC CTC    904

 V   R   H   R   K   I   R   D   R   F   N   D   S   M   I   T   R   G     264
GTG CGT CAT AGG AAA ATT AGA GAT CGA TTC AAT GAT TCC ATG ATC ACA AGA GGT    958
```

Figure A1: Complete nucleotide sequence of BTMV-G, determined from a full-length clone. Nucleotide and amino acid residue numbering is indicated on the right. Nucleotide changes and amino acid changes compared to the firstly determined BtMV-G sequence are indicated below and above the sequences, respectively.

```
  W   S   G   L   I   L   P   K   N   I   A   S   T   C   G   R   R   Y     282
TGG AGT GGT CTG ATT CTG CCC AAG AAC ATT GCA TCA ACA TGC GGA CGA CGC TAT   1012

                                                            V
  D   E   V   I   V   R   G   R   L   Y   G   R   V   E   D   A   R   T     300
GAT GAG GTA ATA GTA AGA GGA AGG CTT TAT GGA AGA GTG GAG GAT GCT CGA ACG   1066
                                                            T

                                        V
  K   L   P   A   G   D   V   G   R   T   M   H   Y   S   S   G   E   E     318
AAG TTA CCT GCA GGC GAT GTT GGT AGA ACC ATG CAC TAC AGT AGT GGA GAG GAA   1120
                                        G

  R   F   F   A   G   W   K   E   G   F   E   K   L   V   P   A   Q   K     336
AGA TTC TTT GCT GGA TGG AAA GAG GGA TTT GAG AAG CTT GTT CCA GCT CAG AAA   1174

  E   H   V   C   K   I   V   Q   D   N   K   F   C   G   K   L   A   A     354
GAG CAT GTC TGC AAG ATA GTA CAA GAC AAC AAG TTT TGT GGA AAA CTG GCA GCA   1228

  S   I   V   Q   I   A   F   P   C   H   K   M   A   C   D   V   C   R     372
TCA ATA GTG CAA ATC GCT TTT CCA TGC CAC AAA ATG GCC TGT GAT GTG TGT AGG   1282

  N   K   F   N   E   M   T   P   E   A   Y   S   E   L   V   E   K   H     390
AAT AAG TTT AAT GAG ATG ACT CCA GAG GCA TAC TCT GAA CTT GTT GAG AAA CAC   1336

  I   D   Q   R   M   S   E   I   N   E   A   V   A   Q   F   P   G   L     408
ATA GAC CAA CGA ATG AGC GAG ATT AAT GAA GCA GTT GCT CAG TTT CCA GGT TTG   1390

  K   Q   V   V   S   N   F   R   S   K   H   T   V   N   T   N   M   K     426
AAA CAA GTT GTT TCA AAT TTC AGA AGC AAG CAC ACT GTT AAC ACC AAC ATG AAG   1444

  D   N   L   E   V   A   K   L   T   Q   G   H   K   A   N   Q   M   M     444
GAT AAT CTG GAA GTT GCG AAA TTG ACA CAA GGG CAC AAG GCC AAC CAG ATG ATG   1498

  Q   L   A   K   I   N   S   I   L   I   K   G   N   T   A   T   P   D     462
CAG CTT GCC AAA ATT AAT TCC ATT CTG ATC AAG GGA AAT ACA GCA ACA CCT GAT   1552

  E   M   G   D   A   S   G   L   L   L   E   I   T   R   W   F   N   N     480
GAA ATG GGC GAT GCT AGT GGC TTG TTA CTC GAG ATA ACA AGA TGG TTC AAC AAC   1606

  H   L   S   I   V   D   K   G   S   L   R   A   F   R   N   K   R   S     498
CAT TTG AGC ATA GTT GAT AAA GGT TCC TTG CGG GCT TTT AGA AAT AAA CGT TCA   1660

  N   K   A   L   V   N   P   S   L   L   C   D   N   Q   R   D   K   N     516
AAC AAA GCA CTA GTC AAT CCG TCG CTT TTG TGT GAC AAT CAA AGG GAT AAG AAT   1714

  G   N   F   I   W   G   E   R   G   Y   H   S   K   R   F   F   A   S     534
GGG AAC TTC ATT TGG GGT GAA AGA GGA TAT CAT TCG AAG CGT TTC TTT GCT AGT   1768

  Y   F   D   E   V   T   P   G   D   G   Y   K   E   Y   I   I   R   K     552
TAC TTC GAT GAA GTA ACT CCT GGG GAT GGA TAC AAG GAA TAC ATA ATT CGC AAA   1822

  G   P   Q   G   Q   R   K   L   A   I   G   N   L   I   V   S   F   D     570
GGA CCA CAG GGT CAA AGG AAG CTT GCA ATT GGA AAC TTG ATC GTC TCG TTT GAT   1876

  L   E   K   T   R   Q   A   L   K   G   E   E   V   E   K   L   L   L     588
TTA GAG AAA ACA CGA CAA GCA CTT AAG GGA GAA GAG GTT GAG AAA CTC CTT TTG   1930
```

Figure A1: continued

