Sequence Variability of *Cucumber mosaic virus* (CMV) and its Effects on CMV-Resistance of *Capsicum* sp.

Dissertation

A thesis submitted to the
Fachbereich Biologie, Universität Hamburg
for the degree of
doctor rerum naturalium

By

Deyong Zhang

Hequ, China

Hamburg, 2005

Genehmigt vom
Fachbereich Biologie
der Universität Hamburg

Auf Antrag von Herrn Professor Dr. Günter Adam
weiterer Gutachter der Dissertation:
Prof. Dr. Edgar Maiss

Tag der Disputation: 27. Mai 2005
Hamburg, den 16. März 2005

Prof. Dr. Arno Frühwald
Dekan des Fachbereiches Biologie

To my parents Mitan Zhang and Runying Wang

To my wife Jing Dai

Table of Content

Abbreviations

µ	micro (10^{-6})
°C	centigrade
A	Adenine
AA, aa	amino acid
Ala	alanine
Amp	Ampicillin
AMV	*Alfalfa mosaic virus*
AS	antiserum
Asn	asparagine
AV genes	avirulence genes
AVRDC	The World Vegetable Centre
BCIP	5-bromo-4-chloro-3-indoxyl-phosphate
bp	basepair
C	Cytidine
cDNA	complementary DNA
ChiVMV	*Chili veinal mottle virus*
CLCV	*Cabbage leaf curl virus*
CMV	*Cucumber mosaic virus*
CP	coat protein
cv.	cultivar
DAS	double antibody sandwich
ddH_2O	double distilled water
DI	defective interfering
DIECA	Diethyldithiocarbamate
d.p.i.	days post inoculation
DMSO	Dimethylsulfoxide
DNA	desoxyribonucleic acid
dNTPs	dATP, dGTP, dCTP, dTTP
ds	double strand
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
EDTA	Ethylene diamintetraacetic acid
ELISA	enzyme linked immunosorbent assay
EtBr	Ethidium bromide
G	guanine
GTZ	Deutsche Gesellschaft fuer Technische Zusammenarbeit GmbH
ha	hectare
HC	helicase
HC-Pro	helper component-proteinase
HR	hypersensitive reaction
Ile	isoleucine
INSV	*Impatiens necrotic spot virus*
IPTG	Isopropylthio-ß-D-galactoside
K	kilodalton
kb	kilobasepair
Mabs	monoclonal antibodies
min	minute
MOPS	4-morpholino propanesulfonic acid
MP	movement protein
mRNA	messenger RNA
MT	methyltransferase
Mt	million tons

NA	nucleic acid
NBT	Nitroblue tetrazolium chloride
nm	nanometer
nts	nucleotides
OD	optical density
ORF	open reading frame
PBS	phosphate-buffered saline buffer
PBS-T	PBS Tween
PCR	polymerase chain reaction
PDR	pathogen derived resistance
PEG	Polyethylene glycol
PMMV	*Pepper mild mottle virus*
PeVMV	*Pepper veinal mottle virus*
PPV	*Plum pox virus*
PSV	*Peanut stunt virus*
PTGS	post transcriptional gene-silencing
PVP	Polyvinylpyrrolidone
PVX	*Potato virus X*
PVY	*Potato virus Y*
R gene	resistance gene
RdRp	RNA-dependent RNA Polymerase
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNAi	RNA interference
RT	reverse transcriptase
rt	room temperature
RT-PCR	reverse transcription and polymerase chain reaction
rpm	rounds per minute
SA	Salicylic acid
satRNA	satellite RNA
SDS	Sodiumdodecylsulfat
Ser	serine
sgRNA	subgenomic RNA
SNP	single nucleotide polymorphism
SS	single strand
T	Thymine
TAS	Triple Antibody Sandwich
TAV	*Tomato aspermy virus*
TE	Tris-EDTA
TEV	*Tobacco etch virus*
Thr	threonine
TMV	*Tobacco mosaic virus*
t	ton
TSWV	*Tomato spotted wilt virus*
TRIS	Tris(hydroxymethyl) aminomethane
TRV	*Tobacco rattle virus*
u	unit
v	volume
Val	valine
VIGS	virus induced gene silencing
w	weight
X-gal	5-Bromo-4-chloro-3-indolyl-ß-D-galactoside

1 INTRODUCTION

Botanically, chili peppers are classified among the *Solanaceae* family and the genus *Capsicum*. Five domesticated species, *C. pubescens* R. et P., *C. baccatum* (Willd.) Eshbaugh, *C. chinense* Jacq, *C. frutescens* and *C. annuum* L., have been described and studied extensively (Greenleaf, 1986; Smith *et al.*, 1987). *C. annuum* L., which originated in Mexico and contains both, the large-fruited bell peppers and small pungent types, is the most widely cultivated species. It has become globally the predominant chili, in part because it was the first chili discovered by Columbus and introduced to the rest of the world. More than 450 years ago Portuguese and Spanish traders introduced this fruit to the Asian continent.

Chili is consumed dried or fresh and has many uses. Its popularity as a condiment, spice and vegetable is growing rapidly. Today, chili is an important part of the diet in Asia, from the curries of India to the sambals of Indonesia, from the spicy soups of Thailand to the hot dishes of the Szechuan and Hunan provinces in China. It is an extremely good source of many essential nutrients and is richer in vitamins A and C than the usual recommended sources. Chili has also found application in the medical field, with pungency (capsaicin) being an important pharmacological property and as a colouring agent in the food industry to colour a wide variety of processed foods. An estimated one billion people consume chili in one form or another on a daily basis, making it one of the most widely consumed vegetables on earth.

Worldwide about 23.7 million Mt of chilies are produced on 1,645,985 ha (FAOSTAT data, last updated December, 2004; http://faostat.fao.org), 998,508 ha (60 % of total yield in the world) of which are grown in Asia (FAOSTAT data, last updated December, 2004; http://faostat.fao.org). Chili is an economically important crop because it generates a significant income in local as well as export markets. However, yields in Asian countries are unstable and low (about 5.5 t/ha compared to 10-17 t/ha elsewhere), largely due to its susceptibility to many viral, bacterial and fungal diseases.

The major diseases contributing to low yields and low quality of fruits include bacterial wilt (caused by *Ralstonia solanacearum*), phythophthora blight (caused by *Phytophthora capsici* Leon.), powdery mildew (caused by *Leveillula taurica*) and anthracnose (caused by *Colletotrichum acutatum*, *C. capsici*, and *C. coccoides*, *C. gloeosporioides*) (Hadden and Black, 1987, 1989). In addition, many viruses are known to infect chili (Villalon, 1981; Green and Kim, 1991) such as *Chili leafcurl virus* (CLCV), *Chili veinal mottle virus*

(ChiVMV), *Cucumber mosaic virus* (CMV), *Pepper mild mottle virus* (PMMV), *Pepper veinal mottle virus* (PeVMV), *Tobacco etch virus* (TEV), *Tobacco mosaic virus* (TMV), *Tomato spotted wilt virus* (TSWV) and *Potato virus* Y (PVY). These viruses can infect either singly or in combination and an infection results in various symptoms, which range from a mild to severe mottling, leaf puckering, leaf distortion, necroses of leaves and fruit to extreme plant stunting.

CMV has been described as one of the five most important viruses infecting vegetable species worldwide (Palukaitis and Garcia-Arenal, 2003; Palukaitis *et al.*, 1992). CMV is also one of the most prevalent viruses of the chili. The virus causes severe mosaic symptoms, stunting, various types of necrosis, leaf deformation and leaf shoestring. Fruits are often malformed and necrotic lesions are common, thereby drastically reducing marketable yield (Green and Kim, 1991; Palukaitis *et al.*, 1992).

CMV is the type species of the genus *Cucumovirus (*family *Bromoviridae)*, which contains three distinct species *Cucumber mosaic virus* (CMV), *Peanut stunt virus* (PSV) and *Tomato aspermy virus* (TAV) (Roossinck *et al*., 2000). CMV has icosahedral virions which encapsidate three linear plus-sense single-stranded (ss) genomic RNAs (RNA 1, RNA 2, RNA 3) and the subgenomic (sg) RNAs 4 and 4A (Fig. 1). RNA 1 and RNA 2 are encapsidated in distinct particles, whereas RNA 3 and RNA 4 (Lot and Kaper, 1976) and possibly RNA 3 and RNA 4A are encapsidated together (Gallitelli, 2000). For the initiation of an infection three types of particles containing either RNA 1 or RNA 2 and a combination of RNA 3 and RNA 4 are required.

RNA 1 (about 3300 nucleotides, nts) and RNA 2 (about 3000 nts) code for the 1a (110 K) and 2a (98 K) proteins, respectively, which are the core elements of the replicase complex (Hayes and Buck, 1990). Two functional domains can be distinguished in the 1a protein: the N-terminal region shows sequence homology to methyltransferases (MT) while the C-terminal region contains motifs characteristic of viral and cellular helicases (HC) (Kadare and Haenni, 1997). The 2a protein contains the conserved amino acid sequence motif (GDD) present in many viral polymerases (Habili and Symons, 1989) and is the other viral component of the replicase complex (Ishihama and Barbier, 1994; O'Reilly and Kao, 1998). On RNA 2 a second open reading frame (ORF) encodes the 2b protein, which is translated by subgenomic RNA 4A. The two ORFs of the 2a and 2b genes are partially overlapping (Fig. 1). The 2b protein is a multifunctional protein, which is a host range determinant (Shi *et al.*, 2002) and a suppressor of post-transcriptional gene silencing

(PTGS) (Brigneti *et al.*, 1998; Baulcombe, 2002); it is involved in interfering salicylic acid (SA)-mediated plant defense (Ji and Ding, 2001), has a role in promoting cell-to-cell movement (Shi *et al.*, 2003) and influences the long distance movement (Ding *et al.*, 1995a; Soards *et al.*, 2002). RNA 3 (about 2200 nts) codes for two proteins. The 5´ proximal 3a ORF is translated directly from the genome and encodes the movement protein (MP) of about 30 K in size, which is essential for long-distance movement within the plants (Canto *et al.*, 1997; Li *et al.*, 2001). The 3´ proximal ORF encodes the coat protein (CP) of about 24 K in size, which constructed the virus particles. The CP is translated from the subgenomic RNA 4 with a size of about 1000 nts. Both RNA 3 gene products, MP and CP, are essential for the short-distance movement of CMV (Canto *et al.*, 1997). The two ORFs are separated by a non-translated intergenic region of apporximately 300 nts large.

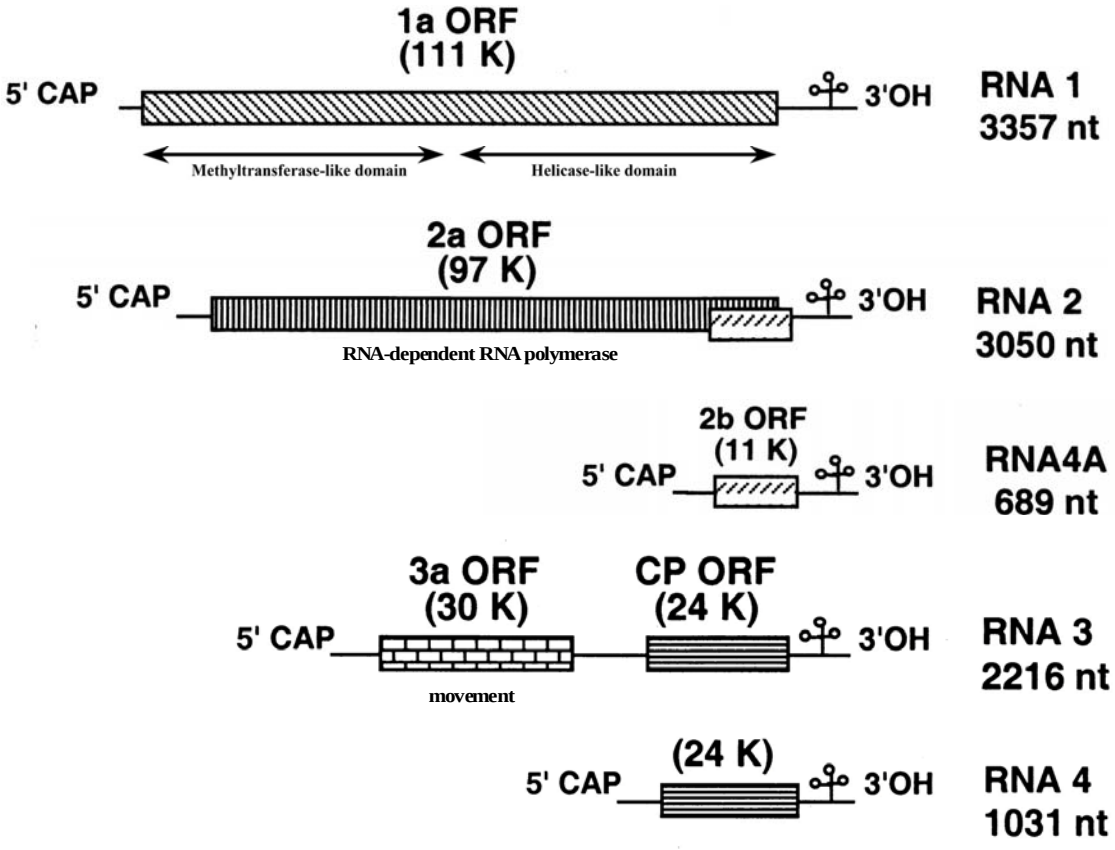


Fig. 1: Schematic genome organization of *Cucumber mosaic virus* (CMV).
Modified from Roossinck (2002) and Moury (2004). Nucleotide (nt) numbers and the sizes for encoded proteins are given for the Fny isolate. K, kilodalton. RNAs 1, 2 and 3 can act as mRNAs. The CP and 2b gene are expressed from a sgRNA, RNA 4 and RNA 4A, respectively. The sg RNAs (RNA 4 and RNA 4A) are encapsidated together with RNA 3 in one virus particle.

Some CMV isolates may contain additional RNAs (Palukaitis and Garcia-Arenal, 2003), either satellites or defective interferings (DI) RNAs.

Satellite RNA (satRNA) are small, linear and encapsidated RNA molecules that are completely dependent on the viral genome for their replication and spread, but do not supply

the helper virus with any essential function. CMV satRNAs range from 332 to 405 nts in size, apparently do not contain any functional ORF, and occur in nature as several variants which can sometimes attenuate or aggravate disease symptoms induced by the helper virus in specific host plants. For example, D-satRNA, B-satRNA and WL1-satRNA induce necrosis, chlorosis and attenuation, respectively, in tomato plants in the presence of any CMV helper virus (Xu and Roossinck, 2000). In tobacco, however, D-satRNA and WL1-satRNA attenuate symptoms, whereas minor nucleotide sequence variants of the B-satRNA either attenuate symptoms or induce chlorosis in a helper virus–specific manner (Garcia-Arenal *et al.*, 2000; Garcia-Arenal and Palukaitis, 1999).

DI RNAs have been described for a number of plant viruses (Adam *et al.*, 1983; White, 1996; Simon and Nagy, 1996; Graves *et al.*, 1996; White and Morris, 1999). DI RNAs are derived from the helper virus genome, by deletions within one or more genes or by complex rearrangements of the genomic sequences (Graves *et al.*, 1996). The DI RNA retains the cis-acting components necessary for efficient replication but depends upon the helper virus for replication and encapsidation functions. The presence of a DI RNA can have profound effects on the symptoms produced by the helper virus, ranging from symptom amelioration to symptom enhancement to no effect at all (Graves and Roossinck, 1995; Graves *et al.*, 1996). In the case of CMV, two DI RNAs derived from the 3a gene of RNA 3 have been described (Graves and Roossinck, 1995). However, the maintenance of DI RNAs was host-dependent (Kaplan *et al.*, 2004).

The cell-to-cell and long-distance movement of CMV involves the participation of each of the five encoded proteins, expressed from the three genomic RNAs and two sgRNAs (Suzuki *et al.*, 1991; Gal-On *et al.*, 1994; Hellwald and Palukaitis, 1995; Ding *et al.*, 1995a,b; Kaplan *et al.*, 1995). Moreover, the MP is involved in a number of functions associated with movement (Ding *et al.*, 1995b; Gal-On *et al.*, 1995; Kaplan *et al.*, 1995; Li and Palukaitis, 1996; Canto and Palukaitis, 1999a,b; Li *et al.*, 2001). This virus-encoded protein interacts with plasmodesmata and facilitates the movement of viral RNA from one cell to another, as well as into and out of the vasculature (Ding *et al.*, 1995a, b; Canto *et al.*, 1997; Blackman *et al.*, 1998).

Plant viruses are often transmitted from one plant to another by vectors, which include insects, nematodes, mites and fungi. The majority of viruses have arthropod vectors (Hull, 2002). CMV is transmitted by more than 85 species of aphids in a nonpersistent manner (Edwardson *et al.*, 1991). The CP has been shown to be the primary determinant for vector

transmission (Gera *et al.*, 1979; Chen and Francki, 1990). Molecular and genetic analyses of isolates that are defective in aphid transmission have revealed that different regions of the CP variably affect transmission and that the efficiency of transmission of CMV varies with the species of the aphid vector. A conservative negatively charged ßH-ßI loop structure on the CP surface is assumed to facilitate aphid vector transmission (Liu *et al.*, 2002).

CMV has been an evolutionarily successful virus with both, a wide host range of more than 885 plant species in 65 families (Palukaitis *et al.*, 1992) and a worldwide distribution.
RNA viruses deserve their reputation as "Nature's swiftest evolvers" because of their high variability, which is thought to be due to three main evolutionary forces, namely, mutation, reassortment (for viruses with a segmented genome) and recombination (Domingo and Holland, 1997; Roossinck, 1997; Aaziz and Tepfer, 1999; Worobey and Holmes, 1999).
Viral RNA replication is characterized by a high mutation rate, due to the lack of proof-reading-repair of viral RNA-dependent RNA polymerases (RdRp). This, in conjunction with short replication time and a high multiplication rate, leads towards a genetically heterogenous population, a unique feature of RNA viruses termed as virus quasi-species (Eigen, 1993) and corresponding to a cloud of sequence variants (Roossinck, 1997). Reassortment occurs only in multipartite viruses and involves swapping one or more of the discrete RNA molecules that make up the segmented viral genome. Artificially constructed reassortants of these viruses, also designated as pseudorecombinants, have proven to be very useful for genetic mapping of functional genes, and in some genera RNA segments can be exchanged even between species (Palukaitis *et al.*, 1992). Phylogenic estimations of the *Cucumovirus* genus were different when the ORFs for each RNA were used separately, supporting the idea that reassortment had occurred during their evolutionary history (Roossinck, 2002). Reassortment may not be a common event, as suggested by a survey of CMV in Spain (Fraile *et al.*, 1997) and in America (Lin *et al.*, 2004), but even a rare occurrence could have a dramatic impact on the evolution of new viral species, especially if the reassortant confers a selective advantage, such as an expanded host and vector range.
Recombination can occur in either segmented or unsegmented viruses when a 'donor' nucleotide sequence is introduced into a single, contiguous 'acceptor' RNA molecule to produce a new RNA containing genetic information from more than one source. The recombination rates in RNA viruses are probably high (Bruyere *et al.*, 2000), and a growing amount of evidence for natural RNA recombination showed that recombination plays an important role in the speciation of viruses. This is also the case for CMV (Aaziz and

Tepfer, 1999; Bonnet *et al*., 2005). In addition, RNA recombination is thought to rescue viral genomes by repairing mutation errors in essential viral genes or in structures that could be introduced during RNA replication (Lai, 1992; Carpenter and Simon, 1996). However, heterogeneity is restricted by the necessity to maintain a functional viral RNA genome, whereas divergence into a new species is selected by environmental pressures.
Variability of CMV is a result of evolution. A number of CMV isolates that have been described previously were classified into the two serotypes I and II according to serological relationships, peptide mapping of the CP, nucleic acid (NA) hybridization, reverse transcriptase (RT)-polymerase chain reaction (PCR) (Palukaitis *et al.*, 1992) as well as by RT-PCR followed by enzymatic digestion of the amplified product (restriction fragment length polymorphism, RFLP) and nucleotide sequence identity (Rizos *et al.*, 1992). Alignment of the 5′ nontranslated regions of RNA 3 of 26 CMV isolates suggests the division of serotype I into subgroups Ia and Ib. Further phylogenetic analyses of CMV using the CP ORF of 53 isolates strongly support this division of serotype I into subgroup Ia and Ib. In addition, isolates within each serotype and subgroup radiate from a single point of origin, indicating that they have evolved from a single common ancestor for each serotype and subgroup (Roossinck *et al.*, 1999). A complete phylogenetic analysis covering all ORFs, 1a, 2a, 2b, 3a and 3b, of 15 CMV isolates indicated that different RNAs had independent evolutionary histories. The phylogenetic trees estimated for ORFs located on the different RNAs were not congruent and did not completely support the subgrouping indicated by the CP ORF. The evolutionary trees of the 1a and 3a ORFs were more compact and displayed more branching than did those of the 2a and CP ORFs (Roossinck, 2002).

To initiate an infection, three RNAs of CMV are required to invade a single cell. This provides an advantage for genetic mapping of the symptom determinants with bioactive transcripts. Indeed, symptom determinants in CMV genomes have been extensively studied by constructing reassorted viruses using CMV isolates that cause different symptoms in the same host plant, mainly tobacco (Palukaitis *et al.*, 1992). The reassortment experiments indicated that some host responses to CMV infection were determined by RNA 1, RNA 2 or RNA 3, but others resulted from the interaction of two or more viral RNAs (Palukaitis and Garcia-Arenal, 2003; Choi *et al*., 2003; Takeshita *et al.*, 2004). Furthermore, many responses have been mapped to specific amino acids (aa) on virus-encoded proteins by reverse genetics (Palukaitis and Garcia-Arenal, 2003; Diveki *et al.*, 2004).

Viral diseases are difficult to control and most measures are preventive. Among those, the use of virus-free or virus resistant plant material, the avoidance of planting contaminated material and the testing of planting material for a pathogen-free status are important, although not exclusive methods for controlling viral diseases of plants (Lopez *et al.*, 2003). The RT-PCR has been proven to be a sensitive and simple method for testing virus presence/absence in a large number of samples. It has been used extensively in virus detection (Choi *et al.*, 1999; Letschert *et al.*, 2002). Although RT-PCR and other methods based on specific sequence identification are undoubtedly effective and highly sensitive, they are labourious for routine testing of very large sample numbers. ELISA and dot blot hybridization are more suitable for this purpose, especially the latter, as blotted membranes prepared elsewhere can easily be mailed to a central laboratory for processing. Since control of CMV must be preventive, accurate diagnosis of the disease causing agent is an essential prerequisite for effective control. To this end, a wide spectrum of methods has been developed and is currently in use in diagnostic laboratories on a day-to-day basis (Anonymous, 1998 and references therein; Boonham *et al.*, 2003).

The volume of sequence data currently available in sequence databases is so large that it is becoming to pose a complex technological challenge, namely, how can such data be mined and examined in a high throughput and cost effective manner. At present, microarray technology is leading the way in offering researchers the ability to examine simultaneously the expression levels of hundreds or thousands of genes in a single array test (Wallace, 1997; Schena *et al.*, 1998; Gerhold *et al.*, 1999). Microarrays or 'gene chips' are being used that have thousands of genes arrayed on the surface (Schena *et al.*, 1998), with current protocols allowing reliable detection of target down to approximately ten copies per cell (Kane *et al.*, 2000). So far the technology has been applied mainly for research, in areas such as expression level profiling, single nucleotide polymorphism (SNP) mapping and studying host pathogen interactions (Marathe *et al.*, 2004). The microarray technology might allow the detection of a large number of different plant viruses in a single generic assay, however, so far only few reports of use of that technique are available (Boonham *et al.*, 2003; Lee *et al.*, 2003).

Another way to control CMV is the use of resistant lines. A passive and constitutive defense of plants against pathogens comprises barriers like the rigid cell wall present in plants prior to any contact with pathogens. Furthermore, plants exhibit active induced defence mechanisms upon recognition of pathogens. "Induced resistance" depends on the

recognition of a pathogen by the plant, which evokes active defence, the so-called hypersensitive response (HR). During this cells, surrounding the primary infection site of the virus, die due to an induced rapid programmed cell death, which results in a visible necrotic local lesion. The induction of this response is preceded by a specific recognition of the virus, and in many cases this is based on matching (dominant) gene products of the plant (produced from dominant resistance genes = R genes) and the virus (avirulence genes = AV genes). To date, a few dozen single dominant R genes recognizing different categories of plant pathogens have been cloned and sequenced (Takken and Joosten, 2000). About half of the resistance alleles studied were dominant, the remainder were either incompletely dominant or recessive. Dominant resistances were the result of active mechanisms, and such efforts correspond to incompatible interactions between viruses and hosts that result in a HR controlled by dominant R genes. Recessive or incompletely dominant resistances were thought to result from loss of factor(s) essential for virus multiplication in the host plant. This kind of resistance is thought to allow the spread of virus, but to reduce multiplication or symptom development. Fully recessive alleles may be associated with complete immunity (Fraser, 1990, 1992, 1999).