```
 S   N   S   C   I   S   K   R   N   G   N   Y   V   Y   T   S   C   C     606
AGC AAC TCA TGC ATC AGT AAA CGT AAT GGA AAT TAC GTT TAC ACT AGC TGT TGT   1984

 V   T   L   D   D   G   T   P   L   Y   S   N   I   K   N   P   T   K     624
GTC ACT CTA GAT GAC GGA ACC CCA CTG TAT TCC AAT ATT AAG AAC CCA ACA AAG   2038

 R   H   L   I   V   G   T   S   G   D   P   K   I   V   D   L   P   A     642
AGA CAC TTG ATT GTT GGA ACA AGT GGA GAT CCC AAG ATA GTC GAC CTA CCT GCT   2092

 T   D   T   D   K   M   Y   I   A   K   E   G   Y   C   Y   L   N   I     660
ACA GAC ACA GAC AAG ATG TAC ATA GCG AAA GAG GGG TAT TGC TAC TTG AAT ATC   2146

 F   L   A   M   L   I   N   V   N   E   N   E   A   K   A   F   T   K     678
TTT TTA GCT ATG CTC ATT AAT GTC AAT GAG AAT GAG GCA AAG GCC TTC ACA AAA   2200

 M   V   R   D   I   I   I   P   M   L   G   T   W   P   T   M   Q   D     696
ATG GTC CGT GAC ATC ATT ATC CCA ATG CTT GGA ACA TGG CCA ACC ATG CAG GAT   2254

 L   A   T   A   C   F   M   M   T   A   F   F   P   E   T   S   S   A     714
CTT GCA ACA GCA TGC TTC ATG ATG ACA GCT TTC TTT CCA GAA ACG AGT AGT GCT   2308

                                     T
 E   L   P   R   I   L   V   D   H   A   N   Q   T   M   H   V   I   D     732
GAG CTT CCA AGA ATA CTA GTT GAC CAC GCA AAT CAG ACA ATG CAT GTC ATT GAT   2362
                                     A

                                                             T
 S   F   G   S   L   T   T   G   Y   H   V   L   K   A   G   A   A   A     750
TCA TTT GGC TCA TTG ACA ACT GGG TAT CAT GTG CTC AAG GCA GGA GCA GCC GCT   2416
                                                             A

                 F
 Q   L   I   D   S   A   S   T   E   L   D   G   E   M   K   W   Y   R     768
CAG CTC ATT GAC TCT GCT TCA ACA GAG CTT GAC GGA GAG ATG AAA TGG TAC AGA   2470
                 T

 V   G   G   H   G   L   P   V   K   E   K   M   I   S   A   L   I   T     786
GTT GGC GGT CAT GGA CTA CCC GTG AAA GAG AAG ATG ATT AGT GCA TTG ATC ACG   2524

 S   I   Y   R   P   K   K   L   V   Q   L   I   E   E   D   P   Y   V     804
AGC ATT TAC AGA CCC AAG AAA TTG GTG CAA TTG ATT GAA GAA GAT CCA TAC GTT   2578

 L   I   M   A   M   S   S   P   R   L   I   I   N   L   F   N   N   G     822
CTA ATC ATG GCT ATG AGC TCA CCA AGG TTA ATT ATC AAT TTA TTC AAT AAT GGT   2632

 A   L   E   L   A   A   K   H   W   I   S   R   D   K   N   V   S   A     840
GCC TTA GAG CTA GCA GCA AAA CAT TGG ATA TCA CGT GAT AAA AAT GTC AGT GCA   2686

 I   F   A   M   L   M   D   L   S   T   E   M   S   K   A   E   L   L     858
ATC TTC GCG ATG CTT ATG GAC CTA TCT ACG GAG ATG TCT AAG GCA GAA CTT CTC   2740

 I   E   Q   H   R   M   I   N   E   C   A   K   R   I   H   D   T   Q     876
ATT GAG CAG CAT CGT ATG ATA AAC GAA TGT GCC AAG AGA ATT CAT GAC ACT CAG   2794

 N   Y   L   D   E   V   G   P   H   Q   Q   E   V   R   T   F   L   A     894
AAC TAC CTT GAT GAA GTG GGT CCA CAC CAG CAA GAG GTT CGC ACT TTC TTG GCA   2848
```

Figure A1: continued

```
  L   I   S   D   E   L   E   V   D   K   E   L   H   K   T   G   F   A     912
TTG ATT TCG GAT GAA CTG GAA GTG GAT AAG GAA CTG CAC AAG ACG GGA TTT GCA   2902

          G
  N   F   S   E   R   F   H   S   L   T   E   K   M   Y   V   D   A   L     930
AAT TTC AGT GAG CGA TTT CAC TCC CTC ACT GAA AAA ATG TAT GTG GAC GCA TTA   2956
          G

  E   E   E   W   R   G   L   S   L   L   D   K   F   S   Y   A   T   F     948
GAA GAG GAA TGG CGA GGT TTA AGC TTG TTA GAC AAA TTC TCC TAT GCC ACC TTT   3010