Different viral proteins, including RdRp (Meshi *et al.*, 1988; Hamamoto *et al.*, 1997; Padgett *et al.* 1997; Erickson *et al.* 1999), MPs (Meshi *et al.*, 1989; Weber and Pfitzner, 1998; Weber *et al.* 1993), and CPs (Bendahmane *et al.*, 1995; Berzal-Herranz *et al.*, 1995) have been identified as avirulence determinants (Culver, 1997). In one case, a RNA sequence per se also has been described as an elicitor of an HR-like resistance response (Szittya and Burgyan, 2001).

In contrast, much less information is available for incompatible interactions controlled by recessive R genes. Recessive mutations suppressing efficient multiplication of tobamoviruses and potyviruses have been identified and fully characterized in *Arabidopsis thaliana* (Yamanaka *et al.*, 2000, 2002; Duprat *et al.*, 2002; Lellis *et al.*, 2002), but only two natural recessive R genes to potyviruses and one natural recessive R gene to carmoviruses have been characterized to date (Ruffel *et al.* 2002; Nicaise *et al.* 2003; Diaz *et al.*, 2004). The available studies indicate that recessive genes providing resistance to viruses do not belong to the same classes as those controlling HR (Yamanaka *et al.*, 2000, 2002; Duprat *et al.*, 2002; Lellis *et al.*, 2002; Ruffel *et al.*, 2002).

Commercial chili varieties with an adequate level of CMV resistance are not yet available (Green and Kim, 1991; Monma and Sakata, 1997) although some resistance mechanisms

have been proposed for *C. annuum* cultivar (cv.) "Perennial" against CMV. These include a partial resistance to initial viral infection as evidenced by reduced development of local necrotic lesions (Caranta *et al.*, 1997), a reduction in viral replication (Nono-Womdim *et al.*, 1993a) and a restriction of long-distance movement of the virus (Nono-Womdim *et al.* 1993b). Resistance in the cultivar "Perennial" has been reported to be controlled by one to several genes that show recessive or partially dominant inheritance (Lapidot *et al.*, 1997). The wide range of resistance mechanisms and the reported differences in inheritance may be a consequence of the effects of environmental conditions, CMV isolates or genetic backgrounds of the host plants.

Genetic engineering of crop plants using virus-derived genes to confer protection is a well-established procedure known as pathogen-derived resistance (PDR) (Sanford and Johnson, 1985). Expression of different virus sequences has proven highly effective in preventing or reducing viral infections (Beachy *et al.*, 1990, Palukaitis and Zaitlin, 1997; Miller and Hemenway, 1998). Most successful and widely applied has been the exploitation of plant viral CP genes, thus achieving CP-mediated resistance. For a number of RNA viruses, including *Potato virus* X (PVX), *Alfalfa mosaic virus* (AMV), *Tobacco rattle virus* (TRV), TMV and CMV, it has been shown that host plants transgenically expressing their CP, exhibit high levels of operational resistance (Wilson, 1993; Baulcombe, 1994, 1996; Beachy, 1994), which can only be overcome by extremely high inoculation pressure or by inoculation with unencapsidated viral RNA. So far, transgenic crop plants expressing (parts of) functional genes or genome parts of plant viruses have not been widely introduced due to legislation problems related to biosafety issues. Indeed under laboratory conditions, it has been shown that recombination may occur between transgenically expressed viral RNA transcripts and infecting viral genomes (Greene and Allison, 1994; Canto *et al.*, 2001).

An important mechanism involved in PDR is referred as RNA silencing or PTGS. PTGS was first reported in plants 15 years ago as a coordinated and reciprocal inactivation of host genes and transgenes encoding the same RNA (Napoli *et al.*, 1990). Shortly after, a similar phenomenon was observed in the filamentous fungus *Neurospora crassa* and termed quelling (Romano and Macino, 1992; Cogoni *et al.*, 1996; Cogoni and Macino, 1999). In animals, PTGS can be induced by double stranded (ds) RNA in a process called RNA interference (RNAi) (Fire *et al.*, 1998; Fire, 1999). Almost all eukaryotic organisms possess this sequence-specific RNA-degradation system, designated as RNA silencing

(Voinnet, 2001; Waterhouse *et al.*, 2001). It is not only triggered by transgenes but can also be initiated by viruses, a process known as virus induced gene silencing (VIGS). Thus, the discovery that viruses are inducers, and targets, and may carry suppressors of PTGS provide the most compelling evidence that RNA silencing functions as an inducible, host RNA-surveillance system (Marathe *et al.*, 2000; Voinnet, 2001; Li and Ding, 2001; Waterhouse *et al.*, 2001). Among others, the helper component proteinase (HC-Pro) of potyviruses, the p19 protein of tombusviruses, the P1 protein of sobemoviruses and the 2b protein of cucumoviruses have been identified as the suppressors of PTGS (Ding *et al*., 1994; Anandalakshmi *et al*., 1998; Brigneti *et al*., 1998; Voinnet *et al*., 1999; Kasschau and Carrington, 2001; Qu and Morris, 2002; Li *et al*., 2002; Soards *et al.*, 2002; Roth *et al*., 2004).

Although several CMV isolates from chili were well characterized by host range, serology and nucleic acid sequence, there is still a lack of detailed information, e.g. biological and sequence data, for Asian chili isolates. Increased knowledge will lead to a better understanding of viral biodiversity and provide a good basis for molecular-based breeding. There is a need for the development of a parallel, sensitive, high-throughput virus detection which can handle large numbers of isolates for diagnostics and disease resistance breeding programs. So far all methods used at present have been designed for single targets. To improve chili production in Asian countries, it is necessary to develop and/or improve resistant lines by conventional breeding since there are no CMV-resistant lines on the market. Furthermore, the resistance mechanism of chili remains to be investigated, which will enrich the germplasm resources and accelerate breeding programs. The study of genetic determinants of CMV responsible for inducing resistance breaking in chili may open a new panel for molecular breeding and also provide help for a better management of this disease in the farmer's field.

2 Material and Methods

2.1 Material

2.1.1 Plant material

Five chili (*Capsicum annuum* L.) lines selected from the global germplasm collection of the AVRDC were received from S. K. Green (AVRDC, The World Vegetable Centre, Taiwan, China). Four of the lines, VC246, PBC370, PBC549, PBC495, were resistant to cucumber mosaic virus (CMV), and one line, VC27a was susceptible.

2.1.2 CMV isolates

Thirty-eight CMV isolates used in this study, were obtained (a) from Asia (China, Indonesia, India, Thailand) through cooperators of the international project "Development of locally-adapted, multiple disease-resistant, high-yielding chilli (*Capsicum annuum*) cultivars for targeted countries in Asia" (Project No. 2001.7860.8-001.00) funded by the Deutsche Gesellschaft für Technische Zusammenarbeit GmbH (GTZ, Germany), (b) from the "Ringtest" organized by COST Action 823 (Anonymous, 1998), (c) from AVRDC, and (d) from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany). The detailed informations about the isolates are summarized in Table 1.

Table 1: Details of CMV isolates, their designation, country and organization, original host plant and year of isolation

No.	[1]designation	country / origin	original host	[2]year of isolation
1	KB8	Thailand, GTZ	chili	1990
2	KS44	Thailand, GTZ	chili	1991
3	UB50	Thailand, GTZ	chili	1990
4	RS30	Thailand, GTZ	chili	1989
5	TR15	Thailand, GTZ	cucumber	2002
6	CH1	Thailand, GTZ	chili	2002
7	CM1	Thailand, GTZ	chili	2002
8	AN	India, GTZ	chili	2002
9	P3613	China, AVRDC	chili	na
10	P522	China, AVRDC	chili	na
11	Peet's	China, AVRDC	tomato	na
12	CN02	China, GTZ	chili	2002
13	CN03	China, GTZ	chili	2002
14	CN04	China, GTZ	chili	2002
15	CN05	China, GTZ	chili	2002
16	CN06	China, GTZ	chili	2002
17	CN07	China, GTZ	chili	2002
18	hn	China, GTZ	chili	2001
19	141N	China, GTZ	tomato	2003
20	RT6	Greece, ringtest	melon	1995
21	RT52	USA, ringtest	squash	na
22	RT54	China, ringtest	cornflower	1982
23	RT55	China, ringtest	tomato	1989
24	RT67	Netherlands, ringtest	lettuce	1978
25	RT68	Netherlands, ringtest	spinach	1974
26	RT88	Spain, ringtest	melon	1996
27	RT138	Iran, ringtest	bean	1989
28	RT144	Bulgaria, ringtest	bean	1988
29	PV0418	USA, DSMZ	pepper	na
30	PV0419	USA, DSMZ	pepper	na
31	PV0420	USA, DSMZ	pepper	na
32	PR	Indonesia, GTZ	chili	2002
33	PW	Indonesia, GTZ	chili	2002
34	KRS	Indonesia, GTZ	chili	2002
35	BRB	Indonesia, GTZ	chili	2002
36	CA	Indonesia, GTZ	chili	2002
37	A4	India, GTZ	chili	2002
38	A1	India, GTZ	chili	2002

[1]The isolates with RTXXX designation numbers are the same as those used in Anonymous (1998); codes in the form PVXXX are the DSMZ numbers. [2]na: information not available

2.1.3 Chemicals

The chemicals and enzymes were bought from the following companies:

Duchefa (Haarlem, the Netherlands)
MBI Fermentas (St. Leon-Rot, Germany)
Promega (Mannheim, Germany)
Merck (Darmstadt, Germany)
New England Biolab (Frankfurt am Main, Germany)
Roth (Karlsruhe, Germany)
Sigma (Münich, Germany)
Serva (Heidelberg, Germany)
All enzymes were used according to manufacturer's specification.
All solutions and reagents were prepared with double distilled water if not indicated otherwise.

2.1.4 Oligonucleotides (primers and probes)

The primers for RT-PCR used in this study were synthesized by MWG Biotech (Ebersberg, Germany). The sequences of the primers are shown in Table 2.

Table 2: Primers used for RT-PCR amplification

No.	name	sequence (from 5′ to 3′) [1]	position[4]	genomic segment	annealing temperature[5]
1	5′-RNA1 (Forward)	CGATGATGTTCGATGGTGC	723—741	RNA1	61℃
2	3′-RNA1 (Reverse)	TCTAGTGATGGGGGAACG	1846—1829		
3	5′-RNA2 (Forward)	GATGAATTCYTGTTTGCTCAC	2130—2150	RNA2	52℃
4	3′-RNA2 (Reverse)	GGATGGACAACCCGTTC	2864—2848		
5	5′-2a (Forward)	GTTTATTTACAAGAGCGTACG	1-21	RNA2	52℃
6	3′-2a (Reverse)	CTTCGTYGTACCTACTCTCT	1142-1123		
7	RNA2-NcoI (Forward)	ACCATYGTCA<u>CCATGG</u>CTGARTTTGC[2]	1842—1867	RNA2	61℃
8	RNA2-CpoI (Reverse)	GTACA<u>CGGACCG</u>RAGTCCTTCCG	2980—2958		
9	Fny209 (V/I) (Forward)	TTATCAA**ATA**GATGGTTCGGAGCTG[3]	2589—2613	RNA2	61℃
10	Fny209 (V/I) (Reverse)	CAGCTCCGAACCATC**TAT**TTGATAG	2613—2589	RNA2	61℃
11	Fny209 (T/A) (Forward)	CGTCGAACTCCA**G**CTGGCTC	2448—2467	RNA2	61℃
12	Fny209 (T/A) (Reverse)	GAGCCAG**C**TGGAGTTCGACG	2467—2448	RNA2	61℃
13	5′-MP (Forward)	ATGGCTTTCCAAGGTACC	120—137	RNA3	59℃
14	3′-MP (Reverse)	ACCGTTAACCACCTGCG	953—937		
15	5′-CP (Forward)	ATGGACAAATCTGRATCWMCC	1257—1277	RNA3	59℃
16	3′-CP (Reverse)	CTGGATGGACAACCCGTTC	2020—2012		

[1]M=A or C, R=A or G, Y=C or T and W=A or T; [2]Underlined sequences containing a restriction enzyme recognition site; [3]Bold letters indicate nucleotides for *in vitro* mutagenesis; [4]Nucleotide positions of the primers in the genomic RNAs of CMV Fny isolate (Rizzo *et al*., 1989); [5]Annealing temperature was calculated according to Thein and Wallace (1986).

Probes for microarray detection were synthesized by Microsynth (Switzerland). All probes have a 15-T spacer to improve the hybridization behaviour to glass slides and an amino-linker on the 3'end to bind the probes to the aldehyde-coated surface of glass slides. The sequences of the probes are given in Table 3.

Table 3: Serogroup specific probes for microarray detection

probe name	specificity	sequence(from 5´ to 3´)[1]	% GC[2]	Tm[3]
SG1-2r	I	GCG GAC GGA GCC TCA CCG GTA CTG $(T)_{15}$	70.8	68.1
SG1a2_1r	IA	CAG TCA CIG AAT ATG ATA AGA AGC$(T)_{15}$	37.5	49.7
SG1b2_1r	IB	CAG TCA CIG AGT TCG ATA AGA AGC$(T)_{15}$	45.8	54.1
SG2_3f	II	GTG GGA CGA CCA ATG GCG AGG GTT$(T)_{15}$	62.5	67.3
SG2_4f	II	GAG TCA AAG CAC GCA ACC CTG CAT$(T)_{15}$	54.2	63.1

[1]I= inosine; [2]%GC=percent bases that are either guanine or cytosine; [3]Tm=melting temperature (thermodynamic Tm based on nearest neighbour theory as implemented in Vector NTI Advance 9.0, InforMax, UK).

2.1.5 Antibodies and antisera

Polyclonal CMV antiserum AS-0475 and monoclonal CMV antibodies AS-0487, AS-0488, AS-0489, AS-0490, and AS-0491 were purchased from DSMZ. Monoclonal antibody AS-0489 is specific for serotype I isolates and AS-0490 for serotype II isolates. AS-0487, AS-0488 and AS-0490 and polyclonal antiserum AS-0475 do not differentiate between serogroups.

2.1.6 CMV full-length clone

Fny 109, Fny209 and Fny309, which are based on the phagemid pIBI76 (International Biotechnologies), contain a full-length copy of RNA1, RNA2 and RNA3 segment of CMV Fny isolate, respectively, and are described in Palukaitis *et al.* (1990). The clones were kindly provided by P. Palukaitis (Scottish Crop Research Institute, Dundee, Scotland). The detailed map of Fny 209 is presented in Appendix 7.1.

2.1.7 Media

All media were prepared according to Sambrook *et al*. (2001).

SOB-Medium (per litre)	20 g	Tryptone
	5 g	Yeast extract
	0.5 g	NaCl
	0.2 g	KCl
	pH 7.5	

SOC-Medium	20 mM	Glucose
	20 mM	$MgCl_2$
	in	SOB-Medium
LB-Medium (per litre)	10 g	Tryptone
	5 g	Yeast extract
	10 g	NaCl
LB-Agar (per litre)	15 g	Micro-agar
	in LB-Medium	
AIX-Agar (per litre)	150 mg	Ampicilline
	47 mg	IPTG
	40 mg	X-Gal
		(dissolved in 1ml Dimethylformamid)
		in LB-Agar

All media were autoclaved for 20 min at 121 °C. Glucose, $MgCl_2$, Ampicilline, IPTG and X-Gal were added after the temperature reached about 60 °C.

2.2 Methods

2.2.1 Plant cultivation

Chili (*Capsicum* spp.), *Vigna unguiculata* L. Walp., *Chenopodium quinoa* Willd., *Nicotiana glutinosa* L. and *N. benthamiana* Domin. were grown in a temperature-water-and light-controlled greenhouse at 20-25°C during day and night with a photoperiod of 16 h.

2.2.2 Plant hybridization

Two chili genotypes, VC246 and PBC370, were used as female or male parents for reciprocal hybridization. The stamens were removed by hand and pollinations were done in a greenhouse. After pollinations, the plants were separated. The F1 generation was used for resistance testing.

2.2.3 Purification of CMV particles

CMV particles were purified following the procedure originally described by Lot *et al.* (1972).

Extraction buffer	500 mM	Sodium citrate, pH 6.5
	5 mM	EDTA
	0.5% (v/v)	Thioglycolic acid
Virus buffer	5 mM	Boric acid, pH 9.0
	0.5 mM	EDTA
	2% (v/v)	Triton X-100

Infected leaves were homogenized in an equal volume of extraction buffer (w/v) and filtered through cheesecloth. The filtrate was clarified by addition of one volume pre-cooled chloroform (4°C), and centrifuged (3000 rpm, 4°C, 20 min, rotor HB-4, Sorvall). Virus was precipitated from the aqueous phase with 10% (w/v) PEG (MW 6000), gently stirring for 30-45 min at 0-4 °C and sedimented (11000 rpm, 4°C, 15 min, rotor SS34, Sorvall). The pellet was resuspended in 50 ml virus buffer and stirred for 30 min at 4°C before centrifugation (14500 rpm, 4°C, 15 min, rotor SS34). The supernate was centrifuged again at high speed (32000 rpm, 4°C, 3 hours, rotor Ti 70.1, Beckman) and the virus pellet was dissolved in ddH_2O.

The virus concentration was determined with the BCA Protein Assay Reagent kit (Pierce Biotechnology, Rockford, USA).

2.2.4 Plant inoculation with virus particles or viral RNA

Particle inoculation buffer (PIB)	0.02 M	NaH_2PO_4 / Na_2HPO_4, pH 7.0
	2 % (w/v)	PVP 15
	0.2 % (w/v)	Na_2SO_3
	10 mM	DIECA
RNA inoculation buffer (RIB)	0.05 M	PBS buffer, pH 8.0
	5 % (w/v)	Carborundum (600 mesh) autoclaved together

For plant inoculation with virus particles, infected plant material was placed in a precooled mortar and homogenized in PIB at 1:50 (w/v) for dried and 1:10 (w/v) for fresh leaf material, respectively. This suspension was rubbed with fingers onto plant leaves which had previously been dusted with carborundum (600 mesh). Following inoculation, the plants were rinsed with tap water and incubated in the greenhouse.

To obtain reassortants between different virus isolates, inoculation with purified viral RNA (2.2.7) was performed. RNA (5-10 µl) was rubbed by glove-covered fingures onto *C. quinoa* leaf at a concentration of 0.01-0.05 µg /µl in RIB.

For reverse genetic experiments, plants were inoculated with a mixture of 1-2 µg RNA derived from purified viral particles (2.2.3) and 20 µg *in vitro* transcripts (2.2.21). The mixture was diluted to 100 µl with RIB and rubbed by glove-covered figures onto *C. quinoa* leaves (5-10 µl inoculum per leaf).

To check the primary replication with plant leaflets, each leaf disk (1×1cm) was rubbed by using autoclaved brushes with 25 µl purified RNA (0.2 µg/µl in RIB) (2.2.3).

2.2.5 Inoculation by grafting

Grafting among different chilli lines was done by using inverted saddle grafts. The grafts were constructed by making a V-shaped notch in the stem of the rootstock and a corresponding V-shaped wedge in the stem of the scion. The cut surfaces were placed together and wrapped tightly with Parafilm (American National Can, USA). Grafted plants were covered with plastic bags for one week to prevent dehydration during healing of the grafts.

2.2.6 Silica-based plant RNA extraction

Total plant RNA was extracted according to Rott and Jelkmann (2001).

Grinding buffer (GB)	4.0 M	Guanadine thiocyanate
	0.2 M	Na-Acetate, pH 5.2
	25 mM	EDTA
	1.0 M	K-Acetate
	2.5 % (w/v)	PVP 40
	store at 4°C	
Washing buffer (WB)	10 mM	Tris-HCl, pH 7.5
	0.5 mM	EDTA
	5 mM	NaCl
	50 % (v/v)	Ethanol
	store at 4°C	
NaI	0.15 M	Na_2SO_3
	6 M	NaI
		store at 4°C in a dark bottle

Preparation of silica:
60 g silica particles (Sigma S 5631) were prepared by mixing with 500 ml water. The suspended particles were allowed to settle for 24 h. The upper 470 ml of the supernate was discarded, and the procedure was repeated by resuspending the sediment in 500 ml water and settling for another 5 h. The upper 440 ml of the supernate was removed and the remaining 60 ml slurry was adjusted to pH 2.0 with HCl, autoclaved and stored at 4℃ in aliquots of 200 µl.

Leaf tissue (300 mg) was homogenized in a plastic bag (Bioreba, Reinach, Switzerland) with 3 ml GB. 500 µl of homogenate was incubated with 100 µl 10% (w/v) N-Laurylsarcosyl at 70℃ for 10 min with intermittent shaking and subsequently placed on ice for 5 min. After centrifugation (13000 rpm, 10 min, rt), 300µl of the supernate was mixed with 150 µl Ethanol, 300 µl 6 M NaI and 25 µl of resuspended silica. The mixture was incubated at RT for 10 min with intermittent shaking before the silica was sedimented (6000 rpm, 1 min, rt). After discarding the supernate, the silica pellet was resuspended in 500 µl WB and sedimented again. The washing step was repeated once, and the pellet was finally allowed to dry for several minutes at room temperature before resuspension in 150 µl water. Following incubation at 70℃ for 4 min, the silica was sedimented (13000 rpm, 3 min, rt), and the supernate was transferred to a fresh reaction tube for storage at -20℃.

2.2.7 Phenol extraction for DNA/RNA purification

An equal volume of phenol (TE-saturated, pH 7.5-8.0, Roth, warmed up to rt) was added to an aqueous DNA/RNA sample, vigorously mixed and centrifuged for phase separation (13000 rpm, 5 min, rt). The upper aqueous layer was transferred to a new reaction tube and extracted twice with an equal volume of Chloroform/Isoamylalcohol (24:1, v/v), following centrifugation to allow phase separation (13000 rpm, 5 min, rt). The DNA/RNA from the upper aqueous layer was concentrated by Ethanol precipitation (2.2.8).

2.2.8 Ethanol precipitation of DNA/RNA

TE buffer	10 mM	Tris-HCl, pH 8.0
	1 mM	EDTA

Cold Ethanol (2.5 - 3 volumes) and one tenth volume of Sodium acetate (pH 4.8) were added to the DNA/RNA solution and incubated at -80°C for at least 30 min or at -20°C overnight. The precipitated DNA was recovered by centrifugation at 15300 rpm for 30 min at 4°C. The supernate was discarded and the nucleic acid pellet was washed with cold 70% (v/v) ethanol for 5 min on ice with intermittent shaking. After centrifugation at 15300 rpm for 15 min, the supernate was discarded and the washing step repeated once. The DNA/RNA pellet was dried in a Speed-Vac concentrator (Savant Instruments Inc., USA) and resuspended in water or TE buffer.

2.2.9 Determination of DNA and RNA concentration

DNA or RNA concentration was determined photometrically. The DNA or RNA sample was diluted 1:100 with ddH_2O. The absorbance of solution was measured at 260 and 280 nm, using water as blank. An OD_{260} of 1 corresponds to a DNA or RNA concentration of 50 (DNA) or 40 (RNA) µg/ml. The ratio 260/280 provides an indication of the purity of the DNA/RNA. The value should be between 2.0 and 2.2.

2.2.10 Agarose-gel electrophoresis

2.2.10.1 Native DNA electrophoresis

TAE-buffer	0.04 M	Tris-Acetate pH 8.3
	1 mM	EDTA
Loading buffer	50 % (v/v)	Glycerol
	0.1 % (w/v)	Bromphenol Blue

DNA was separated using 0.8 to 2.0 % (w/v) agarose gels in TAE buffer containing Ethidium bromide (0.2 µg/ml) with 4 V/cm and examined by UV light at 254 nm, using a transilluminator (Kappa-Messtechnik, Germany). Gels were photographed to record results.

2.2.10.2 Denaturing RNA electrophoresis

MOPS buffer	200 mM	MOPS, pH 7.0
	50 mM	Na-Acetate
	5 mM	EDTA

Loading buffer	50 % (v/v)	Glycerol
	1 mM	EDTA
	0.4 % (w/v)	Bromphenol Blue
	0.1 % (w/v)	Ethidium bromide (EtBr)
Sample buffer	2 ml	MOPS buffer
	3.5 ml	37 % (v/v) Formaldehyde
	10 ml	deionized Formamide

RNA was separated using 1 to 2 % (w/v) agarose gels in TAE buffer. Two to four µg RNA was mixed with 10 µl Sample buffer and 3 µl loading buffer. The RNA in the mixture was denatured at 70℃ for 10 min and cooled down on ice. The samples were processed as described in 2.2.10.1.