  V   C   K   H   K   P   R   S   T   P   V   L   P   P   R   K   S   E     966
GTG TGC AAG CAC AAA CCG CGT TCA ACG CCC GTT TTG CCC CCG AGA AAG TCA GAA   3064

  D   I   D   A   K   F   V   I   S   P   S   W   F   V   G   K   T   R     984
GAT ATC GAC GCC AAA TTC GTC ATA TCG CCC AGC TGG TTC GTT GGA AAG ACG AGG   3118

  E   H   L   N   G   G   R   K   Y   V   T   S   Q   I   T   Q   F   T    1002
GAA CAC CTG AAT GGA GGC CGG AAG TAT GTT ACA AGT CAG ATC ACT CAA TTC ACA   3172

  N   Y   I   K   R   A   T   L   D   K   A   M   R   I   M   C   S   C    1020
AAC TAC ATC AAG CGC GCC ACA CTC GAC AAA GCC ATG CGC ATC ATG TGC AGC TGC   3226

  L   K   D   L   A   Y   F   M   N   V   A   L   V   T   H   L   L   I    1038
CTA AAG GAC CTC GCA TAC TTT ATG AAC GTT GCT TTG GTA ACT CAT TTG CTC ATT   3280

                                                        I
  S   M   I   A   A   V   Y   K   M   L   N   D   H   R   V   A   K   H    1056
AGC ATG ATT GCA GCA GTA TAT AAA ATG CTG AAT GAT CAC AGG GTT GCG AAG CAC   3334
                                                        A

  R   L   R   I   L   E   M   Q   A   V   D   H   T   I   Y   C   L   Y    1074
CGT TTG CGC ATA CTC GAG ATG CAA GCG GTA GAC CAC ACA ATC TAT TGC CTT TAT   3388
                                            G

                  N
  D   T   W   K   T   V   H   N   R   E   P   T   S   K   E   F   R   Q    1092
GAC ACA TGG AAA ACT GTT CAC AAC AGG GAG CCA ACC TCA AAA GAG TTC AGA CAG   3442
                  A

  Y   I   A   T   I   N   K   D   L   L   K   Y   L   P   E   E   E   G    1110
TAC ATC GCC ACC ATC AAC AAG GAC CTT TTG AAG TAC TTA CCC GAG GAA GAA GGC   3496

  K   A   E   V   E   Y   Q   A   N   K   V   Y   E   K   K   L   E   K    1128
AAA GCG GAG GTC GAG TAT CAA GCA AAC AAA GTC TAC GAG AAG AAG CTA GAA AAA   3550

  A   V   A   L   M   A   L   F   T   M   I   F   D   T   E   K   S   G    1146
GCA GTC GCT CTT ATG GCT TTG TTC ACT ATG ATA TTT GAT ACA GAG AAG AGT GGG   3604

  A   V   F   S   I   L   R   N   I   K   S   V   F   S   T   L   G   E    1164
GCA GTG TTC AGC ATT TTA CGC AAC ATT AAA TCA GTG TTT AGC ACA CTC GGA GAA   3658

  E   V   K   Y   Q   S   L   D   E   I   Q   S   I   E   D   E   K   K    1182
GAA GTG AAG TAT CAA AGT CTT GAT GAA ATA CAA TCG ATC GAA GAT GAG AAG AAG   3712

  L   T   I   D   F   D   L   D   T   E   I   T   A   E   H   T   T   M    1200
CTC ACA ATT GAC TTT GAT TTG GAT ACT GAA ATC ACG GCT GAG CAC ACC ACG ATG   3766
```

Figure A1: continued

```
 D   V   Q   F   E   K   W   W   D   K   Q   L   G   Q   N   R   V   V    1218
GAT GTG CAA TTT GAG AAG TGG TGG GAC AAG CAA TTG GGT CAG AAC AGA GTT GTT   3820

 P   H   Y   R   V   G   G   T   F   I   E   F   T   R   H   T   A   A    1236
CCG CAC TAC AGA GTG GGC GGA ACA TTC ATT GAG TTC ACG AGG CAC ACA GCA GCT   3874
                       A

 S   V   C   N   T   I   C   A   S   S   E   Q   E   F   V   V   R   G    1254
AGT GTT TGT AAC ACT ATT TGC GCA AGT TCT GAA CAG GAA TTC GTT GTT CGT GGG   3928

 A   V   G   S   G   K   S   T   G   L   P   S   H   L   S   R   K   G    1272
GCA GTT GGT TCA GGG AAA TCT ACG GGT TTG CCG TCG CAT TTA AGT CGA AAA GGG   3982

                                                                E
 R   V   L   L   L   E   P   T   R   P   L   A   E   N   V   C   K   Q    1290
CGA GTT TTA TTA CTC GAA CCA ACA CGA CCA CTT GCA GAG AAC GTG TGC AAG CAG   4036
                                                                G

 L   R   K   E   P   F   H   L   S   P   T   L   R   M   R   G   L   T    1308
CTT CGA AAG GAG CCA TTC CAC TTA TCT CCA ACG CTC AGA ATG CGA GGA TTA ACA   4090