The *in vitro* transcripts were examined by mixing with 10 µl sample buffer, 3 µl loading buffer (without EtBr). The mixture was denatured at 70℃ for 10 min and cooled down on ice. The samples were processed as described in 2.2.10.1 and Gel-Star (Cambrex Bio Science Rockland, Inc., USA) was added at a final concentration of 1 ‰ (v/v) to the gel.

2.2.11 Reverse transcription (RT) and Polymerase chain reaction (PCR)

2.2.11.1 cDNA synthesis (RT)

5 × M-MuLV buffer	250 mM	Tris-HCl, pH 8.3 at 25℃
	250 mM	KCl
	20 mM	$MgCl_2$
	50 mM	DTT

Total-RNA (0.05-0.5 µg) was denatured at 95 °C for 5 min with 10 µM reverse primer in a total volume of 10 µl and subsequently cooled down on ice to avoid renaturation. cDNA was synthesized by incubation at 42 °C for 45-60 min with the following reagents:

5 µl	5 × MMLV-buffer
2 µl	dNTPs (2mM)
0.1 µl	M-MuLV Reverse transcriptase (200 U/µl)
7.9 µl	H_2O

2.2.11.2 Polymerase chain reaction (PCR)

10 × PCR buffer	200 mM	Tris-HCl, pH 8.75 at 25℃
	100 mM	KCl
	100 mM	$(NH_4)_2SO_4$
	1% (v/v)	Triton X-100

The standard PCR was assembled according to the following conditions unless otherwise indicated:

1-2 µl	cDNA or any other template
2.5 µl	10 × PCR buffer
1.5 µl	$MgCl_2$ (25 mM)
2 µl	dNTPs (2 mM)
2 µl	primer forward (10 µM)
2 µl	primer reverse (10 µM)
0.2 µl	Taq DNA-Polymerase (5 U/µl)
ad 20 µl	H_2O

The PCR was carried out in a Personal Cycler 48 (Biometra, Germany) with the appropriate primers (2.1.4, Table 2):

1	initial denaturation	5 min	95 °C
2	denaturation	30 sec	95 °C
3	annealing	45 sec	50-79 °C
4	elongation	1 min	72 °C
5	final elongation	5 min	72 °C

The steps 2-4 were repeated 29 times.

An aliquot of the PCR products was analyzed by agarose gel electrophoresis (2.2.10.1).

2.2.11.3 Single-tube PCR

10 × IC-PCR buffer	100 mM	Tris-HCl, pH 8.8 at 25℃
	500 mM	KCl
	15 mM	$MgCl_2$
	1% (v/v)	Triton X-100

RT-PCR was carried out in one tube with:

1-1.5µl	total RNA (2.2.6)
2.5 µl	10 × IC-PCR buffer
3 µl	1.7 % (v/v) Triton X-100
2 µl	dNTPs (2 mM)
1 µl	primer forward (10 µM)
1 µl	primer reverse (10 µM)
0.2 µl	Taq DNA-Polymerase (5 U/µl)
0.5 µl	M-MuLV Reverse transcriptase (200 U/µl)
ad 25 µl	H_2O

Synthesis was carried out according to the conditions indicated below.

1	reverse transcription	45min	42 °C
2	initial denaturation	2 min	92 °C
3	denaturation	30 sec	92 °C
4	annealing	45 sec	59 °C
5	elongation	1 min	72 °C
6	final elongation	5 min	72 °C

Steps 3 to 5 were repeated 39 times.

An aliquot of the PCR product was analyzed by agarose gel electrophoresis (2.2.10.1).

2.2.11.4 Immuno-Capture Reverse Transcriptase Polymerase Chain Reaction (IC-RT-PCR)

Polypropylene PCR tubes were coated with 100 µl of purified IgG of AS-0475 (1 mg/ml) in coating buffer (2.2.23), and incubated at 37 °C for 2 h or 4°C overnight. After washing three times with 10 mM Tris-HCl/0.05 % (v/v) Tween-20 (pH 8.0), 100 µl homogenized plant sample (in 10 mM Tris-HCl, pH 8.0, at a 1:10 dilution) was added and incubated at 37 °C for 2 h or 4 °C overnight. The tubes were washed three times as described above and dried briefly by air. PCR was carried out, as described for single-tube PCR (2.2.11.3).

2.2.12 Clone screening by PCR

A single clone (2.2.19) was picked up by a toothpick and dissolved in 50 µl water. One µl was used as a template for the PCR reaction (2.2.11.2).

2.2.13 PCR-based site-directed mutagenesis

To introduce point mutations into the Fny209 full-length clone (2.1.6), a PCR-based, site-directed mutagenesis was carried out according to the procedure of Higuchi *et al*. (1988).

The first PCR was carried out with primer pairs A and C/Reverse or primer pairs B and C/Forward, respectively. Primer C/Reverse and primer C/Forward are two complementary primers, which contained a single nucleic acid mutation. The two PCR fragments were purified by agarose gel electrophoresis (2.2.16) to remove the template and primers from the first PCR. In a final PCR the mutated fragment was amplified from a mixture (1:1) of both purified fragments using primer pairs A and B (Fig. 1).

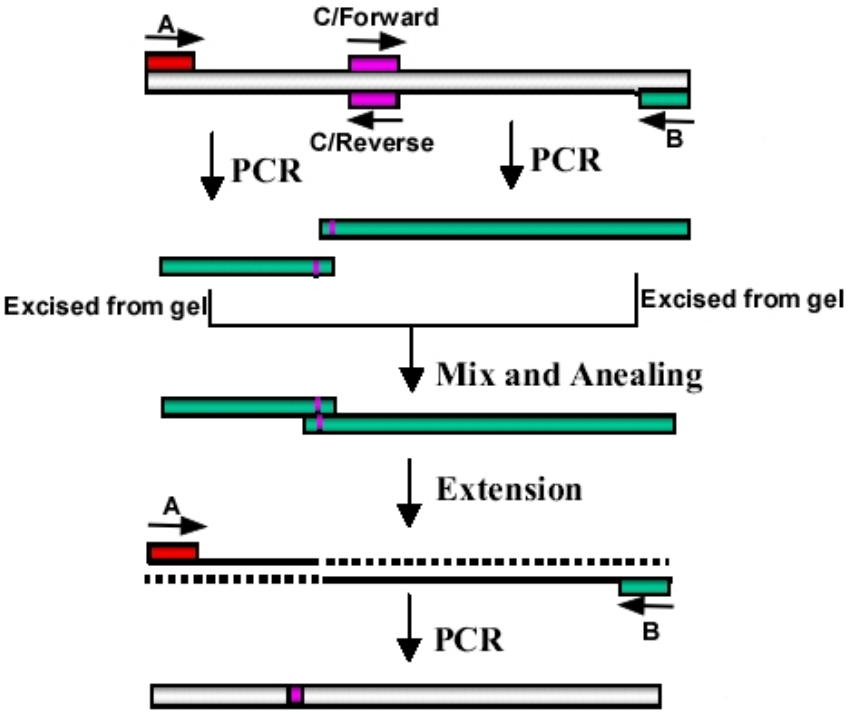


Fig. 2: Schematic of PCR-based mutagenesis
The first PCR was carried out with primer pairs A and C/Reverse or primer pairs B and C/Forward, respectively. Primer C/Reverse and primer C/Forward are two complementary primer, which contained a single nucleic acid mutation. The two PCR fragments were purified with agarose gel electrophoresis and excised from the gel. In a final PCR the mutated fragment was amplified from a mixture (1:1) from both fragments using primer pairs A and B.

2.2.14 Restriction enzyme digestion and restriction fragment length polymorphism (RFLP) analysis

Purified DNA fragment or plasmid was digested with the appropriate restriction enzyme according to manufacturer's recommendation.

The PCR product was directly digested with restriction enzymes without further purification. 15 µl of the PCR product was digested with 1-2 U restriction enzyme overnight.

2.2.15 PCR product purification

PCR products were purified using a E.Z.N.A. Cycle-pure Kit (PEQLAB Biotechnologie GMBH, Erlangen, Germany). DNA was eluted from the column with 40 µl H_2O.

2.2.16 DNA fragment purification from agarose gel

Digested DNA fragments (2.2.14) were separated on agarose gel (2.2.10.1). The fragment of interest was excised from the gel with a razor blade under UV light and purified with the E.Z.N.A. Gel Extraction Kit (PEQLAB Biotechnologie GmbH, Erlangen, Germany). DNA was eluted from the column with 40 µl H_2O.

2.2.17 Preparation of the T-vector

The T-vector for cloning of PCR products (2.2.11) was prepared according to the procedure of Marchuk *et al*. (1991).

The pBluescript SK- plasmid (Stratagene, USA) was linearized with EcoRV (2.2.14), followed by phenol/chloroform extraction (2.2.7) and ethanol precipitation (2.2.8).
The linearized vector (5 µg) was resuspended in 8 µl water. A T-overhang was added to the termini by the Terminal deoxynucleotidyl Transferase (TdT), using the following reagents:

15 µl	linearized pBluescript SK^- (5 µg)
8 µl	5×Tailing buffer (MBI Fermentas)
1 µl	1mM ddTTP
3 µl	5mM $CoCl_2$
5 µl	TdT (25U/µl, MBI Fermentas)

This was followed by incubation for 1h at 37℃.
The vector was extracted with phenol/chloroform (2.2.7), followed by ethanol precipitation (2.2.8) and diluted in water to a 20 ng/µl concentration for the ligation reaction (2.2.18).

2.2.18 Ligation

A 1:3 ratio of vector: DNA fragment (2.2.15 to 2.2.16) was generally used for the ligation reaction.

2 µl	pBluescript SK^--T vector (2.2.17) or other linearized vector (20 ng/ µl)
2-4 µl	purified DNA fragment (～120 ng)
1 µl	10×Ligation buffer (MBI Fermentas)
1-2 µl	T4-DNA Ligase (1U/ µl, MBI Fermentas) 1 µl for cohesive ends, 2 µl for blunt ends
ad 10 µl	H_2O

incubated at 14℃ overnight or 4h at rt.

For self-ligation or blunt end ligation, 1 µl of 50% PEG (MW 4000) solution was added.

2.2.19 Preparation of competent cells and chemical transformation

TFB Ⅰ	30 mM	K-Acetate
	50 mM	$MnCl_2$
	100mM	RbCl
	10 mM	$CaCl_2$
	15% (v/v)	Glycerol
	adjust the pH to 5.8 with 0.2 N HAc	
TFB Ⅱ	10 mM	MOPS, pH 7.0
	10mM	RbCl
	15 mM	$CaCl_2$
	15% (v/v)	Glycerol

Both solutions were sterilized by filtration (0.45 µm) and stored at 4 °C.

All steps were performed on ice with chilled solutions.

An *E. coli* (NM522) stock grown overnight was propagated in 100 ml LB-medium (1:100 dilution, 2.1.7, containing 0.02 M Mg^{2+}) to a density of OD_{550}=0.48~0.52. The bacteria were sedimented by centrifugation (2000 rpm, 12 min, 4°C, rotor 2K15, Sigma) after a 10 min incubation period on ice. The bacteria were resuspended by swirling in 15 ml of TFB Ⅰ buffer, and left on ice for 10 min. After the cells were sedimented again, the pellets were resuspended immediately by swirling in 3 ml of TFB Ⅱ buffer and divided into 200 µl aliquots. The aliquots were snap-frozen in liquid nitrogen and stored at -70°C.

For transformation, the cells were thawed on ice (5~10 min) and the ligation product (2.2.18) was added. The mixture was incubated on ice for 30 min. The cells were shocked in a 42°C circulating water bath for exactly 180 sec and cooled down on ice for 1-2 min. SOC medium (2.1.7) (400 µl, pre-warmed to 37°C) was added, and the tubes were shaken (about 220 rpm) for approximately 1h at 37°C. Cells were spread at different volumes on LB-plates (2.1.7) with the appropriate antibiotics and incubated at 37°C for 12-14 h.

2.2.20 Plasmid isolation from bacteria

A bacterial culture grown overnight in the presence of the appropriate antibiotic was used for the purification of plasmid DNA. For fast screening of clones by PCR and restriction enzyme analysis, the boiling lysis method (2.2.20.1) was used (Sambrook *et al*., 2001); for further processing or manipulation of plasmid DNA, the miniprep method (2.2.20.2) was

chosen (Birnboim and Doly, 1979); for sequencing, the plasmid was isolated with the E.Z.N.A. Plasmid Miniprep Kit I (2.2.20.3).

2.2.20.1 Boiling Lysis method

STET	10 mM	Tris-HCl, pH 8.0
	100 mM	NaCl
	10 mM	EDTA
	5% (v/v)	Triton X-100

Bacteria from 1 ml of the overnight culture were sedimented for 1 min at 12000 rpm at rt before the supernate was removed by gentle aspiration. The pellets were resuspended in 80 µl STET containing 4 µl lysozyme (10 mg/ml in 10 mM Tris-HCl, pH 8.0) and boiled for 1 min. The bacterial lysate was centrifuged (12000 rpm, 10 min, rt) and the pellet was removed. The plasmids from the remaining supernate were precipitated with 80 µl of ice-cold isopropanol and collected by centrifugation (12000 rpm, 5 min, 4°C). The pellet was washed twice with 70% (v/v) ethanol, dried and dissolved in 20 µl TE.

2.2.20.2 Minipreps

Solution A	25 mM	Tris-HCl, pH 8.0
	50 mM	Glucose
	10 mM	EDTA
Solution B	200 mM	NaOH
	1 % (w/v)	SDS
Solution C	3 M	Na-Acetate, pH 4.8
Solution D	50 mM	Tris-HCl, pH 8.3
	100 mM	Na-Acetate

Cells from the overnight bacterial culture (1.5 ml) were sedimented (12000 rpm, 5 min, rt), resuspended in 200 µl solution A, and incubated for 5 min at rt before 400 µl of solution B and 300 µl solution C were added. After incubation for at least 15 min on ice, the suspension was centrifuged for 10 min at 12000 rpm at rt. The supernate was transferred to a fresh tube and centrifuged again. Plasmids were precipitated from the supernate by adding 600 µl isopropanol and sedimented (14000 rpm, 10 min, rt). The pellet was dissolved in

200 µl solution D, precipated again with 400 µl 100 % (v/v) ethanol and sedimented (14000 rpm, 10 min, rt). The pellet was allowed to dry in a Speed-Vac concentrator (Savant Instruments Inc., USA) and resuspended in 50 µl H_2O containing RNase A (1mg/ml).

2.2.20.3 Plasmid preparation for sequencing

For sequencing, plasmid was isolated with the E.Z.N.A. Plasmid Miniprep Kit I (PEQLAB Biotechnologie GMBH, Erlangen, Germany). After elution from the columns with water, 2-2.5 µg plasmid was precipitated (2.2.8) and dried on a heating block at 50-55℃ before sequencing by MWG Biotech (Ebersberg, Germany). The plasmid was sequenced by both directions.

2.2.21 *In vitro* transcription

To prepare the template DNA for *in vitro* transcription, plasmid was linearized with Pst I (2.2.14). The transcription mixture (see below) was incubated at 37°C for 2 hours.

10 µg	RNase-free linearized template
10 µl	5 × Transcription Buffer (MBI-Fermentas)
10 µl	DTT (100 mM)
2 µl	RNasin (40 U/µl, MBI-Fermentas)
5 µl	NTPs without GTP (5 mM)
5 µl	GTP (0.5mM)
5 µl	m^7G Cap Analog (5 mM, Promega or New England Biolabs)
3 µl	T7 RNA-Polymerase (20 U/µl, MBI Fermentas)
ad 50 µl	H_2O

The amount of synthesized full-length RNA transcript was estimated by comparison with known amounts of purified CMV RNA in 2 % (w/v) agarose gels (2.2.10.2). The transcribed RNA was directly used for inoculation (2.2.4) without further purification.

2.2.22 Hybridization of labelled probes into a CMV-microarray

PCR products for hybridization were labelled by using 5'-cy3-labelled forward and reverse CP primers (Table 2, primers 15 and 16) in the PCR reaction (2.2.11.2). Amplified labelled PCR products were purified using the E.Z.N.A. Cycle-pure Kit (2.2.15). The purified PCR products labelled with cy3 was denatured for 10 min at 96 °C, cooled on ice for 3–5 min and then kept on ice until hybridization. After priming the slides with 2×SSC, hybridization was conducted on a Lucidea Slidepro (Amersham Biosciences, Switzerland) by inject-

ing 200 µl of hybridization mix (Bodrossy et al., 2003) containing 1 µg of the purified PCR product, 60 µl 20×SSC, 4 µl 50× Denhardt's solution (Sigma, Switzerland), 2 µl 10 % (w/v) SDS. The DNA was hybridized at 40 °C for 2 h, followed by three washes under low stringency at 42 °C (6 min in 2×SSC and 0.2 % (w/v) SDS, 2 min in 0.2×SSC and 0.2 % (w/v) SDS, and 2 min in 0.075×SSC). After hybridization the slides were analysed in a GenePix Personal 4100 A microarray scanner (Axon Instruments, USA) at a wavelength of 532 nm with a PMT (photomultiplier tube) gain between 800 and 1000 to avoid overexposure.

2.2.23 Double Antibody Sandwich (DAS) and Triple Antibody Sandwich (TAS) Enzyme-Linked Immunosorbent Assay (ELISA)

Coating buffer	15 mM	Na_2CO_3
	35 mM	$NaHCO_3$
	pH 9.6	
PBS-T	137.0 mM	NaCl
	2.7 mM	KCl
	8.1 mM	Na_2HPO_4
	1.5 mM	KH_2PO_4
	0.05 % (v/v)	Tween-20
	pH 7.4	
Sample buffer	2 % (w/v) in PBS-T	PVP 15
Conjugate buffer	0.2 % (w/v) in Sample Buffer	Ovalbumin
Substrate buffer	9.7 % (v/v) pH 9.8 with HCl	Diethanolamine

To perform DAS ELISA, the microtiter plates (Greiner, Germany) were coated with 100 µl IgG (AS-0475, 1 mg/ml diluted in coating buffer) at 37°C for 4 h before excess antibody was removed by washing three times with PBS-T. Leaves were homogenized 1:30 in sample buffer, from which 100 µl was added per well. After incubation overnight at 4 °C, plates were washed again with PBS-T and incubated at 37°C for 4 hours with anti-CMV IgG conjugated with alkaline phosphatase (1 mg/ml diluted in conjugate buffer). After a final washing step, p-Nitrophenylphosphate (1 mg/ml diluted in substrate buffer) was added to the wells and colour development was measured photometrically (Dynatech MR5000, USA) at 405 nm and 630 nm as reference, against buffer as a blank.

The same procedure was applied for the TAS-ELISA with the exception that the plates were incubated with monoclonal antibodies (2.1.5) at 37°C for 4 hours after the sample incubation step. The plates were washed again, and bound anti-CMV antibodies were detected by addition of rabbit-anti-mouse alkaline phosphatase-conjugated IgG (Sigma A-2429, 1:10000 in conjugate buffer). Following incubation for 4 hours at 37°C, colour development step was measured.

2.2.24 Tissue print immunoblots

	137.0 mM	NaCl
PBS-T	2.7 mM	KCl
	8.1 mM	Na_2HPO_4
	1.5 mM	KH_2PO_4
	0.05 % (v/v)	Tween-20
	pH 7.4	

Tissue print immunoblots were performed as described by Lin *et al.* (1990) with some modifications. The leaves of chili plants were detached, rolled into a tight core and cut at the middle with a new razor blade for each sample. In the study on the systemic spread of virus, root tissues of the plants were also subjected to the immuno-tissue blotting. The newly cut surface of leaf or root sections was pressed onto Nitrocellulose membrane (Protran®, Schleicher & Schuell GmbH, Dassel, Germany) to obtain tissue-print. The membranes were removed, dried, and incubated in blocking buffer (5 % (w/v) fat-free milk powder in PBS-T) for 30 min at rt. The membrane was incubated with anti-CMV polyclonal antibody from rabbit (AS-0475, 1 mg/ml in PBS-T with 1 % (w/v) fat-free milk powder) for 1-2 h. Unbound antibody was removed by washing with PBS-T. This was followed by incubation for 1-2 h with goat-anti-rabbit alkaline phosphatase-conjugated IgG (Sigma A-3937, 1:10000 in PBS-T with 1 % (w/v) fat-free milk powder). The membrane was washed with PBS-T and detection of virus was accomplished by Fast-red staining (2.2.25).

2.2.25 Chemical detection (Fast-red)

Fast Red-buffer	0.2 M	Tris-HCl, pH 8.0
	2 mM	$MgCl_2$
Fast Red staining solution 1	6 mg	Naphtol AS-MX-Phosphat-disodium salt in 15 ml H_2O
Fast Red staining solution 2	90 mg	Fast Red TR salt in 15 ml Fast Red buffer

Fast Red staining solution 1 and 2 were mixed immediately before staining. The membrane was developed at rt or overnight at 4 °C.

2.2.26 Multiple sequence alignments and phylogeny estimations

Sequences were aligned using CLUSTAL W with default parameters (Thompson *et al*., 1994). The alignment was corrected by hand (manually) in some inappropriate gaps that were not multiples of 3 nt in coding regions to maintain the alignment of the encoded amino acids. Phylogenetic relationships between the isolates were determined and trees were constructed by neighbour-joining (NJ) method (Saitou and Nei, 1987) where the nucleotide pair-wise genetic distances were corrected by using Kimura two-parameter methods (Kimura, 1980), including transitions and transversions unweighted and handling gaps by pairwise deletion. Phylogenetic analyses were implemented with MEGA, Version 2.1 (Kumar et al, 2001). The homologous regions of the genome of *Peanut stunt virus* (PSV) was used as outgroup for these analyses since PSV is the most distantly related taxon within the same taxonomic group. The significance of the internal branches was evaluated by using 1000 bootstrap replications. All branches with <70% bootstrap support were judged as inconclusive and were collapsed (Efron *et al.*, 1996; Hillis *et al.*, 1993). Branch lengths for all trees were normalized to 5% divergence, except for the 2b analysis, wherein branch lengths are compressed to represent twice as much divergence as in the other trees, in order to fit the figure into a similar space.

3 Results

3.1 Characterization of CMV isolates

3.1.1 Symptomatology

Thirty-eight isolates of CMV were collected in Asia, USA and Europe, mainly from chili as well as from other hosts. For biological characterization, CMV isolates (Table 1) were inoculated onto three systemic hosts (*N. glutinosa*, *N. benthamiana* and *N. tabacum* Samsum NN) and two local lesion hosts (*Chenopodium quinoa* Willd. and *Vigna unguiculata* L. Walp.) using four plants of each species (2.2.4). Symptoms were recorded over a 2-week period and at the end virus presence was confirmed by DAS-ELISA (2.2.23). All of the isolates induced uniform sized local lesions in *C. quinoa* and different sized ones in *V. unguiculata* depending on the serotype of isolates. Serotype I isolates induced large size brown local lesions, whereas the serotype II isolates caused small size local lesions.

No remarkable differences in symptoms among the isolates were observed on the systemic hosts *N. benthamiana* and *N. tabacum* Samsum NN. However, the isolates could be classified into four pathotypes based on their ability to infect and cause symptoms on *N. glutinosa*. All isolates were able to infect *N. glutinosa*, except five collected from Indonesia and two from India, which were not infectious on this host. Three parameters were evaluated at 14 d.p.i.: (a) plant height; (b) reduction of leaf size and (c) type and severity of leaf symptoms. Pathotype I isolates (e.g. RT88) induced mild symptoms including mosaic and mottling; plant height and leaf size were not affected. Pathotype II isolates (e.g. RT138) induced mosaic, mottling and curling of the leaves, plant height and leaf size were slightly reduced compared with non infected plants. Pathotype III isolates (e.g. PV0420) induced severe mosaic, mottling, yellowing and etched rings on leaves and plants were dwarfed. Pathotype IV isolates (e.g. RT55) induced severe mosaic, mottling and necrosis at 14 d.p.i.; the plants were severely dwarfed. The results, which are summarized in Table 4, show that the Asian and other CMV isolates vary in their biological diversity, especially in their ability to infect *N. glutinosa*.