 T   F   G   S   S   N   I   S   V   M   T   S   G   Y   A   L   H   F    1326
ACA TTT GGC TCA AGC AAC ATA AGT GTA ATG ACT AGT GGC TAT GCG CTT CAC TTC   4144

 H   A   N   N   P   Q   R   L   E   E   F   D   F   I   M   I   D   E    1344
CAC GCA AAC AAT CCA CAG AGG CTT GAA GAG TTT GAC TTC ATA ATG ATC GAT GAA   4198

 S   H   T   M   D   S   S   T   M   A   F   Y   C   L   L   R   E   Y    1362
AGC CAT ACA ATG GAC TCT TCC ACT ATG GCT TTT TAT TGC CTC CTT CGT GAA TAC   4252

 E   F   K   G   K   I   L   K   V   S   A   T   P   P   G   R   E   C    1380
GAA TTT AAG GGC AAG ATA CTC AAA GTC TCG GCA ACG CCA CCA GGG AGA GAG TGT   4306

 E   F   K   T   Q   H   D   V   L   I   K   I   E   E   S   L   S   Y    1398
GAA TTC AAG ACA CAG CAC GAT GTT CTT ATC AAG ATT GAA GAG TCT TTG TCA TAC   4360

 N   S   F   I   T   A   Q   G   T   G   S   N   A   D   V   V   Q   N    1416
AAC TCT TTT ATC ACA GCT CAA GGC ACT GGC TCA AAC GCT GAC GTC GTT CAA AAT   4414

 G   D   N   I   L   V   Y   V   P   S   Y   N   D   V   D   Q   L   S    1434
GGT GAC AAC ATT CTT GTA TAC GTC CCA AGC TAC AAT GAT GTT GAT CAA CTG AGC   4468

 K   G   L   M   E   K   G   Y   L   V   T   K   V   D   G   R   T   M    1452
AAG GGG CTA ATG GAG AAG GGC TAC CTG GTC ACA AAG GTG GAT GGA CGC ACA ATG   4522

 K
 R   M   G   N   V   E   I   P   T   K   G   T   S   S   K   K   H   F    1470
AGG ATG GGC AAC GTT GAG ATC CCA ACG AAA GGA ACT TCT TCA AAG AAG CAC TTT   4576
 A

 I   V   A   T   N   I   I   E   N   G   V   T   L   D   I   D   V   V    1488
ATA GTT GCA ACA AAC ATC ATT GAG AAT GGT GTG ACA TTA GAC ATT GAT GTT GTT   4630

 V   D   F   G   L   K   V   V   A   E   L   D   S   D   S   R   C   M    1506
GTG GAC TTT GGT CTG AAG GTA GTT GCC GAA TTG GAT TCT GAC TCA CGT TGC ATG   4684

 R   Y   K   K   V   S   I   S   Y   G   E   R   L   Q   R   L   G   R    1524
CGC TAC AAG AAA GTG TCA ATT AGC TAT GGT GAA AGA CTT CAA CGA CTT GGT AGA   4738
```

Figure A1: continued

```
 V   G   R   V   K   Q   G   T   A   L   R   I   G   H   T   E   T   G    1542
GTA GGA CGC GTC AAG CAG GGC ACA GCT CTT CGC ATT GGC CAT ACT GAG ACA GGC   4792

 M   T   E   I   P   V   A   I   A   T   E   A   A   F   I   C   F   A    1560
ATG ACT GAG ATC CCT GTT GCG ATT GCA ACT GAA GCA GCC TTC ATT TGT TTT GCC   4846

 Y   N   L   P   V   M   T   H   N   V   T   S   S   L   L   S   R   C    1578
TAC AAC TTG CCA GTG ATG ACA CAC AAT GTA ACA TCA AGT CTT CTC TCG CGA TGC   4900

 T   N   R   Q   A   R   T   M   M   Q   Y   E   L   S   P   F   F   M    1596
ACA AAC AGG CAA GCC AGA ACA ATG ATG CAA TAT GAG CTG TCC CCA TTC TTT ATG   4954
                                            A

 V   E   L   V   H   F   N   G   C   V   H   P   Q   I   E   S   K   L    1614
GTC GAA CTA GTT CAT TTT AAT GGC TGT GTG CAC CCA CAA ATA GAG AGC AAA CTG   5008

 K   A   Y   K   L   R   D   S   E   T   Q   L   S   T   L   A   I   P    1632
AAG GCG TAC AAG TTA CGA GAC TCT GAA ACA CAA CTG AGC ACG CTT GCC ATT CCG   5062
                                G

 N   S   G   T   S   R   W   K   T   V   G   E   Y   K   K   L   G   V    1650
AAT AGC GGG ACT TCC CGA TGG AAG ACA GTC GGA GAA TAC AAG AAG CTG GGA GTT   5116

                                                                I
 R   I   E   A   D   D   N   V   R   V   P   F   A   A   N   G   V   P    1668
CGT ATT GAA GCA GAC GAT AAT GTT CGC GTT CCG TTT GCA GCA AAT GGA GTT CCT   5170
                                                                A