Table 4: Symptoms on *N. glutinosa* induced by CMV isolates

pathotype/typical species	[1]No. of isolates	[2]symptoms on *N. glutinosa*		
		[3]general symptoms	plant height	reduction of leaf size
I /RT88	3	mm, m	-	- or +
II/RT138	16	mo, m,	-	+ or ++
III/PV0420	9	sm or e, y, d	++	++
IV/RT55	3	sm, d	+++	+++

[1] CMV isolates corresponding to respective pathotype are: Pathotype I: P522, CN05, RT88; Pathotype II: KB8, UB50, RS30, TR15, CH1, AN, Peet's, CN02, CN03, HN, 141N, RT6, RT54, RT68, RT138, RT144; Pathotype III: KS44, CM1, P3613, CN06, CN07, RT52, PV0418, PV0419, PV0420; Pathotype IV: CN04, RT55, RT67.[2]The experiments were repeated at least once. -=no reduction of plant height or leaf size; +=slight reduction of plant height or leaf size; ++=severe reduction of leaf size or plant dwarfing; +++= strong reduction of leaf or plant size, necrosis and plant stopped growing (almost died). [3]d=dwarfing; e=etched rings; sm=severe mosaic; m=mosaic; mm=mild mosaic; mo=mottling; n=necrosis; y=yellowing.

3.1.2 Serology

All CMV isolates could be grouped into the two known serotypes by testing them by TAS-ELISA (2.2.23) against a set of five monoclonal antibodies (Mabs) (Anonymous, 1998). Antibodies AS-0488 and AS-0491 produced the same reaction with all isolates. AS-0487 did not react with serotype I and reacted strongly with serotype II isolates, whereas AS-0490 did not react with serotype II and reacted weakly with serotype I isolates. On the other hand, AS-0489 reacted weakly with serotype I but strongly with serotype II isolates.

Only two European isolates (RT67 and RT68) and two American isolates (PV0418 and PV0420) were identified to belong to serotype II. None of Asian isolates were found to belong to serotype II. Another set of 135 fresh field samples from several Chinese chili production areas were also tested by TAS-ELISA with the above-mentioned antibodies. All isolates belonged to the serotype I.

3.1.3 RT-PCR-RFLP analysis

RFLP patterns were compared to investigate a possible relationship with serological data, geographical origin and original host.

Total RNA of infected *N. glutinosa* plants or of dried leaves from field samples was extracted (2.2.6) and RT-PCR (2.2.11) was conducted. With primers 5′-CP and 3′-CP (Table 2) a fragment of about 800 bp that comprised the complete coat protein (CP) ORF and a part of the 3′ untranslated region, and with primers 5′-MP and 3′-MP (Table 2) a fragment of about 800 bp that comprised the MP ORF excluding the six 3′ terminal nucleotides were amplified. Primers 5′-RNA 2 and 3′-RNA 2 (Table 2) were used for the amplification of a

fragment of approximate 700 bp that comprised the complete ORF of the 2b gene (2b) with approximate 300 nucleotides in the 5´ terminus of 2a ORF and approximate 110 nucleotides of the 3´ non-coding region. The obtained amplicons were separately digested with *Alu* I, *Mbo* I, *Dde* I, *Hinf* I, *Rsa* I, *Taq* I and *Tas* I (2.2.14) and the resulting RFLP patterns were evaluated by agarose gel electrophoresis (2.2.10.1).

The RFLP patterns of the CP gene of 38 isolates showed that the two serotypes I and II could be differentiated by the two enzymes *Mbo* I and *Taq* I (Fig. 3). The CP amplicon of both serotypes was digested into several bands by either *Mbo* I or by *Taq* I. However, serotype I and serotype II isolates could be differentiated by the largest band after digestion. The band of serotype II isolates was less than 400 bp, while that of serotype I isolates was more than 500 bp in size when the amplicons were digested with *Mbo* I. Likewise, the resulting largest band of serotype I isolates was around 560 bp whereas that of serotype II isolates was less than 300 bp in size when digested with the enzyme *Taq* I. Isolates of the two serotypes could not be differentiated by the other five enzymes. A further differentiation of serogroup I into subgroup Ia and Ib was not possible with the enzyme tested and no correlation with original hosts or geographical origin could be found.

RFLP analysis of the MP amplicon showed that the two serotypes could be distinguished by the enzyme *Mbo* I, but not by the other six enzymes. The amplicons of the MP gene of both serotypes were digested into several bands by enzyme *Mbo* I, but the largest band was significantly different. The resulting largest band of serotype I isolates was less than 700 bp, while the respective one of serotype II isolates was more than 800 bp in size. However, a correlation of RFLP patterns with country of origin, host or serological groups Ia and Ib, was not possible.

RFLP analysis of the 2b amplicon showed that the two serotypes I and II could be classified by the enzyme *Tas* I, but not by the other six enzymes. The resulting band of serotype I isolates was about 550 bp, while the one of serotype II isolates was about 160 bp and 200 bp. Again, as with MP and CP, a correlation of RFLP patterns with country of origin, host or serological subgroups Ia, Ib was not possible.

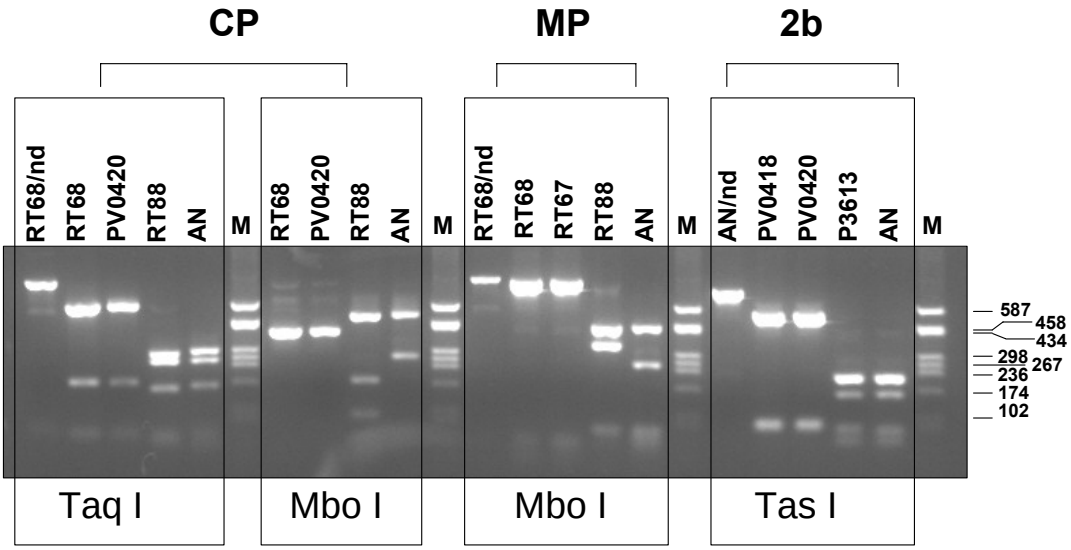


Fig. 3: Variability of amplicons expressed by RFLP patterns of the CP, MP and 2b genes.
Differentiation of serotype I and II was according to the largest band. As a control the CP fragment of isolate RT68 and AN was not digested (RT68/nd; AN/nd). Amplicon derived from the CP gene region was digested with *Taq* I and Mbo I, the amplicons derived from the MP gene region and 2b gene region were digested with *Mbo* I and *Tas* I, respectively. The isolates RT67, RT68, PV0418 and PV0420 belong to the serotype II, the isolates P3613, AN and RT88 to the serotype I. M= pUC9/HaeIII. The numbers on the right site give the size of the marker in bp. Gel was stained by EtBr.

3.1.4 Sequencing of CP, MP and 2b genes

The RT-PCR (3.1.3) fragments were purified (2.2.15) and transformed into *E. coli* NM522 (2.2.19). Five randomly selected positive clones from each transformation were checked for uniformity by RFLP-analysis with the enzymes *Alu* I, *Mbo* I, *Dde* I, *Hinf* I, *Rsa* I, *Taq* I and *Tas* I (3.1.3). If at least 4 colonies showed the same pattern, one of the four colonies was chosen for sequencing.

To establish a possible relationship between sequences and serological data (3.1.2), isolates, the representative isolates, from either different RFLP patterns (3.1.3) or different geographical origin, were selected for sequencing. A total of 28 CP genes, 30 MP genes as well as 2b genes of 8 isolates were sequenced (Table 5). All 66 sequences are deposited in the Appendix 7.2 .

Table 5: Designation, origin, length of CP, MP and 2b genes, biological data and taxonomical grouping of CMV isolates

[1]No.	[2]origin	isolate designation	sequence length (bp)			[3]pathotype	[4]serotype	[5]subgroup
			MP	CP	2b			
1	TH	KB8	834	657	336	2	I	IB
2		KS44	834	657	336	3	I	IB
3		UB50	834	/	/	2	I	IB
4		RS30	834	/	/	2	I	IB
5		TR15	834	657	/	2	I	IB
6		CH1	834	657	/	2	I	IB
7		CM1	/	/	/	3	I	/
8	IN	AN	837	657	336	2	I	IB
9	CN	P3613	834	657	336	3	I	IB
10		P522	834	657	333	1	I	IB
11		Peet's	831	657	/	2	I	IB
12		CN02	/	/	/	2	I	/
13		CN03	834	657	/	2	I	IB
14		CN04	834	657	336	4	I	IB
15		CN05	/	/	/	1	I	/
16		CN06	/	/	/	3	I	/
17		CN07	834	657	/	3	I	IB
18		HN	834	657	/	2	I	IB
19		141N	/	/	/	2	I	/
20	GR	RT6	834	657	/	2	I	IB
21	US	RT52	834	657	/	3	I	IA
22	CN	RT54	834	657	/	2	I	IB
23	CN	RT55	834	657	/	4	I	IB
24	ES	RT88	834	657	/	1	I	IA
25	NL	RT67	834	657	/	4	II	
26	NL	RT68	834	657	/	2	II	
27	IR	RT138	834	657	/	2	I	IB
28	BG	RT144	834	657	/	2	I	IB
29	US	PV0418	/	657	303	3	II	
30		PV0419	/	/	/	3	I	/
31		PV0420	/	657	303	3	II	
32	ID	PR	834	657	/	nt	I	IB
33		PW	834	/	/	nt	I	IB
34		KRS	834	/	/	nt	I	IB
35		BRB	834	657	/	nt	I	IB
36		Ca	834	657	/	nt	I	IB
37	IN	A4	837	657	/	nt	I	IB
38		A1	837	657	/	nt	I	IB

[1] The number corresponds to the respective numbers of Table 1. [2] Those isolates, which have a different MP gene length than most of the isolates, were underlined. The origin of isolates is indicated as follows: AU=Australia; CN=China; BG=Bulgaria; ES=Spain; GR= Greece; ID=Indonesia; IN=India; IR=Iran; NL=The Netherlands; TH=Thailand; US=United States. /= not sequenced. Fourteen CP sequences were deposited in the EMBL database and their accession numbers are shown in Table 7. All other sequences are shown in Appendix 7.2. [3]Pathotype is indicated based on reactions on *N. glutinosa*. nt=not tested (see also Table 4). [4]Sorting by serotypes according to the method of Anonymous (1998). [5]Sorting by subgroups according to Roossinck (2002).

The CP gene of all sequenced isolates was 657 nt, the 2b gene was 333 nt for subgroup I and 303 nts for subgroup II in length. In contrast, the length of the MP genes was variable. Most of the MP genes were 834 nts in length. However, the MP of isolates A1, A4 and AN was 837 nts, and that of isolates Peet's 831 nts in length, respectively. Computer-assisted sequence comparisons confirmed the patterns of the RFLP analyses (3.1.3).
Comparing the data shown in Table 5, no correlation between pathotype, serotype, subgroup and geographical origin could be made.

3.2 Phylogenetic analysis of CP, MP and 2b genes

The phylogenetic analysis of CP, MP or 2b gene sequences was carried out with the neighbour-joining method, respectively. Fifteen reported sequences were used as reference sequences and are described in Table 6.

Table 6: EMBL accession numbers of selected CMV isolates

isolate	accession no.		origin[1]
	RNA2	RNA3	
Fny	d00355	d10538	United States
IA	ab042293	ab042294	Indonesia
Ix	u20218	u20219	Philippines
Leg	d16406	d16405	Japan
LS	af416900	af127976	United States
Ly	af198102	af198103	Australia
Mf	aj276480	aj276481	South Korea
Nt9	d28779	d28780	Taiwan
O	d10209	d00385	Japan
Q	x00985	m21464	Australia
S	y10885	u37227 and af063610	South Africa
SD	d86330	ab008777	[China]
Tfn	y16925	y16926	Italy
Trk7	aj007934	l15336	Hungary
Y	d12538	l12499	Japan
ER-PSV	u15729	u15730	United States

[1] In case of isolate SD the country where the origin of sequence data was attained is indicated in brackets.

3.2.1 Analysis of the coat protein gene

The complete CP gene of 43 isolates including 15 reference isolates from the EMBL database (Table 6) were used to calculate the phylogenetic relationship (2.2.26). All of the isolates could be grouped into the two subgroups, Ia (red), Ib (blue) and serotype II (green) (Fig. 4). All of the Asian isolates characterized in this study belonged to the subgroup Ib.

However, two European isolates (RT6 and RT144), which were distributed into two separate clusters were also found in subgroup Ib. RT6 shows very close evolutionary ties to two subgroup Ib isolates, d28780 from China and y16926 from Italy. Likewise, RT144 shows very close evolutionary ties to ab008777 and RT54, two isolates of subgroup Ib from China.

The unrooted tree showed a small polytomy serotype II cluster and a large heterogeneous cluster, which contained the subgroups Ia and Ib. Close examination of the subgroup Ib cluster, to which most of the Asian isolates belong, revealed that within subgroup Ib there are several independent clusters with high bootstrap support as seen for isolates from China (designated as subgroup CN), from Indonesia (designated as ID) and from India and Thailand (designated as IT) (Fig. 4). To confirm the formation of these additional subgroups, a more detailed phylogenetic analysis with a larger number of sequence data (142 CP-sequences available until 5th December 2004 at the EMBL database) was carried out to construct a complex phylogenetic tree (data not shown). The topology of this tree was essentially identical with that in Fig. 4, however with more internal branching within the groups. In summary, the groups have identical relationship patterns to the complete analysis, and the Asian isolates formed three more independent clusters within subgroup Ib.

To find out any possible relationship between the CP gene and the original plant host, the sequences of CMV isolates, whose original hosts were known either from the EMBL database or from the literature, were used to construct another phylogenetic tree. The results showed that the CP gene had no relationship with most of the original hosts, except isolates from lily and squash. Two single conservative clusters with high bootstrap support in subgroup Ia including a lily isolate cluster could be defined, which was consistent for all 8 sequences of lily isolates reported up to date and a squash isolate cluster, which consisted of all the 7 squash isolates reported up to date (data not shown).

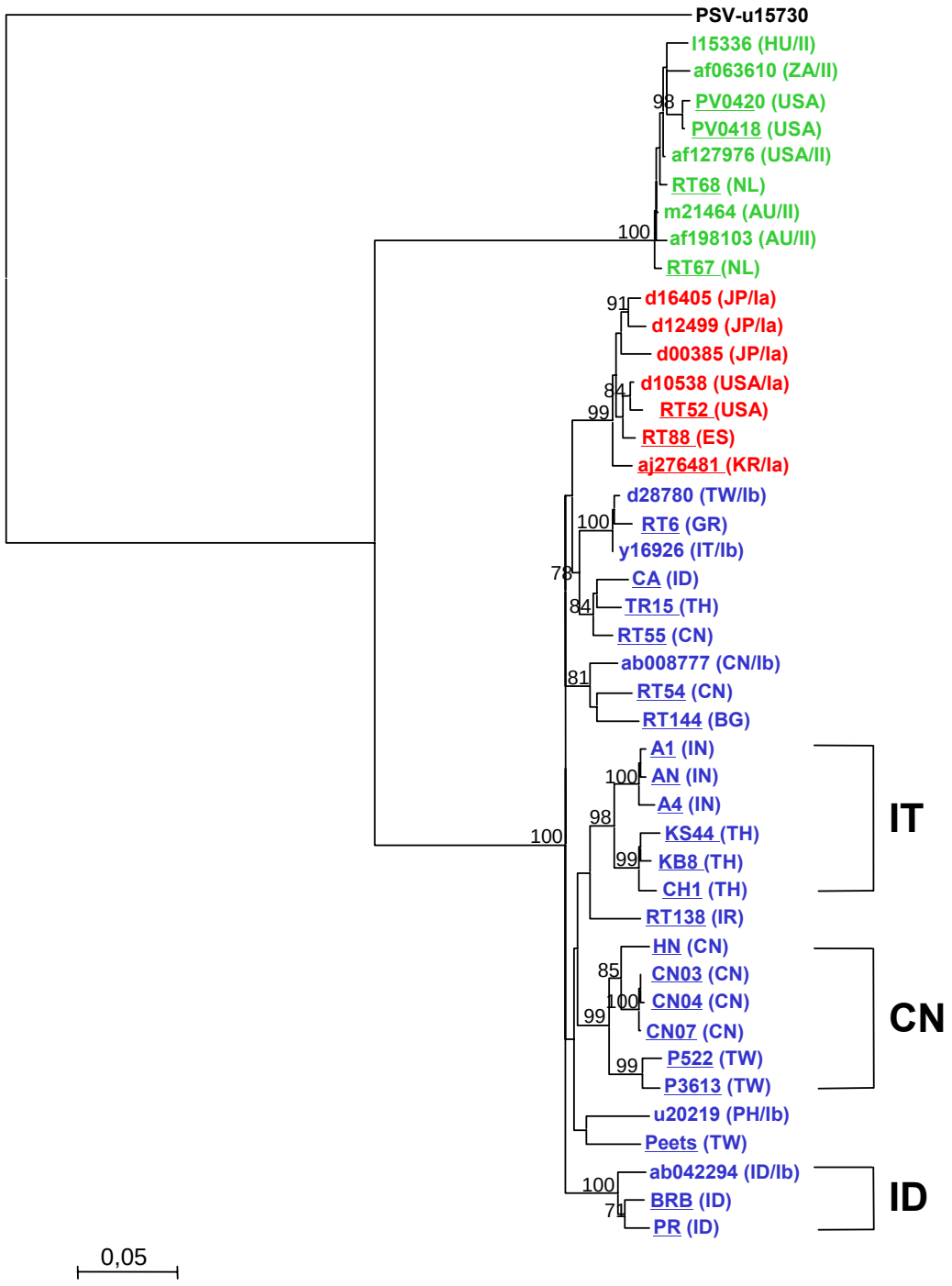


Fig. 4. Phylogenetic analysis of the CP ORF, constructed from 43 aligned nucleotide sequences.
Isolates highlighted in green are from serotype II, isolates highlighted in blue are from subgroup Ib, and isolates highlighted in red are from subgroup Ia. Bootstrap percentage values are shown above the branches. The subgroup classification of the reference sequences and the geographical origin of the isolates were indicated in brackets. The isolates sequenced in this study were underlined. Country codes are as follows: AU=Australia; CN=China; BG=Bulgaria; ES=Spain; GR=Greece; HU=Hungary; ID=Indonesia; IN=India; JP=Japan; IT=India+Thailand; KR=Korea; NL=The Netherlands; PH=Philippines; TH=Thailand; TW=Taiwan, China; USA=United States; ZA=South Africa. The tree is unrooted and the scale measures genetic distances in substitution per nucleotides. PSV was used as outgroup. See also material and methods (2.2.26).

3.2.2 Analysis of the movement protein gene

The almost full length sequence of the MP gene (3.1.3) of 46 CMV isolates including 15 reference isolates (Table 6) from the EMBL database were used to calculate the phylogenetic relationship (2.2.26). The nucleotide sequence divergence in the phylogenetic analysis of the MP gene showed a pattern almost identical to that of the CP ORFs. However, more clusters within the subgroup IB were present (Fig. 5).

Close examination of the subgroup Ib cluster revealed that there are five additional clusters within subgroup Ib, which are consistent with the geographical origin, with high bootstrap support as seen for isolates from Taiwan (designated as subgroup TW), from China (designated as subgroup CN), from Indonesia (designated as subgroup ID), from India (designated as subgroup IN) and Thailand (designated as subgroup TH) (Fig. 5).The detailed phylogenetic analysis with a set of 53 MP sequences (available at the EMBL database until 5th December 2004) confirmed the existence of five additional clusters (data not shown).

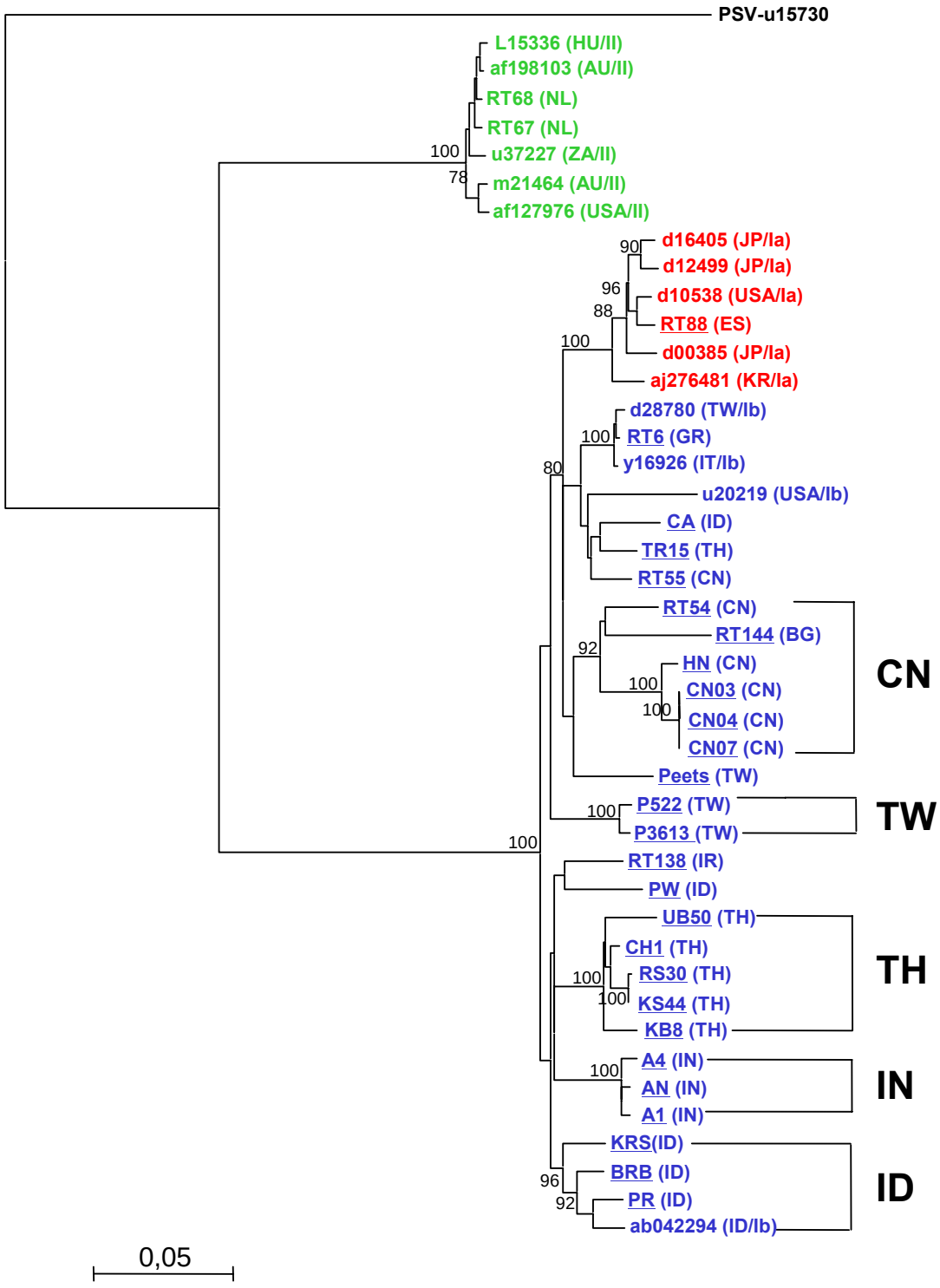


Fig. 5. Phylogenetic analysis of the MP ORF constructed from 46 aligned nucleotide sequences.
Isolates highlighted in green are from serotype II, isolates highlighted in blue are from subgroup Ib, and isolates highlighted in red are from subgroup Ia. Bootstrap percentage values are shown above the branches. The subgroup classification of the reference sequences and the geographical origin were indicated in brackets. The isolates sequenced in this study were underlined. Country codes, tree, outgroup and scale are the same as in Fig.4.

3.2.3 Analysis of 2b gene

The complete 2b ORF of 23 isolates including 15 reference isolates (Table 6) were also used to estimate the phylogenetic relationship (2.2.26). The unrooted tree showed a small polytomy serotype II cluster and a large heterogeneous cluster, which contains the subgroup Ia and Ib. Most of the Asian isolates were found in subgroup Ib. However, the isolates were grouped into different clusters as compared to those based on the CP and MP (Fig. 5). The four isolates AN, KS44 and KB8 and isolate IA (ab042293) formed a single cluster. The isolates CN04, P3613 and Ix (u20218) formed a second cluster, while isolates P522, NT9 and Tfn (y16925) fell into a third cluster. Another two isolates PV-0420 and PV-0418 were clearly grouped into serotype II.

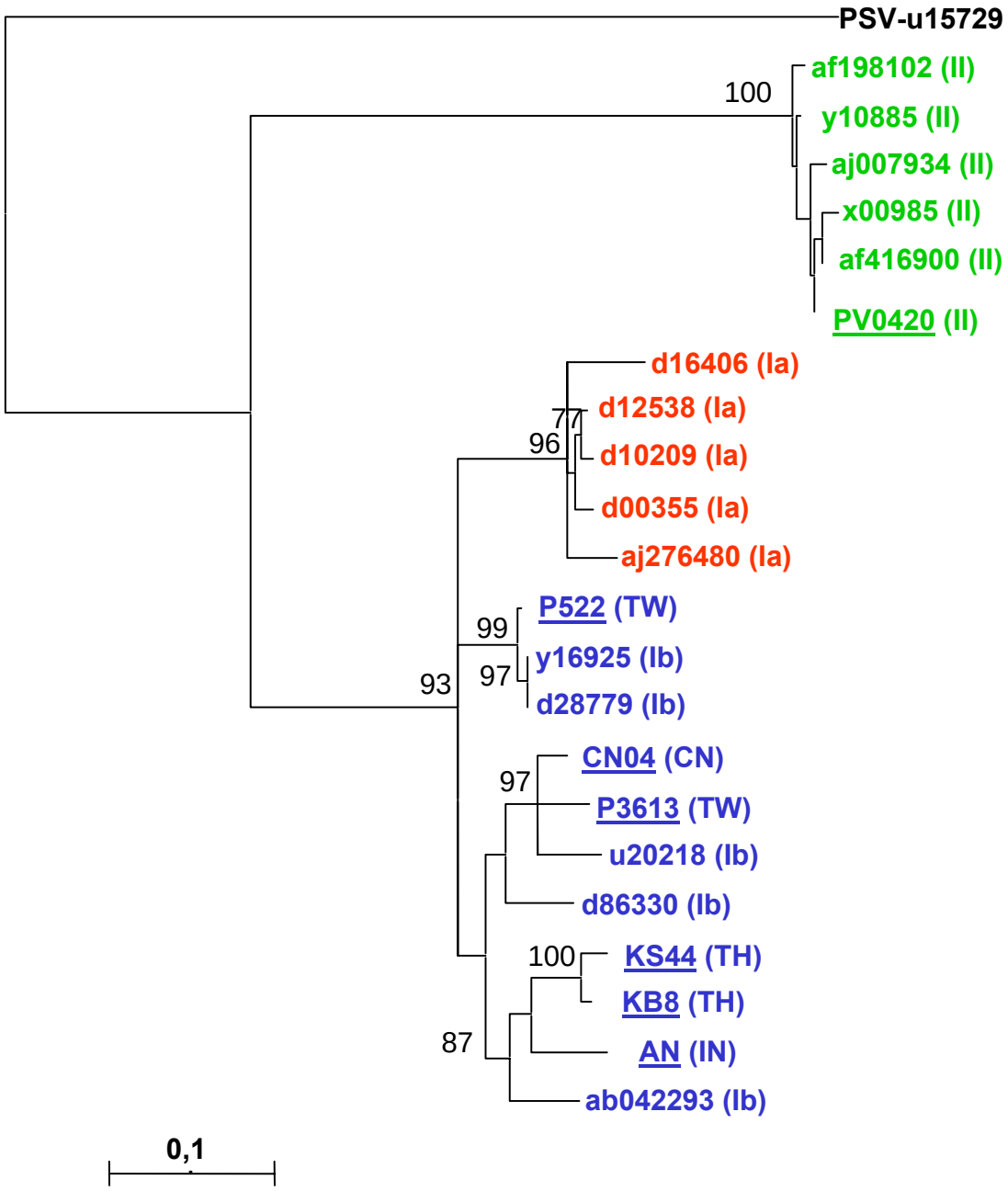


Fig. 6: Phylogenetic analysis of the 2b gene, constructed from 23 aligned nucleotide sequences.
Isolates highlighted in green are from serotype II, isolates highlighted in blue are from subgroup Ib, and isolates highlighted in red are from subgroup Ia. Bootstrap percentage values are shown above the branches. The subgroup classification of the reference sequences and the geographical origin of the isolates are indicated in brackets. The isolates sequenced in this study were underlined. Country codes, tree, outgroup and scale are the same as indicated in Fig. 4.

3.3 Differentiation of subgroups Ia, Ib and serotype II by an oligonucleotide-microarray

3.3.1 Selection of isolates

Fourteen isolates, four from serotype II, two from subgroup IA and eight from subgroup IB, were selected as representative isolates to develop an oligonucleotide- microarray detection method based on the phylogenetic analysis of the CP gene (Table 7, Fig. 4). These isolates represented the extensive heterogeneity of 28 CP gene sequences in this study.

Table 7: CMV isolates used for differentiation by an oliogonucleotide-microarray

[1]virus isolate	[2]serotype	[3]subgroup	accession No.
RT67	II		aj 810253
RT68	II		aj 810254
PV0420	II		aj 810255
PV0418	II		aj 810256
RT88	I	a	aj 810257
RT52	I	a	aj 810258
KS44	I	b	aj 810259
AN	I	b	aj 810260
CN03	I	b	aj 810261
RT144	I	b	aj 810262
RT54	I	b	aj 810263
TR15	I	b	aj 810264
RT6	I	b	aj 810265
P3613	I	b	aj 810266

[1] Codes starting with RT are the same as those used in Anonymous (1998); codes starting with PV#### are DSMZ numbers. The same codes are used in Fig. 4. [2] Sorting by serogroups according to Anonymous (1998). [3] Sorting by subgroups according to Roossinck (2002).

3.3.2 Design of the capture probes

Five serogroup-specific probes were designed, each 24 bp long and containing a 15 T spacer at the 3´ end to improve the hybridization behaviour (Table 3). At the very 3´ end an aminolinker was attached to bind the probes to aldehyde-coated glass slides (Genetix, UK). Probes were selected manually based on sequence alignments to optimize serogroup and subgroup differentiations in a hybridization reaction, i.e., mismatches were positioned towards the probe centre. Two probes targeted serotype II and three targeted serotype I, two of which were selectively designed to hybridize only subgroup Ia or Ib, respectively (Fig. 7). Because it was not possible to avoid a wobble position in the subgroup-specific probes, an inosine was included at the respective position to avoid signal intensity loss due

to dilution effects if using degenerate probes (Table 3, Fig. 7). All selected probes were tested for thermodynamic properties using the program Vector NTI Advance 9.0 (Infor-Max Inc., UK) and the best probes were used for the experiments. Because of the stringent constraints on the selection of subgroup-specific probes, their specifications were rather tolerant within 37.5–70.8 % GC content and a melting temperature (Tm) between 49.7 and 68.1°C (thermodynamic melting temperature calculation with Vector NTI Advanced 9.0, based on "Nearest Neighbour" theory, with constant parameters of DNA and salt concentration).

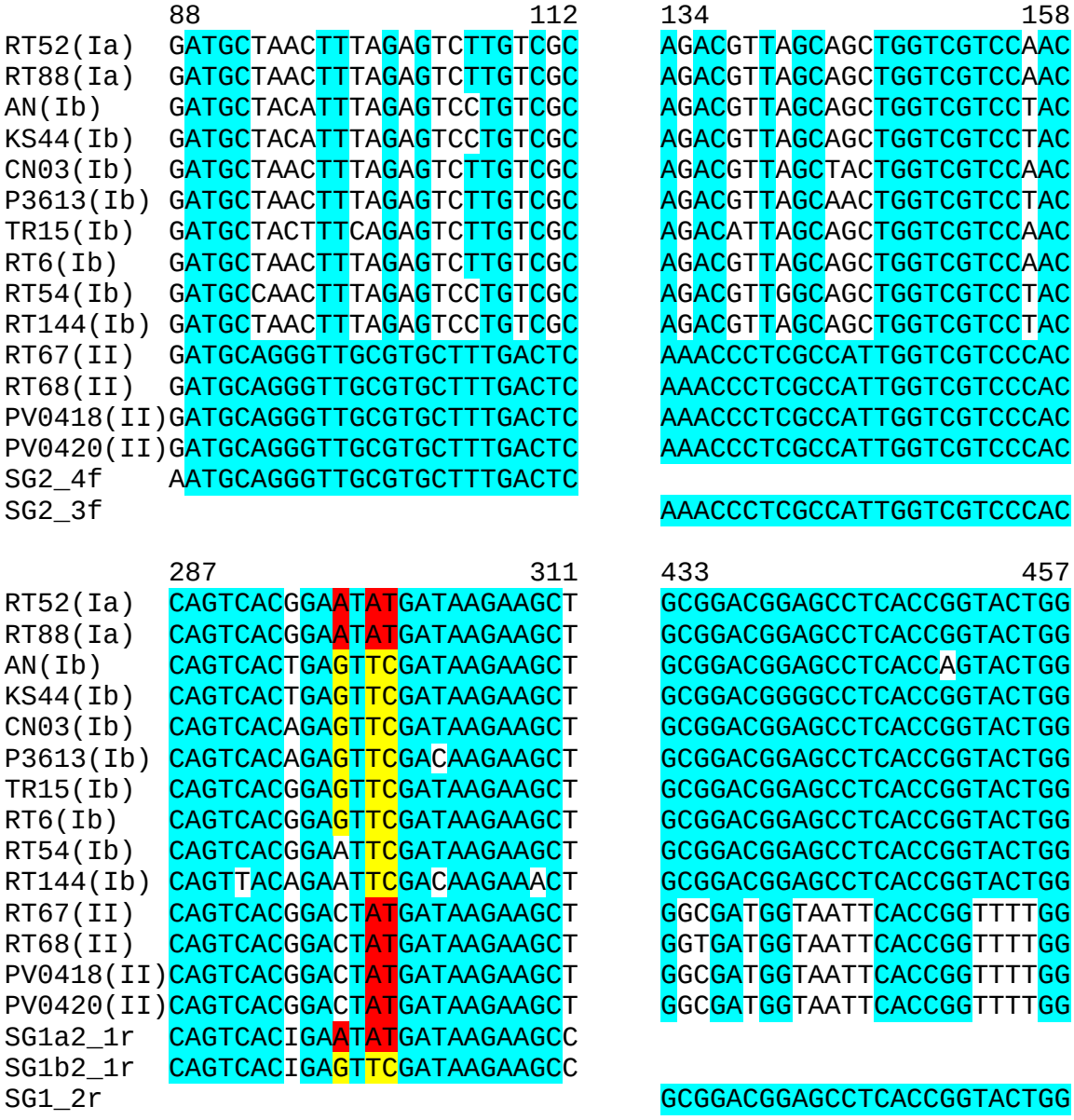


Fig. 7: The alignments of subgroup Ia, subgroup Ib and serotype II isolates for probe design.
Probe SG1_2r and SG2_3f, SG2_4f are specific for serotype I and serotype II isolates, respectively. Probe SG1a2_1r and SG1b2_1r are specific for subgroup Ia and subgroup Ib isolates.

3.3.3 Microarray printing, design and hybridization with labelled probes

Fifty microlitres of the 3´-6-C-aminoterminated probes were diluted 1:1 in Genetix Microarray Spotting Solution to a final concentration of 25 mM and spotted onto glass slides (Genetix, UK) using a QArrayMini automated spotter (Genetix, UK). Each slide included three replicates of each block (of which only one is shown in Fig. 8), with one block including two adjacent replicates of each probe, as well as a cy3-labelled amino-linked probe as standard. The spots are organised such that if positive hybridization in the left half is observed, the isolate belonged to serotype I, whereas if positive hybridization was observed in the right half of the array, the isolate belonged to serotype II (Fig. 7). Each spot was stamped three times and spotting was conducted in 75 % relative humidity. After spotting, the slides were washed following the protocol for aldehyde-coated slides of the manufacturer (Genetix, UK). Purified labelled PCR products were hybridized with arrays (2.2.22) and the array was further scanned.

The capture probes were manually designed based on an alignment of 170 different CP sequences obtained from the EMBL/GenBank or from our own sequencing data (Appendix 7.2). The probes targeted regions optimal for the differentiation between serogroup I and II or between subgroups 1a and 1b. As shown in Fig. 8, the serotype I and II specific probes worked perfectly and detected the respective virus isolates without any false positive result. The differentiation into the two subgroups of serotype I was possible with one exception. The two subgroup Ia isolates hybridized correctly, whereas one (RT144) of the eight subgroup Ib isolates, did not react with its subgroup specific capture probe. Nevertheless, the hybridization pattern of the isolate RT144 clearly allows correct allocation to serotype I.

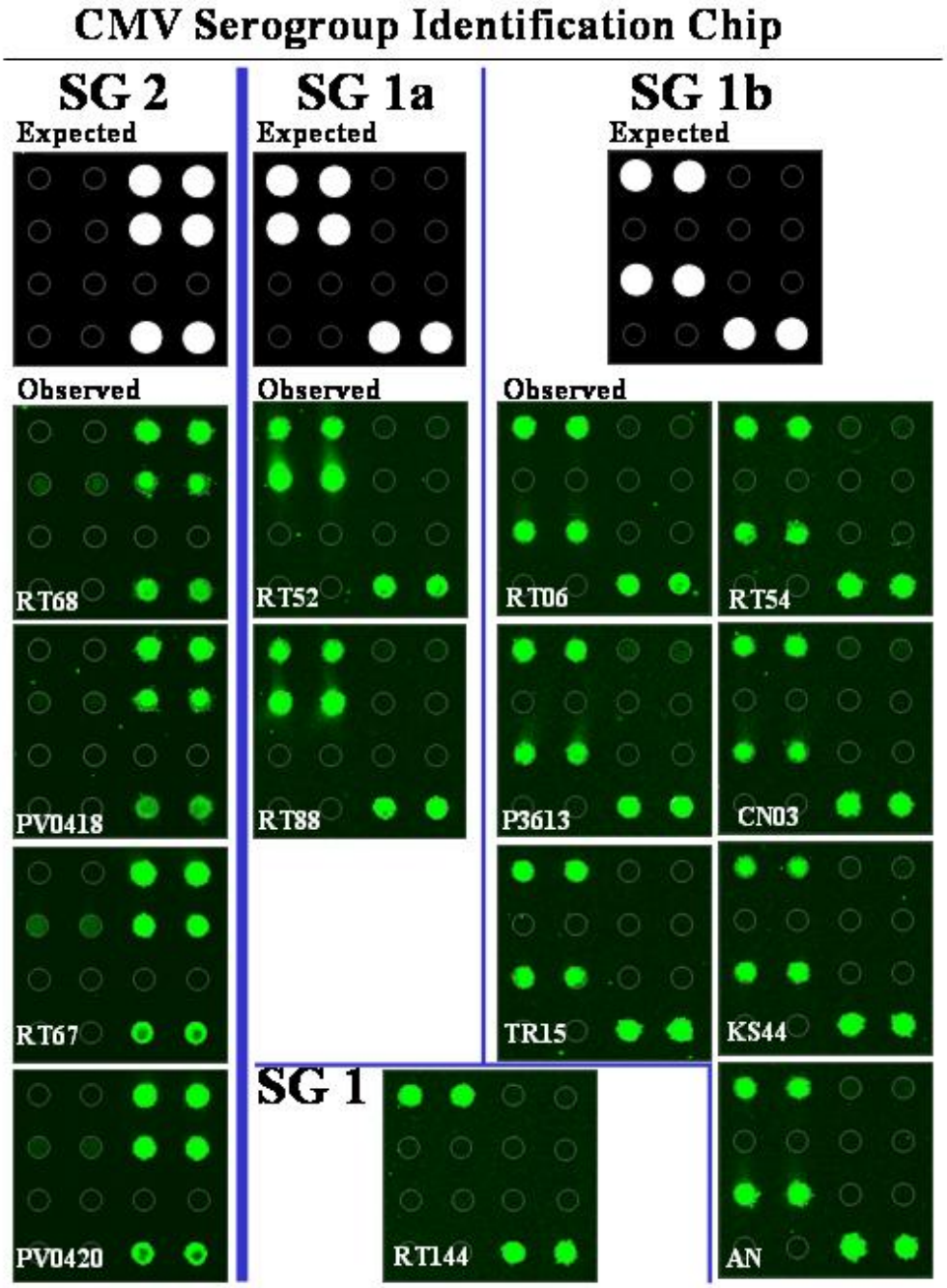


Fig. 8: Hybridisation results of 14 different CMV isolates on a two-dimensional oligonucleotide-microarray.
Cy3-labelled DNA amplicons were hybridised to oligonucleotide capture probes bound onto glass slides. Fluorescent patterns were recorded with a GenePix Personal 4100 A microarray scanner (Axon Instruments, US) at a wavelength of 532 nm.

3.4 Resistance screening

Twenty-seven CMV isolates (2.3.2) were inoculated to the resistant chili lines VC246, PBC370, PBC549 and PBC459 and to the susceptible line VC27a as a positive control (2.1.1). For each test three plants were inoculated (2.2.4) at the four to five leaf stage using either fresh or frozen with CMV infected *N. glutinosa* leaves as inoculum. The testing was repeated at least once. Twenty-one to 28 d.p.i. the upper, non inoculated leaves were tested by DAS-ELISA (2.2.23) and tissue print immunoblot (2.2.24) for evaluation of virus presence and distribution.

On the susceptible line VC27a, all isolates induced systemic symptoms on the leaves at seven d.p.i. followed by stunting. No symptoms were observed in lines PBC370, PBC549 and PBC495. However, 8 Asian serotype I isolates and two American isolates (serotype II) were able to replicate in these lines as proven by DAS-ELISA and RT-PCR (2.2.11). Tissue prints revealed that virus was present in the vascular tissue and roots (Fig. 9). Most of the isolates could not infect line VC246, which was confirmed by the absence of symptom expression and negative DAS-ELISA results. Only four chili isolates from different countries, KS44 and AN from serotype I and PV0418 and PV0420 from serotype II, could infect chili line VC246 and induced severe mosaic and reduced the size of the lamina at 14 d.p.i. (Fig. 9). These isolates were therefore defined to be resistance-breaking isolates on line VC246. The resistance reaction of different isolates on five chili lines is summarized in Table 8.

Table 8: Reaction patterns of different CMV isolates inoculated onto four different chili lines and a susceptible control line

[1]No	[2]designation	[3]chili lines				
		VC246	PBC370	PBC549	PBC495	VC27a
1	KB8	6 SS	6 S	6 S	6 S	6 SS
2	KS44	12 SS	11 S+1 R	12 S	12 S	12 SS
3	UB50	6 R	6 R	6 R	6 R	6 SS
4	RS30	6 R	6 R	6 R	6 R	6 SS
5	TR15	6 R	6 S	6 S	6 S	6 SS
6	CH1	6 R	6 S	6 S	6 S	6 SS
7	CM1	6 R	6 R	6 R	6 R	6 SS
8	AN	12 SS	12 S	12 S	12 S	12 SS
9	P3613	12 R	12 S	12 S	12 S	12 SS
10	P522	12 R	12 R	12 R	12 R	12 SS
11	Peet's	6 R	6 R	6 R	6 R	3 SS+ 3 R
13	CN03	6 R	6 R	6 R	6 R	6 SS
14	CN04	6 R	6 S	6 S	6 S	6 SS
17	CN07	6 R	6 R	6 R	6 R	6 SS
18	HN	6 R	6 R	6 R	6 R	6 SS
19	141N	6 R	6 S	6 S	6 S	6 SS
20	RT6	6 R	5 R+1S	6 R	6 R	6 SS
21	RT52	6 R	6 R	6 R	6 R	6 SS
22	RT54	6 R	6 R	6 R	6 R	6 SS
23	RT55	6 R	6 R	6 R	6 R	6 SS
24	RT67	6 R	6 R	6 R	6 R	6 SS
25	RT68	6 R	6 R	6 R	6 R	6 SS
26	RT88	6 R	6 R	6 R	6 R	6 SS
27	RT138	6 R	6 R	6 R	6 R	6 SS
28	RT144	6 R	6 R	6 R	6 R	6 SS
29	PV0418	12 SS	12 S	12 S	12 S	12 SS
30	PV0419	6 R	6 R	6 R	6 R	6 SS
31	PV0420	12 SS	12 S	12 S	12 S	12 SS

[1] The number corresponds to the respective numbers of Table 1. [2] Resistance-breaking isolates and the isolates that were used for reassortment studies (3.8, 3.11) are underlined. [3] VC27a is a susceptible control line. R=no visual symptoms and ELISA negative; S=no visual symptoms but ELISA positive; SS=visually severe mosaic and ELISA positive. 12 S=12 plants showed "S" phenotype. The isolates used for reassortment system are indicated with red boxes.

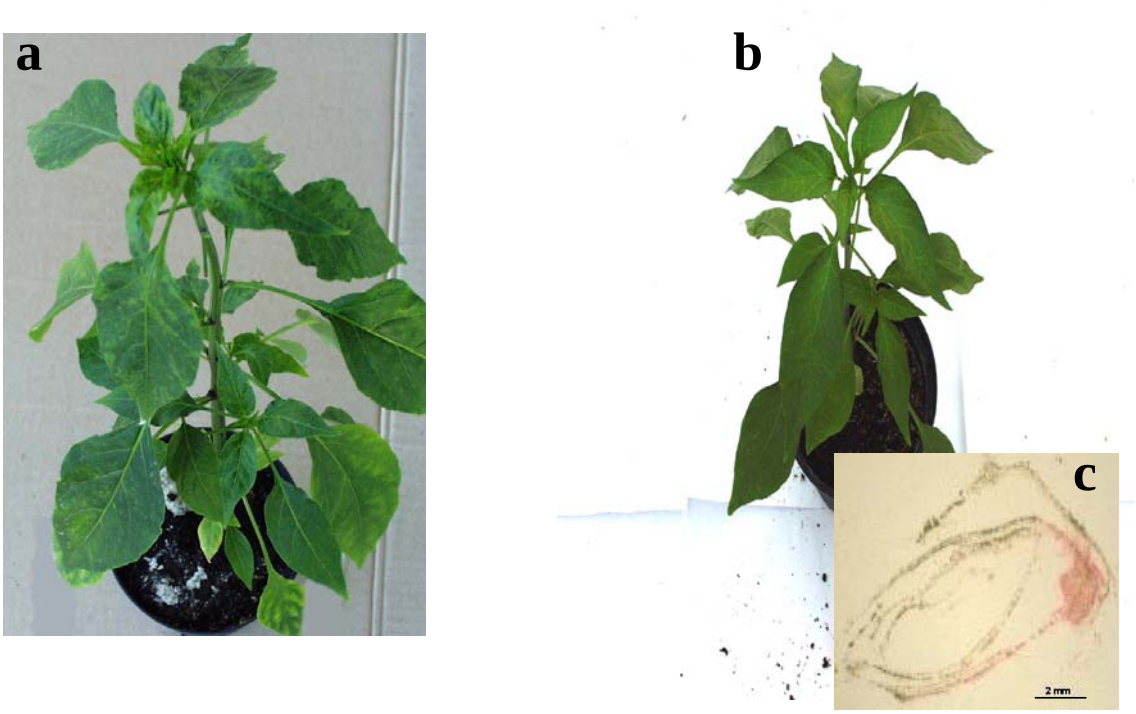


Fig. 9: Phenotypes of chili plants after inoculation and distribution of virus.
(a) and (b): Two typical phenotypes of chilli line / isolate combinations. (a) VC246 / AN: mosaic symptoms; (b) PBC370 / P3613: plants were symptomless, but virus was detected by DAS-ELISA. (c) Tissue-print analysis showed accumulation and spread of CMV in the uninoculated leaves 21 d.p.i. Chili plants were inoculated with CMV particles and tissue-print was carried out according to 2.2.24 with Fast-Red staining.

Although the four resistance-breaking isolates AN, KS44, PV0420 and PV0418 induced the same symptoms in line VC246, they caused different symptoms when inoculated onto *N. glutinosa* (Table 4). AN induced mild mosaic whereas KS44 induced severe mosaic, mottling, leaf malformation and reduction of the size of lamina. PV0418 and PV0420 induced ring etching, followed by leaf necrosis.