 D   R   L   Y   A   D   L   W   E   T   I   Q   Q   H   K   S   D   A    1686
GAT AGG CTA TAT GCT GAT TTG TGG GAA ACC ATT CAG CAG CAT AAA TCT GAT GCG   5224

 G   F   G   R   L   T   S   A   C   A   S   K   I   S   Y   T   L   T    1704
GGG TTC GGA CGG CTC ACA AGT GCT TGT GCG AGC AAG ATT TCA TAC ACT CTG ACA   5278

 T   Q   P   N   A   I   P   R   T   L   A   I   I   E   H   L   L   R    1722
ACT CAA CCA AAT GCA ATT CCT CGC ACT TTG GCA ATA ATT GAG CAT TTA TTA AGA   5332

 E   E   Q   Q   K   K   A   Y   F   E   S   L   N   D   T   L   C   A    1740
GAA GAG CAG CAA AAG AAA GCC TAC TTC GAA TCT CTT AAT GAC ACG CTT TGT GCA   5386
                                                                    G

                                                        R
 T   S   F   S   L   A   G   M   V   N   N   I   R   R   G   Y   L   K    1758
ACA AGC TTT TCG CTA GCA GGT ATG GTC AAT AAC ATA CGA AGA GGA TAT CTG AAA   5440
                                                        A

 D   H   S   A   H   N   I   N   V   L   Q   N   A   K   S   Q   L   N    1776
GAT CAT TCA GCG CAC AAT ATC AAT GTG CTC CAG AAC GCT AAG TCG CAG CTT AAT   5494

 E   F   N   S   K   A   V   D   P   E   R   I   G   D   L   M   G   Y    1794
GAA TTT AAT TCG AAA GCT GTT GAT CCA GAA AGA ATT GGA GAT CTG ATG GGA TAT   5548

 G   V   L   D   T   V   Q   Y   Q   S   A   T   D   I   Q   K   R   L    1812
GGA GTG CTG GAC ACT GTT CAG TAC CAA AGT GCA ACA GAT ATC CAA AAG CGA TTG   5602

 K   L   K   G   R   W   N   G   S   L   A   A   T   D   F   L   I   A    1830
AAG CTG AAA GGT CGA TGG AAT GGT TCA TTA GCT GCA ACA GAT TTT CTA ATA GCA   5656
```

Figure A1: continued

```
 G   T   V   F   A   G   G   C   W   M   L   W   E   Y   A   K   S   G    1848
GGA ACC GTA TTT GCT GGT GGT TGC TGG ATG CTA TGG GAG TAT GCT AAG AGT GGA   5710

 N   E   I   V   Q   Y   Q   G   K   R   R   Q   M   Q   K   L   K   F    1866
AAT GAA ATT GTG CAA TAC CAA GGG AAG AGG CGC CAG ATG CAG AAG CTT AAG TTT   5764

 R   N   A   R   D   N   K   V   G   R   E   V   Y   G   D   D   G   T    1884
AGG AAT GCC AGA GAT AAT AAG GTT GGG CGC GAA GTA TAT GGT GAC GAT GGA ACA   5818

 I   E   H   L   F   G   A   A   Y   T   E   R   G   K   R   K   G   N    1902
ATT GAA CAC CTT TTC GGA GCA GCT TAC ACT GAA AGA GGA AAA AGG AAG GGA AAC   5872

 N   S   T   K   G   M   G   T   K   T   R   R   F   V   H   M   Y   G    1920
AAT AGT ACC AAG GGA ATG GGA ACA AAG ACG AGA CGC TTC GTA CAC ATG TAT GGA   5926

 F   D   P   T   E   Y   S   F   V   R   F   V   D   P   L   T   G   Y    1938
TTT GAT CCT ACA GAG TAC TCA TTT GTG CGA TTC GTA GAC CCA TTG ACT GGT TAT   5980

 S   K   D   E   S   V   Q   T   D   I   A   L   V   Q   S   E   I   G    1956
TCA AAG GAT GAG AGC GTG CAA ACT GAC ATT GCC TTA GTG CAG AGC GAG ATA GGT   6034

 E   Y   R   Q   K   C   M   E   D   D   D   E   L   I   D   F   I   K    1974
GAG TAC AGA CAG AAG TGC ATG GAA GAT GAT GAC GAA TTG ATC GAC TTC ATC AAG   6088

 Q   K   P   G   I   Q   A   Y   F   M   K   N   G   S   D   K   A   L    1992
CAG AAA CCT GGA ATT CAA GCT TAT TTC ATG AAA AAT GGC TCA GAT AAG GCT TTA   6142

 Q   V   D   L   T   P   H   I   P   L   L   S   C   A   K   T   A   T    2010
CAA GTC GAC CTC ACA CCT CAC ATT CCT TTG TTG TCA TGT GCT AAG ACT GCC ACA   6196

 I   A   G   F   P   E   R   E   S   E   L   R   Q   T   G   T   P   I    2028
ATC GCT GGT TTT CCT GAA CGA GAA TCT GAA TTA CGG CAA ACT GGA ACC CCA ATT   6250