3.5 Replication of CMV in chili lines inoculated with RNA

To investigate the primary replication of different isolates (AN, KS44, P522 and P3613) in the three chili lines VC246, PBC370 and VC27a, leaflets (1x1cm) were excised from eight-leaf stage plants. RNA extracted from purified CMV particles (5 µg in 5-10 µl virus inoculation buffer, 2.2.4) was inoculated on leaf discs by gentle rubbing with a brush. The leaf discs were incubated in a moist environment at 25 °C under a 16 h light/8 h dark cycle. The samples were immediately frozen in liquid nitrogen at 0, 1, 3 and 6 d.p.i. and stored at

–20 °C until analysis. To verify presence of virus, the stored leaf disks were ground with 200 µl ELISA sample buffer (2.2.23) and DAS-ELISA was performed (2.2.23). Experiments were repeated twice. VC27a was used as a positive control line.

Virus was detected as early as at one d.p.i. by DAS-ELISA in leaf discs derived from the susceptible control line VC27a. However, no virus could be detected until 6 d.p.i in leaf discs of lines VC246 and PBC370 with any of the tested isolates.

Instead of leaf discs, plants of lines VC246 and VC27a were inoculated at the four to six leaf stage with purified viral RNA (10 µg/ leaf) derived from AN particles (2.2.3, 2.2.7). Mosaic symptom was observed in VC27a 3 d.p.i., but not in VC246. However, mosaic symptoms could occasionally be observed in VC246 (two out of 12 plants) at 14 d.p.i.

3.6 Symptom expression of chili lines inoculated by grafting

Since in some plant lines the virus was restricted to the vascular tissue (3.4) or the virus could not replicate in the original infected cells when inoculated mechanically, virus was "inoculated" directly into the vascular tissue by grafting (2.2.5).

Table 9: Grafting combinations and symptoms

No.	[1]rootstock, mechanically inoculated			scions, uninoculated		[4]symptoms after grafting in line VC246
	line	inoculated virus	[2]symptoms after mechanical inocuation	line	[3]symptoms after mechanical inocuation in line VC246	
1	VC27a	AN	SS	VC246	SS	SS
2	VC27a	P3613	SS	VC246	R	RS
3	PBC370	AN	S	VC246	SS	SS
4	PBC370	RS30	S	VC246	R	SS
		P3613	S	VC246	R	RS

[1]Each grafting combination was consisted by two different chili lines as rootstock or as scion. The rootstock was inoculated with virus before grafting. Line VC27a, which is susceptible to all CMV isolates, shows mosaic after mechanical inoculation. Line VC246 is resistant to isolate P3613 and RS30 but susceptible to AN. Line PBC370 is susceptible to isolates AN, P3613 and RS30, but plants show no symptoms after mechanical inoculation. See also Table 8 for detailed informations.[2]RS=ringspots; SS=plant was systemically infected and showed mosaic symptom; S=virus could infect this line, but plants showed no symptoms.[3]The symptoms present in VC 246 after mechanical inoculation with the same virus isolate used to inoculate the rootstock.[4]Symptoms at 21 days after grafting.

Each grafting combination consisted of two chilli lines, the non-resistant rootstock, which was inoculated with the virus at the 4-6-leaf stage before grafting, and the non inoculated resistant scion of line VC246. The grafting was carried out at the 12-16-leaf stage and all experiments were repeated once. The results were recorded 21 days after grafting. On the meantime, the presence of virus was confirmed by DAS-ELISA (2.2.23) or tissue print immunoassay (2.2.24).

Lines VC27a and PBC370 were used as rootstocks and the resistant line VC246 as scions. Isolate AN induced the same systemic mosaic in the grafted scion as in mechanically inoculated rootstocks (No. 1, Table 9). The same symptoms in VC246 were observed when AN inoculated rootstocks of line PBC370 instead of VC27a plants were used (Table 9, No. 3). Isolate P3613 could not infect line VC246 when mechanically inoculated. However, inoculation by grafting through an infected rootstock of line VC27a induced ringspots (Table 9, No. 2). The ringspots were observed in most of the old and newly emerging leaves, but some of the leaves remained symptomless and virus free. Neither isolate P3613 nor RS30 induced symptoms on PBC370 after mechanical inoculation but virus was detectable in the vascular tissue. On the other hand, no symptoms were observed and none of virus was detected in line VC246 after mechanical inoculation. After inoculation by grafting, isolate P3613 induced ringspots and isolate RS30 induced mosaic (Table 9, No. 4).

3.7 Characterization of resistance in hybrids of two different resistance lines

Two hybrid lines, designated as PV and VP, were produced. VP is the first-generation (F1) of a reciprocal hybrid line of a cross of VC246 (as a female parent) and PBC370 (as a male parent), while PV is the first-generation (F1) of a reciprocal hybrid line of PBC370 (as a female parent) and VC246 (as a male parent).

The two resulting hybrid lines VP and PV (2.2.2, Table 10) were inoculated (2.2.4) with the resistance breaking isolates PV0420, AN and KS44 (3.4).

Isolate AN induced strong mosaic symptom at 21 d.p.i on both hybrid lines and caused a reduction of lamina size identical to the symptoms induced in VC246. Interestingly, newly emerging leaves did not develop any symptoms and the symptoms in older leaves started to disappear at 28 d.p.i. Finally, those plants completely recovered, showing no symptoms at 42 d.p.i. However, tissue print immunoassays (2.2.24) of the newly emerging symptomless

leaves and of the old symptomless leaves revealed that virus was distributed evenly in the whole plant.

Isolate KS44 showed the same symptoms at 21 d.p.i. in both hybrid lines as described above for isolate AN. Compared with isolate AN, infection with KS44 did not result in any recovery. At 42 d.p.i. the whole plant still showed mosaic symptoms.

Although isolate PV0420 showed symptoms in line VC246 and virus was detected in the veins of line PBC370, none of symptoms were observed in the hybrid line PV. In three out of four inoculated plants, virus was detected in the veins at 21 d.p.i., meanwhile, there was no detectable virus in line VP. However, all the leaves of inoculated plants were symptomless and free of virus at 42 d.p.i., which were confirmed by RT-PCR (2.2.11) and DAS-ELISA (2.2.23).

Inoculation results of both hybrid lines with isolate P3613 were different from those obtained with chili line PBC370. Half of the tested plants showed mosaic and a reduction in size of the lamina, whereas the other half of the plants were symptomless and virus was detected by DAS-ELISA at 21 d.p.i. Plants with systemic mosaic started to recover at 28 d.p.i. and were symptomless at 42 d.p.i. However, tissue print immunoblots revealed that virus was still present in the veins of plants.

In total, three kinds of phenotypes were observed. (a) Mosaic in the early stage, and none of symptom when recovered, but virus was detectable (P3613, AN); (b) Mosaic in the whole life of plant, none of recovery (KS44); (c) Mosaic or no symptom in few plants in the early stage, but recovery and no detectable virus in the late stage (PV0420).

Table 10: Symptoms induced by several isolates in parental and hybrid lines

virus isolates	parental line		hybrid line[2]			
	VC246	PBC370	PV		VP	
			21 d.p.i	42 d.p.i	21 d.p.i	42 d.p.i
P3613	12 R	12 S	2 S +2 SS[1]	4 S	2 S +2 SS	4 S
AN	12 SS	12 S	4 SS	4 S	4 SS	4 S
KS44	12 SS	12 S	4 SS	4 SS	4 SS	4 SS
PV0420	12 SS	12 S	3 S+1 SS	4 R	4 R	4 R

[1] Code “2 S+2 R” indicated 2 of 4 tested plants showed “S” phenotype and the other 2 plants showed “R” phenotype. R=resistant, no symptoms and no virus detectable by ELISA or tissue print; S=induced no visual symptoms, but ELISA positive; SS=induced mosaic and reduction of the lamina size; /=not tested. [2] Two hybrid lines, which were designated as PV and VP, were produced. VP is the first-generation of a reciprocal hybrid line between VC246 (as a female parent) and PBC370 (as a male parent), while PV is the first-generation of a reciprocal hybrid line between PBC370 (as a female parent) and VC246 (as a male parent).

3.8 Reassortants between isolate P3613 and AN revealed that RNA 2 is responsible for resistance-breaking in chili line VC246

3.8.1 Establishment of genomic markers of CMV genome segments

The multipartite nature of CMV facilitates studies on the genetics of the virus because various reassortants are easily prepared by exchanging the respective genome RNAs between different isolate. Therefore, a reassortment system consisting of biological distinct isolates was constructed using genomic markers of CMV genome segments. Based on the distinct symptomatology on the chili line VC 246 (Table 8), the CMV isolates P3613 and AN were selected as parents for the generation of reassortants to identify which genomic RNA is responsible for resistance breaking. The PCR-RFLP analysis was used to establish the genetic markers for determining the parental origin of each genome segment in progeny reassortants. The specific regions on each genome segment were selected and amplified by RT-PCR. For the differentiation of RNA 1, RNA 2 and RNA 3, the primers 5´-RNA1 and 3´-RNA1, 5´-RNA2 and 3´-RNA2 and 5´-CP and 3´-CP, respectively, were used for RT-PCR (Table 2). Finally, the three amplicons of each single lesion were investigated by RFLP analysis (2.2.14). RNA 1 of AN resulted in two fragments of about 800 and 300 bp when digested with *Mbo* I, whereas RNA 1 of P3613 was not digested and showed a single band on the agarose gel, RNA 2 of AN resulted in two fragments of 430 and 300 bp upon digestion with *Pvu* II, whereas RNA 2 of P3613 was not digested. RNA 3 of P3613 was digested into two fragments of 190 and 580 bp whereas RNA 3 of AN was not digested with *Acl* I (Fig. 10).

3.8.2 Generation of reassortants

To obtain all six possible genome combinations (Appendix 7.4.1), extracted RNA (2.2.7) of AN and P3613 derived from purified particles (2.2.3), were mixed and inoculated (2.2.4) onto *C. quinoa*. Resulting single lesions with different shape and emerging periods were used separately for serial reinoculation of *C. quinoa*. After three or more passages the local lesions arose as uniform necrotic spots, indicating the presence of a distinct combination. Single lesions were used for a final inoculation of *N. glutinosa* before combinations were identified by RT-PCR-RFLP (3.8.1).

On the average about one third of local lesions from *C. quinoa* failed to infect *N. glutinosa*. In several cases the RFLP analyses revealed the presence of a mixture of both parental

RNAs. These combinations were used either for additional local lesion transfer on *C. quinoa* until only one parental RNA could be detected or were discarded. To increase the probability of generating the six theoretically possible reassortants of each combination, different ratios of the parental isolates were used in the initial coinfection.

In total, the three combinations [$1_P 2_A 3_A$], [$1_P 2_P 3_A$] and [$1_A 2_P 3_A$] (for detailed description, see legend of Fig. 10 and Appendix 7.4.1) were obtained, whereas the combinations [$1_P 2_A 3_P$], [$1_A 2_A 3_P$] and [$1_A 2_P 3_P$] were not achieved in three independent experiments. The frequency of detection of the parental and all possible reassortants genotypes was not a random event revealed by X^2 analysis (Appendix 7.4.1).

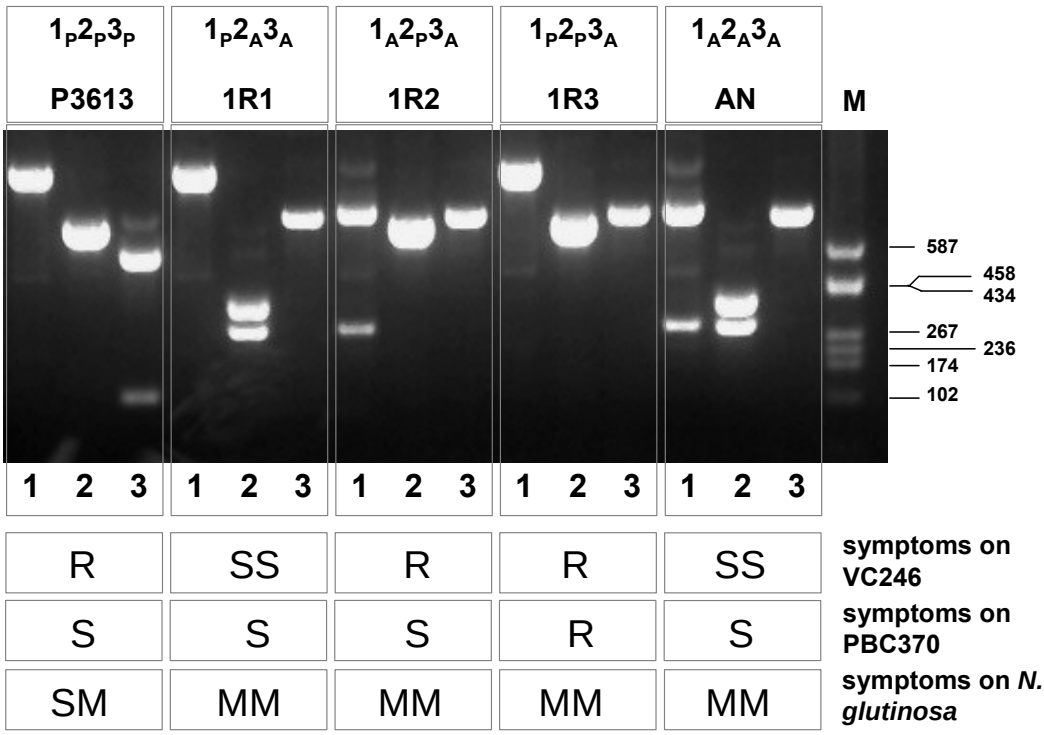


Fig. 10: RFLP analysis of the amplified fragments for determination of the genotypes of parental isolates P3613, AN and progeny reassortants.
Individual segments of CMV isolates P3613 and AN were designated as 1_P, 2_P, 3_P for segments of P3613 and 1_A, 2_A and 3_A for AN, respectively. For instance, $1_P2_A3_A$ contains RNA 1 of P3613 and RNA 2 plus RNA 3 of AN (first line). The names of the parental isolates and for the designation for the reassortants (1R1, 1R2, 1R3) were given in the second line. The corresponding phenotypes on chilli VC246, PBC370 and *N. glutinosa* were indicated as R (resistant), S (no visual symptoms, but virus was proven to be present in plant), SS (systemic infection), SM (severe mosaic) and MM (mild mosaic) in the lower part of the figure. Numbers 1, 2 and 3 under the gel represent the respective RFLP patterns of RNA 1 (digested with *Mbo* I), RNA 2 (digested with *Pvu* II) and RNA 3 (digested with *Acl* I), respectively. M= pUC9/*Hae* III. The numbers on the right site give the size of the marker in bp. Gel was stained by EtBr.

3.8.3 Manifestation of the non-infectivity of some reassortants

Since some reassortants were never obtained, they were artificially constructed by exchanging segments between two parental isolates. The reassortants [1_A 2_A 3_P] and [1_P 2_P 3_A] were achieved by *in vitro* mixing of RNA3 of isolates AN and P3613. The RNA1, RNA2 and RNA3 of both isolates were separated and excised from an agarose gel (2.2.16). The combination of RNA1+RNA2 and RNA3 segments of the two isolates were purified separately and mixed. The combinations, [1_A 2_A 3_P] and [1_P 2_P 3_A], were inoculated onto *C. quinoa* and *N. glutinosa.* Since reassortant [1_P 2_P 3_A] was obtained previously and was shown to infect both hosts it served as a positive control. The results proved that reassortant [1_A 2_A 3_P] was not infectious in both hosts, whereas [1_P 2_P 3_A] induced symptoms.

3.8.4 Correlation of CMV genome segments with phenotypes

The host responses of parental isolates and reassortants were compared. The symptoms on *N. benthamiana* and the chili line VC27a caused by the reassortants 1R1, 1R2 and 1R3 (Fig. 10) were not distinct from those caused by the parental isolates AN and P3613. On *N. glutinosa* AN induced mild mosaic symptoms and P3613 severe mosaic and dwarfing. The three reassortants caused mild mosaic on *N. glutinosa* comparable to those caused by the AN isolate. The mild mosaic symptoms induced by the three reassortants and AN on *N. glutinosa* suggested that RNA 3 of isolate P3613 is involved in the enhancement of symptom expression.

The reassortant 1R1 which included RNA 2 and RNA 3 from isolate AN (Fig. 10) induced systemic symptoms on the chili line VC246 comparable to those caused by the AN isolate. The two other reassortants, 1R2 and 1R3, did not induce any symptoms. The presence of virus in symptomatic tissue and the absence of virus in non inoculated, symptomless leaves were confirmed by DAS-ELISA (2.2.23). The non infectivity of reassortant 1R2 and symptom expression of reassortant 1R1 suggests that RNA 2 of isolate AN is responsible for the systemic infection in chili line VC246. Both of parents as well as reassortants 1R1 and 1R2 induced no symptoms on chili line PBC370, while virus was proven to be present in the symptomless plants. Reassortant 1R3 was not able to infect this line.

3.9 Confirmation of the role of RNA 2 in resistance breaking in chili line VC246 by reverse genetics

3.9.1 Exchange of a part of the viral RNA 2 genome between isolates AN and Fny

In order to verify the reassortment results (3.8) and to determine the section of RNA 2 which is responsible for resistance-breaking, a reverse genetics approach was performed with a full-length clone (Fig. 11).

The full-length CMV RNA 2 clone, named Fny209 (2.1.6), was double digested with the restriction enzymes *Nco* I and *Cpo* I (2.2.14), followed by excision of the resulting 6 kb fragment from the gel (2.2.16). This fragment was purified (2.2.16) and used as a cassette carriage system to clone foreign 2a/2b fragments.

A cassette of RNA 2, which comprised the complete 2b ORF, the 3´ terminal part of ORF 2a and the 3´ terminus of the non-translated region, was amplified from isolate AN by RT-PCR (2.2.11.2) using the forward primer *Nco* I-RNA 2 containing the *Nco* I restriction enzyme recognition site, and the reverse primer *Cpo* I-RNA 2 containing the *Cpo* I restriction enzyme recognition site (Table 2). The amplified fragment was purified by phenol/chloroform extraction (2.2.7) followed by ethanol precipitation (2.2.8) prior to ligation into the T-vector (2.2.18). The ligation product was transformed into *E. coli* NM522 (2.2.19), and transformants were selected by ampicillin resistance. Positive clones were corroborated by clone-PCR (2.2.12) and restriction enzyme digestion (2.2.14). The cloned fragments were verified by sequencing with primers T7 and T3 (2.2.20.3).

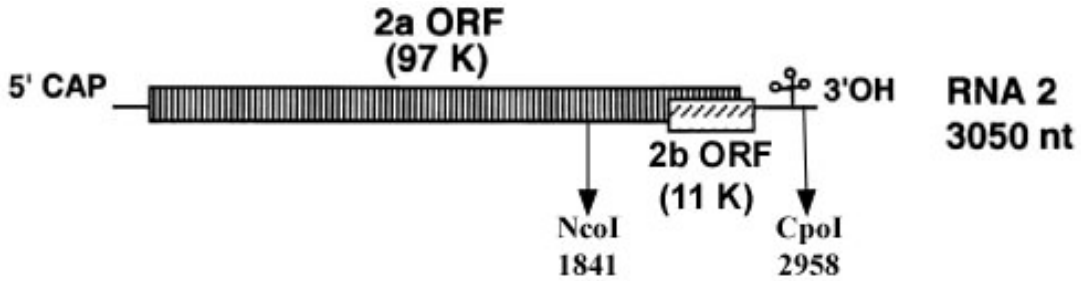


Fig. 11: Map of RNA 2 genome segment and position of restriction enzyme recognition sites *Nco* I and *Cpo* I used for reverse genetics.
A 1100 bp fragment amplified by RT-PCR from the AN isolate was exchanged with *Nco* I and *Cpo* I restriction enzyme sites, which covered the 3´-terminial part of the 2a ORF and the complete 2b ORF. 2a ORF (dark box) and 2b ORF (light box) were indicated separately (from Roossinck, 2002, modified).

The plasmid from a positive clone was extracted (2.2.20.2), followed by a double digestion with the restriction enzymes *Nco* I and *Cpo* I. The resulting band of about 1100 bp in the agarose gel electrophoresis was excised and purified (2.2.16) prior to ligation with the Fny209 vector (see above). The ligation product was transformed into *E. coli* NM522. Positive clones were confirmed by clone-PCR (2.2.12) with primers RNA2-NcoI and RNA2-CpoI (Table 2) and by restriction enzyme digestion with enzymes *Nco* I and *Cpo* I. The achieved positive clone was designated as Fny209ΔAN.

3.9.2 Correlation of the Fny209ΔAN with phenotype

Transcripts from Fny209 (RNA 2) and the chimeric Fny209ΔAN (RNA 2, with isolate AN cassette, 3.9.1) were separately combined with purified RNAs of isolate P3613 prior to inoculation of *C. quinoa* (2.2.4). Two new reassortants, designated as $[1_p2_F3_p]$ and $[1_p2_{F+AN}3_p]$, were obtained as previously described (Fig. 10, 3.8). The primer pairs 5´-RNA 2 and 3´-RNA 2 (Table 2) were used for RT-PCR and the resulting amplicons of Fny209 and Fny209ΔAN were differentiated by digestion with *Msp* I and *Pvu* II, respectively. P3613 was not digested with either enzymes, *Msp* I or *Pvu* II, and showed therefore a single band on the gel. Whereas the amplicon of Fny209 resulted in three fragments of about 50 bp, 90 bp and 600 bp when digested with *Msp* I, and the amplicon of Fny209ΔAN resulted in two fragments of about 430 bp and 300 bp when digested with *Pvu* II.

The two new reassortants, $[1_p2_F3_p]$ and $[1_p2_{F+AN}3_p]$, were inoculated onto the chili line VC246 (2.2.4). The chimeric virus $[1_p2_{F+AN}3_p]$ induced mosaic and mottling, which were the same as those induced by AN. The chimeric virus $[1_p2_F3_p]$ neither induced symptoms nor was virus detected by DAS-ELISA (2.2.23) and RT-PCR (2.2.11). RNA was extracted from $[1_p2_{F+AN}3_p]$ infected VC246 plants and used for RT-PCR. A 1100 bp fragment was amplified with primers RNA2-Nco I and RNA2-Cpo I, further cloned into T-vector (2.2.18, 2.2.19) and sequenced (2.2.20.3). The sequence provided proof that no mutations in the respective region were generated during inoculation passages (data not shown).

3.10 RNA 2 segment mutation revealed that two selected single mutations were not involved in resistance-breaking on chili line VC246

Comparison of the 2a and 2b proteins of the non resistance-breaking CMV isolates CN04, P3613, P522, Fny and the resistance-breaking isolates KS44 and AN (Table 8) revealed

seven amino acid differences in the 1100 nts cDNA region, bordered by the *Nco* I and *Cpo* I sites. Six amino acid differences were located in the 2a protein, whereas two amino acid differences were located on the 2b gene encoded protein (Fig. 12).