 V   V   S   K   N   V   V   P   G   E   H   T   E   V   V   R   E   E    2046
GTT GTG AGC AAA AAT GTG GTG CCG GGT GAG CAC ACA GAG GTT GTC AGG GAA GAG   6304

 G   K   S   I   V   K   G   L   R   N   Y   N   P   I   S   S   I   V    2064
GGA AAG TCC ATA GTC AAA GGA CTA CGG AAT TAT AAT CCA ATC TCT TCA ATT GTT   6358

 C   R   L   T   N   D   S   N   G   N   A   Q   T   L   Y   G   I   G    2082
TGC CGT TTG ACC AAT GAC TCG AAT GGA AAT GCT CAA ACC CTT TAT GGG ATT GGT   6412

 F   G   P   L   I   I   T   N   S   H   L   F   K   A   N   N   G   T    2100
TTT GGT CCA TTG ATC ATC ACG AAC AGT CAT TTG TTC AAG GCA AAT AAT GGC ACA   6466

 L   F   I   R   S   H   Q   G   E   F   T   V   Q   N   T   T   Q   L    2118
TTG TTC ATC AGA TCA CAT CAA GGG GAA TTC ACC GTG CAG AAC ACA ACG CAG CTC   6520

                         E
 Q   V   H   H   V   K   D   K   D   M   I   L   I   R   M   P   K   D    2136
CAA GTC CAT CAT GTG AAA GAC AAA GAC ATG ATC TTG ATC CGT ATG CCA AAA GAT   6574
                           G

                         R
 F   P   P   F   P   M   K   L   K   F   R   A   P   H   S   E   E   R    2154
TTC CCA CCA TTT CCG ATG AAG CTG AAG TTT AGA GCA CCC CAC TCT GAG GAA AGG   6628
                           G
```

Figure A1: continued

```
                                                        S
  A   C   L   V   G   S   R   F   Q   Q   K   G   L   S   S   E   V   S    2172
GCT TGC TTA GTC GGT TCA CGT TTC CAA CAG AAG GGC CTT TCA AGT GAA GTC TCC    6682
                                        A

  D   S   T   L   I   R   P   T   D   S   G   S   G   Y   W   K   H   W    2190
GAT TCA ACC TTA ATA CGA CCC ACA GAT TCA GGA AGT GGA TAT TGG AAA CAT TGG    6736

  V   S   T   K   E   G   D   C   G   L   P   M   V   A   L   K   D   G    2208
GTC TCA ACA AAG GAA GGT GAT TGT GGT TTG CCT ATG GTC GCC TTG AAG GAC GGT    6790

  S   V   I   G   I   H   G   L   T   S   V   R   S   E   L   N   Y   F    2226
AGT GTG ATA GGG ATT CAT GGC TTG ACG AGC GTG CGG TCA GAG TTG AAC TAT TTC    6844

  V   P   F   T   D   D   F   Q   S   K   Y   L   S   N   I   D   S   L    2244
GTT CCC TTC ACT GAT GAC TTC CAG TCA AAG TAT CTT TCA AAC ATA GAT TCG CTT    6898

  E   W   V   K   H   W   R   H   T   P   D   K   V   A   W   N   G   M    2262
GAA TGG GTG AAA CAT TGG AGG CAT ACA CCA GAC AAA GTT GCA TGG AAC GGA ATG    6952

  T   L   R   E   N   G   P   A   S   E   F   S   V   S   K   L   I   A    2280
ACA TTG CGA GAG AAT GGT CCA GCA TCA GAA TTT TCG GTT AGC AAG CTC ATA GCA    7006

  D   L   T   H   G   Y   F   D   E   V   V   E   Q   G   Y   S   S   K    2298
GAT CTA ACA CAT GGT TAC TTT GAT GAG GTT GTT GAA CAA GGG TAC AGT TCG AAA    7060

  W   V   A   N   R   L   D   G   N   L   K   A   V   A   S   S   S   S    2316
TGG GTT GCT AAT CGT CTT GAT GGC AAT CTT AAG GCT GTT GCT AGT TCT TCC AGT    7114

  Q   L   V   T   K   H   V   V   K   G   P   C   V   L   F   Q   E   F    2334
CAG CTT GTG ACG AAA CAT GTA GTA AAG GGG CCA TGC GTT TTA TTT CAA GAG TTT    7168
                              T

  L   A   T   H   E   E   A   A   R   Y   F   V   P   R   M   G   E   Y    2352
CTT GCA ACA CAT GAG GAA GCA GCT AGA TAT TTT GTG CCA CGC ATG GGA GAG TAC    7222

                                                        R
  G   P   S   R   L   N   K   E   A   F   L   K   D   F   L   K   Y   A    2370
GGT CCA AGT CGT CTC AAT AAA GAA GCC TTT CTT AAG GAT TTC TTA AAG TAT GCT    7276
                                                        G

  G   P   I   T   V   G   V   V   N   T   D   S   F   E   D   A   V   A    2388
GGT CCG ATA ACT GTA GGT GTT GTC AAC ACT GAC AGT TTT GAA GAT GCA GTT GCA    7330