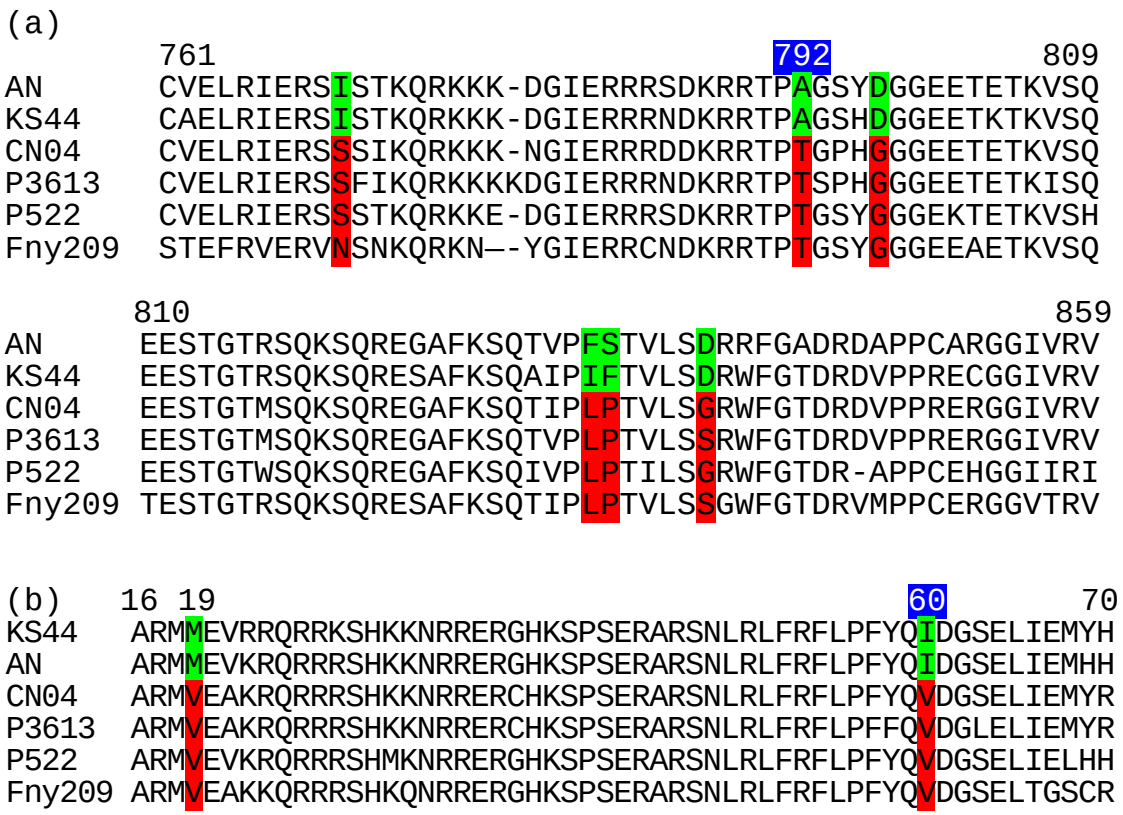


Fig. 12: Parts sequence comparisons of the 2a and 2b genes.
(a) Alignment of part of the predicted 2a protein sequences of non resistance breaking and resistance breaking CMV isolates between positions 760 and 859. The striking amino acid differences were indicated with color. Position 792 (blue) was selected for mutagenesis. (b) Alignment of part of the predicted 2b protein sequences of several non resistance breaking (P3613, P522 and Fny) and resistance breaking CMV isolates (AN and KS44) between positions 16 and 70. Two striking amino acids positions are labelled with colours. Two additional mutations are at amino acid positions 19 and 60, respectively. Amino acid 60 (blue) was selected for mutagenesis.

Based on the comparison of amino acids of non resistance breaking and resistance breaking isolates, two amino acid exchanges (Fig. 12, blue), one in the 2b ORF and one in the 2a ORF, were selected for point mutations.

The two mutated full-length clones were obtained by the PCR-based mutagenesis procedure (Fig. 2, 2.2.13). Two common primers, 7 and 8 (2.1.4, Table 2), were used for constructing all mutants, whereas two complementary mutagenic primers (2.1.4, Table 2) that contained the desired mutation, were used for constructing each of the mutants. Two overlapping fragments, with the overlap defined by the two mutagenic primers (primers 9 and 10 or primers 11 and 12, respectively), were amplified in the first two PCRs using primers 7 or 8 in concert with one of the mutagenic primers. Both products were excised from the agarose gel (2.2.10.1) and purified (2.2.16) before they were combined (1µl of each for 25

µl PCR system) as a template for the second PCR using primers 7 and 8. The resulting 1100 bp fragment was processed further (2.2.7 and 2.2.8), double digested with *Nco* I and *Cpo* I (2.2.14) and cloned back into the Fny209-vector (2.2.18-20), which was linearized with *Nco* I and *Cpo* I. (2.2.14).

The *Nco* I and *Cpo* I fragments in all two mutant constructs were completely sequenced (2.2.20.3) to confirm the introduced mutation and to ensure the absence of any undesired mutations that might have been accidentally introduced by the PCR mutagenesis protocol. The resulting mutants were designated as Fny209 (T/A) and Fny209 (V/I), referring to the Fny209 origin of the molecule with the sequence change in the encoded 2a (T exchanged with A at position 792) and 2b proteins (V exchanged with I at position 60), respectively. These full-length clones were linearized, *in vitro* transcribed (2.2.21) and further used for coinoculation with purified RNA of P3613 (2.2.3) to get the two reassortants M1, whose RNA 1 and RNA 3 originated from isolate P3613, whereas RNA 2 was Fny209(V/I), and M2, whose RNA 1 and RNA 3 originated from isolate P3613, while RNA 2 was Fny209(T/A) (Table 11). However, both new expected reassortants were never attained from the selection of 80 total local lesions each in at least four independent experiments.

To attain any possible biologically active reassortants with single mutation, another strategy was used. Two full-length clones Fny109 and Fny309, which contains the RNA 1 and RNA 3 of isolate Fny, respectively, were *in vitro* transcribed (2.2.21). Both transcripts were combined with the transcript of Fny209, Fny209(V/I), Fny209(T/A) and Fny209ΔAN, respectively, in equal amounts of RNA. Those three RNA mixtures (MR1, MR2 and MR3, Table 11) were firstly inoculated onto *N. benthamiana*, which was further used as inoculum for six tobacco hosts (Table 12) and chili VC246. However, reassortant M3 was never obtained in at least three independent experiments.

Three reassortants, MR1, MR2 and MR3, could not infect chili line VC246. Those results suggested that the two single mutations are not responsible for the resistance breaking on chili line VC246.

Table 11: the genotype and induced symptoms on chili of five reassortants

[1]designation	genotype			[2]symptoms on chili VC246
	RNA1	RNA2	RNA3	
MR1	Fny109	Fny209(V/I)	Fny309	R
MR2	Fny109	Fny209(T/A)	Fny309	R
MR3	Fny109	Fny209	Fny309	R
MR4	P3613	Fny209	P3613	R
MR5	P3613	Fny209ΔAN	P3613	SS
M1	P3613	Fny209(V/I)	P3613	/
M2	P3613	Fny209(T/A)	P3613	/
M3	Fny109	Fny209ΔAN	Fny309	/

[1]Detailed information of reassortants MR4 and MR5 is given in chapter 3.8. Reassortants M1-M3 were never attained in at least three independent experiments in *C. quinoa* and *N. glutinosa*. [2]R=no visual symptoms and ELISA negative; S=no visual symptoms but ELISA positive; SS=visually severe mosaic symptoms and ELISA positive; /, data not available.

MR1 and MR3 induced severe symptoms in the six tobacco hosts, whereas MR2 induced mild symptoms (Table 12). However, virus was proven to be present in the whole plant by PCR. MR2, a mutant with the single mutation in the 2a ORF, which codes for the RdRp, induced weak symptoms on all tobacco varieties.

Table 12: Symptoms induced by Fny isolate with and without single mutation in tobacco

designation	[1]symptom on tobacco					
	N.glutinosa	*N. tabacum cv.* Xanthi	*N. tabacum cv.* Samsum	*N. benthamiana*	*N. clevelandii*	*N. rustica*
MR1	hm, d	hm	m, mo	m	b	b, cr
MR2	nt	nt	nt	nt	nt	b, cr
MR3	hm, d, s	mo	m, mo	c, rl, d	rl, d	b, cr, rl

[1]Symptoms were indicated as: b=blistering; c=curling; cr=crinkling; d=dwarf; hm=heavy mosaic; m=mosaic; mm=mild mosaic; mo=mottle; nt=no visual symptom, but plant was proven to be infected by RT-PCR; rl=reduction of leaf size; s=shoestring.

3.11 Reassortants of several isolates revealed that interaction and/or compatibility among segments may play an important role in symptom expression

To ascertain whether the mapping of resistance breaking with isolates AN and P3613 is case dependent, reassortants of more isolate combinations were tested. For this, firstly genomic markers have been established and then reassortants were generated as described before (3.8). The frequency of detection of the parental and all possible reassortant genotypes revealed to be non random by X^2 analysis in all combinations (Appendix 7.4).

3.11.1 Selection of genomic markers of three RNA segments between different isolates

The genome segments from each pair of CMV isolates were distinguished by RFLP analysis of single tube RT-PCR (2.2.11.3) or IC-RT-PCR (2.2.11.4) amplified from RNA 1, RNA 2 and RNA 3 with the primers described previously (3.8) and additionally with primer pair 5'-MP and 3'-MP for MP gene. The origin of each genome segment in the progeny isolates was determined from the RFLP of the RT-PCR-generated fragments. The amplified RNA 1, RNA 2 and RNA 3 (MP or CP) fragments were directly digested with restriction enzymes (2.2.14) without further purification.

The restriction enzymes for the RFLP analysis were selected either from sequence alignments of the respective region or from randomly selected enzymes. To easily distinguish the genome segments of the progeny isolates, the enzyme, which could digest the segment of one isolate while not the other isolate, was used as suitable genomic marker. For each isolate combination and each RNA segment a differentiating RFLP pattern was obtained. These enzymes are summarized in Appendix 7.3.

3.11.2 Reassortants of P3613 and KS44

Isolate KS44 induces the same severe mosaic on line VC246 as AN. To proof whether RNA 2 of KS44 induces symptom expression in chili line VC 246 reassortants between the isolates P3613 and KS 44 were investigated.

In total the three combinations $[1_p\,2_k\,3_k]$, $[1_p\,2_p\,3_k]$ and $[1_k\,2_p\,3_k]$ (Table. 13, Appendix 7.4.2) were obtained, whereas the combinations $[1_p\,2_k\,3_p]$, $[1_k\,2_k\,3_p]$ and $[1_k\,2_p\,3_p]$ were never achieved in three independent experiments.

Table 13: symptoms induced by reassortants between KS44 and P3613 on two chili lines

genotypes[1]	[2]symptoms on VC246	symptoms on PBC370
$1_p2_p3_p$ (P3613)	R	S
$1_p2_k3_k$ (2R1)	SS	SS, recovery
$1_p2_p3_k$ (2R2)	R	S
$1_k2_p3_k$ (2R3)	R	S
$1_k2_k3_k$ (KS44)	SS	S

[1]Individual RNAs of CMV isolates P3613 and KS44 were designated as 1, 2 or 3. For instance, $1_p2_k3_k$ contains RNA 1 of P3613 and RNA 2 plus RNA 3 of KS44. [2] R=resistant, S=no visual symptoms, but virus was proven to be present in plant, SS=systemic infection.

The symptoms on *N. glutinosa* and *N. benthamiana* and chili line VC27a caused by the reassortants were not distinct from the parental isolates KS44 and P3613 (3.3). On chili line VC246 the reassortants which included RNA 2 and RNA 3 from isolate KS44 induced systemic symptoms comparable to the KS44 isolate, and all other reassortants did not induce any symptoms. The presence of virus in symptomatic tissue and the absence of virus in uninoculated leaves that showed no symptoms were confirmed by DAS-ELISA (2.2.23). The noninfectivity of reassortant 2R3 and infectivity of reassortant 2R1 (Table 13) suggests that RNA 2 of isolate KS44 is responsible for the systemic infection in chili line VC246.

The parental isolates and two of the three obtained reassortants (2R2, 2R3) did not induce any symptoms on chili line PBC370, although virus could be detected in the veins by tissue print immunoassay (2.2.24). Reassortant 2R1 induced mosaic on the leaves 14 d.p.i. However, newly emerging leaves were symptomless 28 d.p.i. Tissue-print immunoassays revealed that virus was present in the veins (data not shown).

3.11.3 Reassortants of KS44 and P522

Generating reassortants between P522 and KS44 in total the two combinations [1_p 2_k 3_k] and [1_p 2_p 3_k] (Table 14, Appendix 7.4.3) were obtained, whereas the combinations [1_k 2_p 3_k], [1_p 2_k 3_p], [1_k 2_k 3_p] and [1_k 2_p 3_p] were never achieved with three independent experiments.

Table 14: symptoms induced by reassortants between P522 and KS44 on two chili lines

genotypes[1]	[2]symptoms on VC246	symptoms on PBC370
$1_p2_p3_p$ (P522)	R	R
$1_p2_k3_k$ (3R1)	R	S
$1_p2_p3_k$ (3R2)	R	S
$1_k2_k3_k$ (KS44)	SS	S

[1]Individual RNAs of CMV isolates P522 and KS44 were designated as 1, 2 or 3. For instance, $1_p2_k3_k$ contains RNA 1 of P522 and RNA 2 plus RNA 3 of KS44. [2] R=resistant, S=no visual symptoms, but virus was proven to be present in plant, SS=systemic infection.

The symptoms on *N. glutinosa* and *N. benthamiana* as well as chili line VC27a caused by the reassortants were not distinct from the parental isolates KS44 and P522 (3.3). The reassortant 3R1 which included RNA 2 and RNA 3 from isolate KS44 and the reassortant 3R2 which included RNA 3 from isolate KS44 induced no symptoms on chili line VC 246. The absence of virus in un-inoculated leaves that showed no symptoms was confirmed by DAS-ELISA (2.2.23). The non-infectivity of reassortants 3R1 and 3R3 suggests that RNA 1 of isolate KS44 is responsible for the systemic infection in chili line VC246.

The reassortants 3R1 and 3R2 could replicate in chili line PBC370, although no symptoms were induced. The capability of replication of reassortants 3R1 and 3R2 suggests that RNA3 of isolate KS44 is responsible for the infection in chili line PBC370.

3.11.4 Reassortants of KS44 and AN

Generating reassortants between AN and KS44 in total, three combinations [$1_k\,2_A\,3_k$], [$1_A\,2_k\,3_A$] and [$1_k\,2_k\,3_A$] (Table 15, Appendix 7.4.4) were obtained, whereas the combinations [$1_k\,2_A\,3_A$], [$1_A\,2_A\,3_k$] and [$1_A\,2_k\,3_k$] were never achieved with three independent experiments.

Table 15: symptoms induced by reassortants between AN and KS44 on two chili lines

genotypes[1]	[2]symptom on chilli VC246	symptom on chilli PBC370
$1_A2_A3_A$ (AN)	SS	S
$1_k2_A3_k$ (4R1)	R	S
$1_A2_k3_A$ (4R2)	R	S
$1_k2_k3_A$ (4R3)	R	S
$1_k2_k3_k$ (KS44)	SS	S

[1]Individual RNAs of CMV isolates AN and KS44 were designated as 1, 2 or 3. For instance, $1_k2_k3_A$ contains RN A1 plus RNA 2 of KS44 and RNA 3 of AN. [2] R=resistant, S=no visual symptoms, but virus was proven to be present in plant, SS=systemic infection.

The symptoms on *N. benthamiana* and chili lines VC27a and VC246 caused by both isolates were not distinct. On *N. glutinosa* AN induced a mild mosaic, whereas KS44 caused a severe mosaic (Table 4). The three obtained reassortants (4R1, 4R2 and 4R3, Table 15) could not induce any symptoms on chili line VC 246, and further ELISA tests revealed that no virus was present in those plants. The non-infectivity of the three reassortants suggests that an interaction and/or compatibility among segments of each isolate is necessary for the systemic infection in chili line VC246.

Both isolates could replicate in chili line PBC370, however do not induce visible symptoms. The infectivity of the three reassortants suggests that in contrast to the infection in

chili line VC246 not a distinct segment combination is necessary for the replication in chili line PBC370.

3.11.5 Reassortants of P522 and P3613

Generating reassortants between P522 and P3613 in total the three combinations [1_{p5} 2_{p3} 3_{p3}], [1_{p5} 2_{p3} 3_{p5}] and [1_{p3} 2_{p3} 3_{p5}] (Table 16, Appendix 7.4.5) were obtained, whereas the combinations [1_{p5} 2_{p5} 3_{p3}], [1_{p3} 2_{p5} 3_{p3}] and [1_{p3} 2_{p5} 3_{p5}] were never achieved with three independent experiments.

Table 16: symptoms induced by reassortants between P522 and P3613 on two chili lines

genotypes[1]	[2]symptom on chilli PBC370	symptom on chilli VC27a
$1_{p5}2_{p5}3_{p5}$ (P522)	R	SS
$1_{p5}2_{p3}3_{p3}$ (5R1)	R	SS
$1_{p5}2_{p3}3_{p5}$ (5R2)	R	R
$1_{p3}2_{p3}3_{p5}$ (5R3)	R	SS
$1_{p3}2_{p3}3_{p3}$ (P3613)	S	SS

[1]Individual RNAs of CMV isolates P522 and P3613 were designated as 1, 2 or 3. For instance, $1_{p3}2_{p5}3_{p5}$ contains RNA 1 of P3613 and RNA 2 plus RNA 3 of P522. [2] R=resistant, S=no visual symptoms, but virus was proven to be present in plant, SS=systemic infection.

The symptoms on *N. benthamiana* caused by the reassortants 5R1, 5R2 and 5R3 (Table 16) were not distinct from the parental isolates P522 and P3613. However, P522 induced a mild mosaic on *N. glutinosa*, while P3613 did induce a severe mosaic (Table 4). In chili line PBC 370 neither of the isolates caused symptoms, although P3613 was replicating in that plants. All of the three reassortants 5R1, 5R2 and 5R3 behaved comparable to isolate P522 and did not replicate in chili line PBC370. The absence of virus in uninoculated leaves that showed no symptoms was confirmed by DAS-ELISA (2.2.23). The noninfectivity of three reassortants suggests that RNA 1 and RNA 3 of isolate P3613 might be responsible for the replication in chili line PBC370.

The parents and reassortants 5R1, 5R3 could induce mosaic on line VC27a, while reassortant 5R2 could not infect this line. Those results revealed that the interaction and/or compatibility among segments is necessary for systemic infection on chili line VC27a.

3.11.6 Reassortants of P522 and AN

Generating reassortants between P522 and AN in total, out of the six possible ones the four combinations [1_p 2_A 3_A], [1_A 2_p 3_A], [1_p 2_A 3_p] and [1_A 2_A 3_p] (Appendix 7.4.6) were ob-

tained, whereas the two combinations [1_A 2_p 3_p] and [1_p 2_p 3_A] were never achieved with three independent experiments.

The inoculation results showed that the four reassortants were unstable, because none of them could be passaged to *N. glutinosa* and *N. benthamiana*.

3.11.7 Reassortants of RT68 and PV0420

Both isolates, RT68 and PV0420 belonged to serotype II (3.1.2). RT68 induces mosaic and PV0420 ring etches on systemically infected leaves from *N. glutinosa*. On chili line VC246 isolate PV0420 induces systemic symptoms while RT68 does not replicate. Generating reassortants between RT68 and PV0420 in total only one combination [1_{pv} 2_R 3_R] (Table 17, Appendix 7.4.7) was obtained, whereas the combinations [1_{pv} 2_{pv} 3_R], [1_R 2_{pv} 3_R], [1_{pv} 2_R 3_{pv}], [1_R 2_R 3_{pv}] and [1_R 2_{pv} 3_{pv}] were never achieved with three independent experiments.

Table 17: symptoms induced by reassortants between RT68 and PV0420 on two chili lines

genotypes[1]	[2]symptoms on line VC246	symptoms on line PBC370	symptoms on line VC27a	symptoms on *N. glutinosa*
$1_{pv}2_{pv}3_{pv}$ (PV0420)	SS	S	SS	ring etching
$1_{pv}2_R3_R$ (7R1)	R	R	R	mosaic
$1_R2_R3_R$ (RT68)	R	R	SS	mosaic

[1]Individual RNAs of CMV isolates RT68 and PV0420 were designated as 1, 2 or 3. For instance, $1_{pv}2_R3_R$ contains RNA 1 of PV0420 and RNA 2 plus RNA 3 of RT68. [2] R=resistant, S=no visual symptoms, but virus was proven to be present in plant, SS=systemic infection.

The reassortant 7R1 induced a serious mosaic turning into necrosis on *N. glutinosa*, which is clearly deviating from the symptoms of the parental isolates PV0420 and RT68 (Table 4). The symptomatology of reassortant 7R1 on *N. glutinosa* suggested that RNA 2 and/or RNA 3 are involved in the expression of ring etching. Moreover, the reassortant 7R1 could not infect chili line VC 246. The presence of virus in symptomatic tissue and the absence of virus in uninoculated leaves that showed no symptoms was confirmed by DAS-ELISA (2.2.23). The noninfectivity of reassortant 7R1 suggests that RNA 2 and/or RNA 3 of isolate PV0420 is/are responsible for the systemic infection in chili line VC246.
Isolate PV0420 could infect PBC370, whereas RT68 and reassortant 7R1 could not. This result suggested RNA 2 or/and RNA 3 is/are responsible for the infection on line PBC370.

Both parents could induce mosaic on line VC27a, while reassortant 7R1 could not infect this line. Those results revealed that the interaction among segments is necessary for systemic infection on chili line VC27a.

3.11.8 Reassortants of AN and PV0420

Inter-serotype (serotype I and serotype II) reassortants were generated with isolates AN (serotype I) and PV0420 (serotype II). Both isolates induced the same mosaic and leaf-size reduction in chili line VC246 and VC27a, severe leaf curling in *N. benthamiana*, whereas PV0420 caused ring etches on leaves in *N. glutinosa*, *N. tabacum* cv. Xanthi NC and *N. tabacum* cv. Samsun NN (Fig. 13).

Generating reassortants between AN and PV0420 in total only one combination [$1_A2_A3_{pv}$] (Table 18, Appendix 7.4.8) was obtained, whereas the other combinations [$1_{pv}\,2_A\,3_A$], [$1_{pv}\,2_{pv}\,3_A$], [$1_A\,2_{pv}\,3_A$], [$1_{pv}\,2_A\,3_{pv}$] and [$1_A\,2_{pv}\,3_{pv}$] were never achieved with three independent experiments.

The obtained reassortant 8R1 could not infect chili line VC246, whereas could induce mosaic symptoms in *N. glutinosa, N. tabacum* cv. Xanthi NC and *N. tabacum* cv. Samsum NN. Those results revealed that RNA 3 of AN or RNA 2 and RNA 1 of PV0420 was/were responsible for systemic infection on chili line VC246. Furthermore, RNA 1 or/and RNA 2 of PV0420 was/were involved in the determinant of leaf ring etching on tobacco.

Table 18: The symptoms induced by AN and PV0420 as well as their corresponding reassortants in chilli and tobacco

[1]genotypes	[2]symptom in line VC246	symptom in tobacco			
		N.glutinosa	*N. tabacum* cv. Xanthi NC	*N. tabacum* cv. Samsum NN	*N. benthamiana*
$1_{pv}2_{pv}3_{pv}$	SS	ring etching	ring etching	ring etching	curling
$1_A2_A3_{pv}$ (8R1)	R	mosaic	mosaic	mosaic	curling
$1_A2_A3_A$	R	mosaic	mosaic	mosaic	curling

[1]Individual RNAs of CMV isolates AN and PV0420 were designated as 1, 2 or 3. For instance, $1_{pv}2_A3_A$ contains RNA 1 of PV0420 and RNA 2 plus RNA 3 of AN-CMV. [2]R, no visual symptoms and ELISA negative; S, no visual symptoms but ELISA positive; SS, visually heavy mosaic symptoms and ELISA positive.

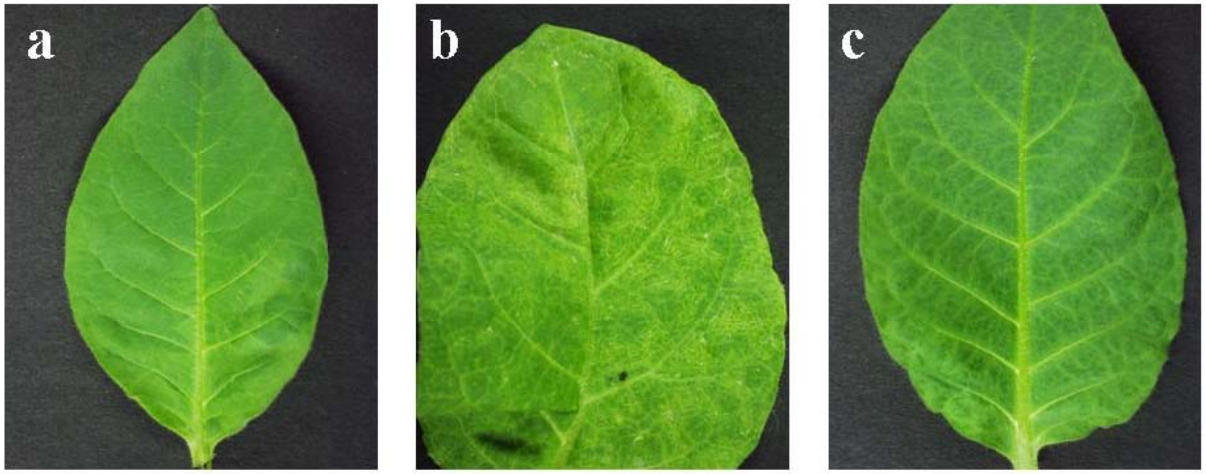


Fig 13: The symptoms induced by AN, PV0420 and reassortant $1_A2_A3_{pv}$ on *N. tabacum* cv. *Xanthi*.
(a). *N. tabacum* cv. Xanthi inoculated with isolate AN; (b). *N. tabacum* cv. *Xanthi* inoculated with reassortant [$1_A2_A3_{pv}$]; (c). *N. tabacum* cv. Xanthi inoculated with isolate PV0420.