  S   V   I   N   M   L   E   D   L   D   Y   G   E   C   A   Y   V   T    2406
AGT GTT ATC AAC ATG CTA GAA GAC CTG GAT TAT GGT GAG TGC GCT TAT GTC ACT    7384

  D   P   D   S   I   F   D   S   L   N   M   K   A   A   V   G   A   L    2424
GAT CCT GAC TCA ATA TTC GAT TCC CTC AAC ATG AAA GCA GCC GTC GGG GCT CTA    7438

  Y   K   G   K   K   K   E   Y   F   E   Q   L   N   S   A   E   R   E    2442
TAT AAA GGG AAG AAG AAA GAG TAC TTT GAA CAA CTC AAC TCT GCA GAG CGC GAG    7492

  N   L   L   R   L   S   C   E   R   L   Y   E   G   K   M   G   V   W    2460
AAT TTA CTA CGG CTT AGT TGT GAG CGC TTG TAC GAA GGC AAG ATG GGC GTA TGG    7546

  N   G   S   L   K   A   E   L   R   P   K   E   K   L   E   Q   N   K    2478
AAT GGA TCT CTC AAA GCT GAG TTA CGT CCC AAG GAA AAA CTG GAA CAA AAT AAG    7600
                                                            G
```

Figure A1: continued

```
 T   R   T   F   T   A   A   P   I   D   T   L   L   G   G   K   V   C    2496
ACT CGT ACA TTC ACT GCA GCA CCT ATT GAC ACT TTG CTT GGC GGG AAA GTT TGT   7654

 V   D   D   F   N   N   R   F   Y   S   L   N   L   K   G   P   W   S    2514
GTT GAC GAT TTC AAC AAC AGA TTT TAC TCG TTA AAC CTC AAG GGT CCA TGG TCT   7708

 V   G   M   T   K   F   Y   G   G   W   N   E   L   L   Q   K   L   P    2532
GTC GGA ATG ACG AAG TTT TAT GGA GGG TGG AAC GAA TTG TTA CAG AAG CTT CCA   7762

 D   G   W   I   Y   C   D   A   D   G   S   Q   F   D   S   S   L   T    2550
GAT GGT TGG ATC TAT TGT GAT GCC GAT GGT TCA CAA TTT GAT AGC TCT CTC ACA   7816

 P   Y   L   I   N   A   V   V   Q   I   R   E   H   F   M   E   D   W    2568
CCT TAC CTG ATC AAT GCT GTG GTG CAG ATA AGG GAG CAT TTC ATG GAG GAC TGG   7870

 E   I   G   R   T   M   L   R   N   F   Y   T   E   I   V   Y   T   P    2586
GAA ATT GGT AGG ACC ATG CTT CGA AAT TTT TAC ACT GAG ATA GTG TAC ACA CCA   7924

 I   L   T   P   D   G   T   I   V   K   K   F   K   G   N   N   S   G    2604
ATT CTC ACA CCA GAC GGA ACA ATA GTA AAG AAG TTC AAG GGG AAT AAC AGT GGC   7978

             S
 Q   P   S   T   V   V   D   N   T   L   M   V   I   L   A   M   H   Y    2622
CAA CCA TCA ACA GTT GTT GAC AAC ACT TTA ATG GTT ATC CTT GCA ATG CAT TAC   8032
              T

 A   M   H   Q   Q   C   W   R   E   E   E   M   K   E   K   I   R   F    2640
GCG ATG CAT CAG CAG TGC TGG AGG GAA GAG GAA ATG AAA GAG AAG ATT CGG TTC   8086

 F   A   N   G   D   D   L   L   I   A   I   H   P   S   K   E   K   F    2658
TTC GCA AAT GGT GAT GAT CTG CTA ATA GCT ATA CAT CCG AGT AAG GAG AAA TTC   8140

 L   N   V   L   S   E   Y   F   H   E   L   G   L   K   Y   D   F   S    2676
TTG AAC GTG CTC TCA GAA TAT TTC CAC GAG CTT GGG TTG AAA TAT GAT TTC TCG   8194

 S   R   S   T   V   R   E   T   L   W   F   M   S   H   R   G   L   Y    2694
AGC AGA AGC ACC GTG CGC GAG ACT TTG TGG TTC ATG TCA CAC AGG GGG CTT TAC   8248

 H   D   D   M   Y   I   P   K   L   E   E   E   R   I   V   S   I   L    2712
CAT GAT GAT ATG TAT ATC CCT AAA CTT GAA GAG GAA CGC ATC GTT TCA ATT CTT   8302

 E   W   D   R   S   N   E   A   T   H   R   A   E   A   I   C   A   A    2730
GAA TGG GAC AGG AGC AAT GAA GCC ACA CAT AGA GCT GAA GCA ATA TGT GCT GCA   8356

 M   I   E   A   W   G   Y   P   E   L   L   K   Y   I   R   E   F   Y    2748
ATG ATT GAG GCA TGG GGA TAC CCA GAA CTT TTG AAG TAC ATT CGT GAG TTT TAT   8410

 L   W   M   M   H   H   E   C   Y   R   D   L   V   R   D   G   K   L    2766
CTA TGG ATG ATG CAT CAT GAA TGT TAT AGG GAT CTG GTG CGT GAC GGA AAG CTA   8464

 P   Y   I   A   E   T   A   L   R   K   L   Y   T   D   K   S   V   D    2784
CCA TAT ATT GCA GAG ACA GCT CTT AGA AAG CTC TAC ACT GAC AAG AGC GTG GAT   8518

 E   S   E   L   V   K   Y   W   K   A   L   A   P   E   E   E   D   G    2802
GAG AGT GAG CTT GTC AAG TAC TGG AAA GCA CTT GCA CCA GAG GAA GAG GAT GGC   8572

 P   D   I   V   T   Y   Q   G   D   E   K   P   S   K   S   S   Q   P    2820
CCG GAT ATA GTC ACA TAC CAA GGT GAT GAG AAA CCC TCA AAG TCT TCA CAG CCA   8626
```

Figure A1: continued

```
  Q   S   S   S   P   Q   V   P   Q   Q   V   D   A   G   A   S   S   Q    2838
CAG AGC TCA AGT CCA CAA GTG CCA CAA CAA GTT GAT GCA GGA GCT AGC TCT CAA    8680
                                          C

  G   K   D   K   Q   S   V   I   K   H   D   S   T   K   S   K   D   V    2856
GGA AAG GAC AAG CAG AGT GTC ATC AAG CAC GAT AGC ACA AAG AGC AAA GAT GTG    8734

  G   Q   S   S   T   A   V   P   R   L   K   Q   I   S   K   M   R   M    2874
GGA CAA TCG AGC ACT GCT GTT CCA AGA TTG AAG CAG ATC TCA AAA ATG AGG ATG    8788

  P   V   S   K   G   R   Q   V   L   A   L   D   H   L   L   D   Y   K    2892
CCT GTT AGC AAA GGA CGT CAA GTT TTA GCA CTT GAC CAT CTA TTG GAT TAT AAG    8842

  P   E   Q   V   D   L   S   N   T   R   A   T   R   E   Q   F   D   N    2910
CCA GAA CAG GTC GAT CTC TCA AAC ACT AGA GCC ACG AGG GAG CAG TTT GAC AAT    8896

  W   Y   E   A   V   M   K   E   Y   D   V   S   D   S   Q   M   G   V    2928
TGG TAT GAA GCA GTC ATG AAG GAA TAC GAT GTG TCA GAT TCC CAA ATG GGT GTG    8950

  I   M   N   G   L   M   V   W   C   I   E   N   G   T   S   P   N   L    2946
ATT ATG AAT GGA CTC ATG GTG TGG TGT ATT GAG AAT GGC ACA TCA CCA AAT CTC    9004

  S   G   D   W   V   M   M   D   G   E   E   Q   V   S   F   P   L   K    2964
AGT GGA GAC TGG GTC ATG ATG GAT GGA GAG GAA CAA GTC TCA TTC CCT TTG AAG    9058

          I
  P   I   M   E   N   A   K   P   S   F   R   Q   I   M   H   H   F   S    2982
CCG ATA ATG GAA AAT GCT AAA CCA TCT TTT CGG CAA ATA ATG CAT CAC TTT TCT    9112
            A

  D   A   A   E   A   Y   I   E   M   R   N   R   E   R   P   Y   M   P    3000
GAT GCA GCA GAA GCG TAT ATT GAA ATG CGC AAC AGA GAA AGG CCA TAC ATG CCT    9166

  R   Y   G   A   Q   R   N   L   R   D   R   T   L   A   R   Y   A   F    3018
CGT TAT GGC GCT CAG AGA AAT CTG AGA GAC AGA ACG CTA GCT CGC TAT GCA TTC    9220

  D   F   Y   E   V   T   S   R   T   T   D   R   A   R   E   A   H   F    3036
GAT TTC TAT GAG GTC ACC TCA CGA ACA ACT GAT CGT GCA CGT GAA GCT CAT TTC    9274

  Q   M   K   A   A   A   L   A   S   V   S   N   K   L   F   G   L   D    3054
CAA ATG AAG GCG GCA GCG TTG GCA AGC GTG TCC AAC AAG CTC TTT GGG CTT GAT    9328

  G   S   V   A   T   T   S   E   D   T   E   R   H   T   A   T   D   V    3072
GGG AGC GTG GCC ACC ACA TCG GAG GAT ACA GAG AGG CAC ACA GCC ACA GAT GTC    9382

  N   A   H   M   H   H   M   M   G   V   R   Q   G                        3085
AAC GCT CAC ATG CAT CAC ATG ATG GGC GTT AGG CAA GGT  taattctgtacctcgttct  9440

atgaatagttaaatatggtaaccatttaaaagagtgaggttttacctccgttgcttatttctatttcgcata  9512

gttccaaaccactaccctcaataggcgtctcatagtgaggttttacctcggaggattctacggacggtacac  9584

aggtttac                                                                  9592
```

Figure A1: continued