4 Discussion

Cucumber mosaic virus (CMV) is an economically important pathogen on chili in Asia, but commercial resistant plants against this virus are still not available. The AVRDC has established a worldwide collection of chili germplasm, some of which were proven to be resistant to two Taiwan isolates (AVRDC, 2000; 2001). For evaluation of these chili lines, knowledge of the virus diversity was needed. For this, CMV isolates were collected from Asian chili fields and characterized serologically, by RFLP analysis and by symptomatology on tobacco and chili. In addition, the phylogenetic background was analyzed, using the sequences of the CP, MP and 2b ORFs. Based on sequence data, a microarray system for the differentiation serogroups and serotypes was developed. As a rule relationships between host symptoms (tobacco and chili) and molecular data were not found.
Resistance screening of the tested chili lines from AVRDC showed that resistance breaking virus isolates were present in the Asian chili production areas. The basic resistance mechanism of these chili lines was investigated by grafting and reciprocal crosses. The genomic determinant for the resistance breaking of one isolate was mapped to RNA 2 by reassortment studies and was confirmed by reverse genetics. Two single amino acid (aa) changes did not prove to be responsible for resistance breaking in chili, but caused changes in symptomatology on tobacco. Additional reassortment studies revealed that symptomatology on several hosts are strongly isolate dependent.

According to biological, serological and nucleic acid properties, CMV isolates are placed into two main serotypes, commonly referred as serotype I and II (Palukaitis *et al.*, 1992). Apparently there are no differences in the host range of the two serotypes (Anonymous, 1998). The isolates investigated in this study could be classified into the two serotypes using five monoclonal antibodies (Mabs) described previously (Anonymous, 1998). Four isolates, from US and the Netherlands, belonged to serotype II, whereas all others, including the Asian ones, belonged to serotype I. An additional set of 135 field samples collected from five chili production areas in southern China were also classified as serotype I. My results indicated that the serotype I isolates are prevalent in Asian countries, at least in tropical and subtropical areas. This is in good agreement with previous reports (Haase *et al.*, 1989; Yu *et al.*, 2005), which showed that higher thermal optima favored the speciation of serotype I. Furthermore, the prevalence of one over another serotype also has been reported from other countries (Daniels and Campbell, 1992; Crescendi *et al.*, 1993; Perry

et al., 1993; Rodriguez-Alvaro *et al.*, 1995; Fraile *et al.*, 1997). Hristova *et al*. (2002) proposed that the host plant might be an important factor for the prevalence of the different CMV serotypes. This host factor might also explain the exclusive detection of serotype I in this study since the majority of isolates investigated was attained from *Capsicum* sp. The prevalence of serotype I in Asian countries might also be explained by the prevalence of certain aphid species because different aphid species may exist in certain niches and may transmit the virus with different efficiencies. However, no data for vector species and their ability for transmission in the respective regions are available. For a better understanding of the incidence and prevalence of the two serotypes in Asian chili production areas, more samples from tropical, subtropical, and temperate regions, different hosts and a series of growing seasons as well as transmission studies would need to be tested.

Several methods for the differentiation of serotypes by Mabs have been reported (Cai *et al*., 1989; Haase *et al.*, 1989; Porta *et al.*, 1989; Anonymous, 1998; Hsu *et al.*, 2000). The differentiation of the Asian isolates investigated in this study, into the two serotypes by Mabs was successful, however variability within a serotype could not be detected by this technique.

RT-PCR became an important tool for the detection of plant viruses in the last decade. It is very suitable for the evaluation of large sample numbers (Huttinga, 1996). RT-PCR-RFLP has a higher sensitivity for the differentiation of viruses as compared to the serological methods. This method was shown useful to differentiate Tobamovirus species (Letschert *et al.*, 2002) or pathotypes (Tenllado *et al.*, 1994; Velasco *et al.*, 2002) such as the *Potato virus Y* (PVY) pathotypes PVY^{O}, PVY^{N} and PVY^{C} (Boonham *et al*., 2002).

The differentiation of CMV isolates in my study into the serotype I and II by RT-PCR-RFLP analysis was carried out with the restriction enzyme *Msp* I, as it was published by Rizos *et al.* (1992). However, this method did not give a clear distinction. This is in good agreement with the findings of other researchers (Anonymous, 1998) who presented several restriction enzyme patterns of Asian CMV isolates using *Msp* I. My results showed that the RFLP pattern of CP amplicons derived from 38 isolates could clearly differentiate the two serotypes I and II with two restriction enzymes, *Mbo* I and *Taq* I. Computer-assisted enzyme analysis of additional 142 CP sequences derived from the EMBL database (data not shown) confirmed the effectiveness of this method. My results also suggest that a reliable differentiation should be based on more than one restriction pattern. Most of the RFLP differentiation methods of (plant) viruses are based solely on one restriction enzyme

(Rizos *et al.*, 1992; Tenllado *et al.*, 1994; Velasco *et al.*, 2002). However, in some cases a single mutation would lead to ambiguous or false results. For example, the differentiation by RFLP of the *Barley yellow dwarf virus* (BYDV) of the serotype PAV into two groups was not possible because of one single mutation (Mastari *et al.*, 1998). Similarly several restriction patterns, which were caused by single nucleic acid mutations, were observed from the N-protein gene of the *Tomato spotted wilt virus* (TSWV) (Heinze *et al.*, 2001). The use of five restriction enzymes to obtain an unequivocal classification of serologically related Tobamoviruses has been proposed (Letschert *et al.*, 2002).

CMV appears to be a rapidly evolving virus (Roossinck *et al.*, 1999; Schneider and Roossinck, 2000, 2001) whose mutation frequency is presumed to be higher than that of other plant viruses (Schneider and Roossinck, 2000, 2001). This is attributed to its high degree of population variation and its segmented genome (Rodriguez-Alvardo *et al.*, 1995; Roossinck, 1997). The RT-PCR-RFLP method with double or more enzymes may offer a more reliable detection of CMV.

It has been proposed to further classify serotype I isolates into subgroups Ia and Ib by using CP sequences (Roossinck *et al.,* 1999). However, the differentiation of serotype I isolates used in my study into subgroups Ia and Ib was impossible using Mabs or by RFLP analysis. Instead, a microarray based on five species-specific probes was developed, which allowed the differentiation of CMV isolates into the serotype I subgroups Ia, Ib and serotype II. The serotype I and II specific probes worked perfectly and did not produce any false positive result. Faint cross hybridization signals of the subgroup Ia specific probe (SG1a2_1r) with serotype II isolates might be explained by the fact that the probe differs from the serotype II sequences by only a single base pair in the centre of the probe sequence. The differentiation into the two subgroups Ia and Ib was possible with one exception. One of the eight subgroup Ib isolates, RT144, did not react with its subgroup specific capture probe. This can be explained by four mismatches. To avoid ambiguous results in the microarray analysis, a clear grouping has to be established and additional subgroups has to be defined before group specific probes are designed. Nevertheless, the hybridization pattern of the isolate RT144 clearly allowed correct allocation to serotype I. The result of differential hybridization against five specifically selected oligonucleotides clearly demonstrates the great potential of oligonucleotide-based microarray technology for virus diagnostics.

The differentiation of CMV serotypes and subgroups by microarray was based on five 24-

mer oligonucleotides. This is the first report using small oligonucleotides as capture probes in the microarray format. So far, cDNAs have been used as capture probes (Boonham *et al.*, 2003; Lee *et al.*, 2003). These cDNA probes are specific for the homologous virus, however, because of their size, are limited in their discriminating ability. Only sequences showing less than 80 % sequence homologies worked specifically (Boonham *et al.*, 2003). In my experiments it was shown for the first time that short oligonucleotides are suitable as capture probes for the discriminative hybridization of isolates that differ by less than 8% in amplified PCR products of more than 700 bp in length. The proven ability to hybridize successfully generic PCR products in parallel against oligonucleotides is one step towards a possibly general plant virus chip, where non-specific random oligonucleotides could serve as capture probes for genus-specific amplification products. This procedure might be faster than sequencing the amplicons to distinguish new isolatess or even species. It could certainly make the fast detection of new viruses of a given genus possible. The microarray method offers a parallel detection to many pathogens in a single test (Boonham *et al.*, 2003). In the field, mixed infection with a number of pathogens is a common phenomenon. A microarray approach offers a great opportunity for diagnosis of plant pathogens as compared with other methods, such as electron microscopy, ELISA and PCR.

The clearest allocations of isolates into serotypes and subgroups are based on sequence data. My phylogenetic analysis of a large number of isolates based on the CP, MP and 2b genes showed a pattern of divergence similar to earlier reports (Roossinck *et al*., 1999, Roossinck, 2002). Serotypes I and II as well as subgroup Ia are clearly monophyletic. Subgroup Ib, as previously defined (Roossinck *et al*., 1999), is composed of several clusters and isolates that do not belong to subgroup Ia.

In general, all sequences obtained in this study confirmed this grouping. The phylogenetic tree based on CP sequences, revealed that all Asian and two European isolates (RT6 and RT144) belonged to the recently defined subgroup Ib (Roossinck *et al.*, 1999). Closer examination revealed that isolate RT6 is related to isolate NT9 (98.9 % identity), reported from Taiwan (Hsu *et al.*, 1995) and isolate Tfn (98.9 % identity), which was probably introduced into Italy in the mid 1980s from Asia (Gallitelli, 2000). CP sequences showed that isolate RT144 is related to the Chinese isolates RT54 and SD. The history of these two isolates is not clear. However, based on the phylogenetic analysis it is justified to assume that RT6 and RT144 may have been introduced to Europe just as isolate Tfn. This agrees

with other data (Roossinck, 2002) which showed that one isolate (isolate Tfn) of subgroup Ib was not of Asian origin.

Based on my CP sequence data and phylogenic tests, three more subgroups besides the subgroups Ia, Ib and serotype II could be defined. The phylogenetic analysis based on MP gene sequences confirmed partly the results of the CP gene-based analysis. However, five more subgroups could be defined with MP sequences. The additional subgroups seemed to be country-dependent. Subgroups, which according to Roossinck (2002) fall into the Ib cluster, are divided further when additional sequences of CMV isolates were included for analysis. In the phylogenetic tree based on CP gene sequences, India, Thailand as well as China (including Taiwan) isolates fall into a single cluster. However, when the MP gene was used for the phylogenetic analysis, they fall into two separate clusters. This suggests that isolates from these geographic areas have the same progenitor before they became predominant in certain chili growing areas. It should be kept in mind, that excessive branching might be the result of the use of too few sequences in the analysis. To confirm the subgrouping, more sequences of Asian isolates need to be analyzed.

Serotype I isolates are thought to evolve faster than serotype II isolates (Roossinck *et al.,* 1999). Again, subgroup Ib seems to be more heterogenous than subgroup Ia. The restricted geographic distribution may imply that those isolates are the result of recent subspeciation events especially in considering that the introduction of chili into Asia was less than 450 years ago.

The differences in the clustering based on CP and MP genes may reflect different constrains of the evolution of each gene. It is generally believed that the CP predominantly interacts with itself or with the viral RNA and has little interaction with host components (Edwardson, *et al.*, 1991; Palukaitis *et al.*, 1992). Also, CP interaction with vector components is minimal and non-specific, since more than 85 species of aphids can transmit this virus (Edwardson *et al.*, 1991; Palukaitis *et al.*, 1992). In contrast, the MP is involved in viral cell-to-cell movement (Canto *et al.*, 1997) and interacts with plasmodesmata (Ding *et al.*, 1995b). Almost all Asian isolates of this study are isolated from chili. Since each country has its own chili cultivar, the subdivision of non Ia and non-II isolates may be cultivar- rather than country-dependent.

Phylogenetic analysis of the 2b gene did not completely support the subgroups defined by CP and MP analysis. Three resistance-breaking isolates (AN, originated from India, KS44 and KB8, originated from Thailand) and three non-resistance breaking isolates (P522, CN04 and P3613, originated from China) were distributed across three seperate clusters. The 2b protein is related to long-distance movement and virulence and is a suppressor of gene silencing (Brigneti *et al.*, 1998; Ding *et al.*, 1995a). Furthermore, this host-specific function of the 2b protein is believed to provide a genetic basis for the extremely wide host range of CMV (Ding *et al.*, 1995a). The 2b gene is assumed to be subject to positive selection (Lin *et al.*, 2004). Considering these facts, it is tempting to assume that the 2b gene evolution has been host-dependent, especially since the clusters formed by the Asian isolates correlate well with their biological behaviour in chili. To confirm these aspects, more 2b-genes of isolates with known biological behaviour and distinct hosts, but from one geopgraphic area should be sequenced and analyzed.

Symptom variability is a generic property of plant viruses. Evidence for genetic variation of plant viruses was described by symptom variants from the *Tobacco mosaic virus* (TMV) by McKinney (1935). Symptom variability of CMV is one of the characteristics of this virus and has been extensively reviewed (Martelli and Russo, 1985; Palukaitis *et al.*, 1992). My results suggest that CMV isolates differ in symptom expression when tested on several hosts. In *C. quinoa* a correlation of serogroup and the type of symptom expression was present in the tested isolates of this study. Serotype I isolates induced large local lesions on that host, while serotype II isolates induced small ones. On *N. glutinosa* the isolates could be divided into four pathotypes. However, no correlation was observed with subgroups determined from sequence analysis. This might be explained by the fact that symptom expression is in some cases based on a combination of distinct genome. For example, the induction of systemic chlorosis has been attributed to RNA 2 as well as to RNA 3 (Rao and Francki, 1982). The symptom expression of other isolates has been mapped to a single genome segment, either RNA 2 or RNA 3 (Palukaitis *et al.*, 1992; Palukaitis and Garcia-Arenal, 2003). Since no detailed information of RNA 2 and RNA 1 sequences from isolates investigated in this study were available, no clear statement can be made with respect to the serotyping of these two RNAs.

The chili lines PBC370, PBC549, PBC459, previously documented as resistant or tolerant to the CMV isolate P522 and line VC246, reported to be resistant to isolates P522 and

P3613 (S.K. Green, AVRDC, personal communication), were found to be resistant to most of the isolates investigated in this study. However, a few isolates were found to be able to break the resistance of those lines. When plants of lines PBC370, PBC549 and PBC459 were inoculated with these isolates, they remained symptomless during the whole growing season although immunoblots revealed that virus was present in both, the roots and veins of the leaves. This suggests that the same resistance gene locus may be present in the three lines. The reaction pattern in line VC246 is different from that of lines PBC370, PBC549 and PBC459. Few CMV isolates could induce mosaic symptoms, but most of isolates could not infect this line and virus was not detectable in the non-inoculated leaves. These findings led to the assumption that the resistance mechanism of line VC246 is distinct from that of the PBC-lines.

In general, resistance to infection by plant viruses occurs at various levels. These include resistance to replication (true immunity or non host), resistance to movement out of the initially infected cell (subliminal infection), resistance to infection outside a zone of a limited number of cells (usually induced by hypersensitive response), resistance to movement beyond the initially infected leaf (systemic resistance), and various degrees of restriction to systemic virus movement accompanied by reduction to disease development (tolerance) (Zaitlin and Hull, 1987; Dawson and Hilf, 1992; Fraser, 1992).
The resistance barrier of the PBC lines and line VC246 against most of the CMV isolates might be based on one or a combination of the above mentioned resistance patterns. To understand which barrier is involved in the blocking of virus replication in line VC246 or in restricting the virus to the vascular system of line PBC370, plants were inoculated by grafting. Scions of VC246 developed clear mosaic symptoms when grafted onto AN-inoculated rootstocks comparable to mechanically AN-inoculated VC246 plants. However, when VC246 was grafted onto a P3613-infected rootstocks, it developed local ringspots in some of the leaves, even though line VC246 is resistant to P3613 (no virus detectable, no symptoms) when inoculated mechanically with this isolate. A similar phenomenon was observed in *Lycopersicon chmielevskii* and *L. hirsutum.* The two species were found to be resistant to CMV by mechanical inoculation, but susceptible to inoculation by grafting (Abad *et al.*, 2000). The virus may occasionally escape from the phloem into parenchyma cells and subsequently forms ringspots in some leaves due to inactivation either by plant defence mechanisms or by virus-induced gene silencing (VIGS). An explantation for the uneven distribution of ringspots in the leaves, it might be the coupling of defense mecha-

nism to the development stage of the invaded leaves. If VIGS is involved in this process, the gene silencing suppressor 2b of the resistance-breaking isolates is thought to be involved. Therefore, aboved results suggested that the replication and cell-to-cell movement of isolate P3613 are still efficient in the resistant VC246 line, and it might be possible that some of the early events before cell-to-cell movement were blocked either by active plant resistance gene(s) or by VIGS in plants of line VC246.

Line PBC370 showed tolerance to CMV infection and may have recessive, partially dominant, or polygenic recessive inheritance. The virus could replicate, move and accumulate, but could not exit from the vascular system into the lamina. Almost all reported cases of genetic resistances to CMV in *Capsicum* spp. (Green and Kim, 1991; Monma and Sakata, 1997), are recessive or partially dominant (Lapidot *et al*., 1997; Grube *et al*., 2000; Suzuki *et al*., 2003).

The inheritance patterns of only two lines, *C. annuum* 'Perennial', the leading natural resistance source in many breeding programs, and *Capsicum frutescens* 'BG2814-6', have been characterized and identified. The resistance in *C. annuum* 'Perennial', a pungent Indian chili, has been described as monogenic recessive, partially dominant, and polygenic recessive (Singh *et al*., 1979; Pochard *et al*., 1989). Quantitative Trait Loci (QTL) has been described and mapped (Ben Chaim *et al*., 2001). The resistance in the small-fruited, pungent pepper line collected from southern Mexico, *Capsicum frutescens* 'BG2814-6', is controlled by at least two major recessive genes. Both lines may share one or more resistance genes (Grube *et al*., 2000).

Line PBC370 has similarities in its resistance behaviour with the described lines and the resistance might be caused by the same or related genes. On the contrary, the gene(s) involved in the resistance of line VC246 may be dominant. Most isolates could not infect this line, but four isolates could induce mosaic symptoms. This kind of symptom seems more likely to be correlated with an active and dominant resistance rather than a recessive one. Hypersensitive or extreme resistance is a rare phenomenon in *Capsicum* spp. and was described only in three chili lines belonging to *C. baccatum* and *C. frutescens* (Suzuki *et al*., 2003). Thus, line VC246 is the first *C. annuum* line that shows dominant resistance against CMV.

To exploit the different resistance gene(s) of lines VC246 and PBC370 and to study the resistance mechanism, the two lines PBC370 and VC246 were reciprocally crossed. Most

of the tested isolates were able to infect the F1. However, recovery was observed at 28 d.pi., but virus could still be detected in the newly emerging leaves. The recovery might be based on VIGS or R-gene(s). The virus replicated, moved from cell to cell and ultimately induced mosaic in the hybrid plants. Then replication was blocked by VIGS or R-gene(s) and the plant recovered. The degree of this blocking seems to be highly isolate-dependent since three kinds of phenotypes were formed.

It is now well known that the homology-dependent virus resistance established in recovered tissues of either wildtype (wt) or transgenic plants following the primary virus infection is based on RNA silencing (Lindbo *et al*., 1993; Covey *et al*., 1997; Guo and Garcia, 1997; Ratcliff *et al*., 1997, 1999; Al Kaff *et al*., 1998; Ruiz *et al*., 1998). However, with the present data it is not possible to decide which mechanism, R gene or VIGS, plays the major role in recovery in the VC246 x PBC370 hybrid lines. It is reasonable to assume that VIGS may be involved, since this phenomenon is more dependent on isolates rather than on the plant lines. A single nucleic acid mutation in the *Tobacco streak virus* (TSV) genome was found to be sufficient to revert the phenotype from recovery to non-recovery (Xin and Ding, 2003). This proves that host recovery can be initiated by minimal genetic changes in the viral genome.

Isolates AN and P3613, which differ in their ability to infect the chili line VC246, were used to investigate the molecular basis to overcome the resistance of line VC246 using reassortments. A set of reassortants derived from P3613 and AN were produced and used to check their ability to systemically infect the line VC246. Reassortants with RNA 2 of isolate AN combined with RNA 1 and RNA 3 of both isolates induced systemic infection in line VC246. However, reassortants with RNA 2 of isolate P3613, combined with RNA 1 and RNA 3 of both isolates failed to systemically infect VC246 plants. From that it can be concluded that RNA 2 of isolate AN is responsible for breaking resistance of line VC246.

In most cases RNA 3 has been reported to be involved in symptom determination (Palukaitis *et al*., 1992; Palukaitis and Garcia-Arenal, 2003 and reference herein). Therefore, the involvement of RNA2 of the two isolates P3613 and AN in symptom determination was an unexpected phenomenon. This function of RNA2 was reported in only a few cases (Palukaitis *et al.*, 1992; Palukaitis and Garcia-Arenal, 2003). RNA 2 was shown to be a determinant for elicitation of HR in cowpea (Marchoux *et al*., 1975; Hanada and Tochihara, 1980; Edwards *et al*., 1983). Furthermore, two amino acids in the 2a protein were

identified to influence the phenotype (Kim and Palukaitis, 1997; Karasawa *et al*., 1999). The involvement of RNA 2 in determining the efficiency of systemic infection in radish (Takeshita *et al*., 1998) and zucchini squash (Choi *et al*., 2003) has been reported.

To determine the segment of RNA 2 of isolate AN which is responsible for breaking the resistance of line VC246, an intraspecies hybrid virus, named $1_p2_{F+AN}3_{p,}$ was generated. The hybrid virus consisted of RNA 1 and RNA 3 of the non resistance-breaking isolate P3613 and an RNA 2 construct, which included the 2 kb 5´-terminal part of RNA 2 from the non resistance-breaking isolate Fny and the 1 kb 3´-terminal part of RNA 2 from the resistance-breaking isolate AN. This chimeric virus $1_p2_{F+AN}3_p$ induced the same symptoms as AN on line VC246. This result confirmed the conclusion drawn from the reassortment experiment, namely that the genetic determinant is coded on RNA 2. Since the exchanged RNA 2 cassette in the chimeric virus $1_p2_{F+AN}3_p$ covered the conserved GDD RdRp motif and the complete ORF of 2b, it is not clear whether mutation(s) in 2a and/or 2b are responsible for the resistance-breaking. To clarify this, detailed studies using reverse genetics will be necessary with deletion of either ORF 2a or 2b and complementation with the respective genes of a non resistance-breaking isolate. It is likely that the systemic infection determinant of isolate AN is expressed at the amino acid level rather than the RNA level. There is only one report describing viral genomic RNA as an elicitor of an hypersensitive reaction (Szittya and Burgyan, 2001). An alignment of the exchanged 1100 bp section with sequences of non resistance-breaking and resistance-breaking isolates revealed seven amino acid differences. Of these, two amino acid positions, one in 2b Val^{60} (Fny) to Ile^{60} (AN) and one in 2a Thr^{792} (Fny) to Ala^{792} (AN), were selected for single point mutations. These two amino acid exchanges were selected because of the different properties of the amino acids, which are thought to be more likely associated with changes in the phenotype rather than changes of amino acids with similar attributes. However, the conversion from Val^{60} to Ile^{60} of 2b protein results in an unavoidable change of an amino acid in the 2a protein, because of the overlapping coding of the two genes.

Unfortunately reassortants of this mutated RNA 2 with RNA 1 and 3 of isolate P3613 were not infectious on chili. Therefore *in vitro* transcripts from the full length clones of isolate Fny from RNA 1 and 3, respectively, were used for complementation with the mutated RNA 2. The wild type Fny isolate could not infect line VC246. The two chimeric viruses with single mutations caused neither symptom nor were they detectable in the non-inoculated leaves. Therefore it is concluded that the two amino acid changes are not alone

responsible for the resistance-breaking of line VC246. Further studies with other amino acid changes may single out the amino acid determinants that induced resistance breaking.

Interestingly, both reassortants with single amino acid changes in RNA 2 induced different symptoms on six tobacco species as compared to reactions with the parental isolate Fny. Symptom severity was less in all tobacco species infected by the virus with the Val^{60} to Ile^{60} exchange in the 2b protein. The virus with the Thr^{792} to Ala^{792} exchange of the 2a protein induced no symptoms on five tobacco species, except on *N. rustica* in which mild symptoms developed. However, in all plants virus replicated since the presence of virus was confirmed by DAS-ELISA.
Obviously the single amino acid changes reduced virulence of the virus in all tobacco hosts. One single amino acid of RNA 2 at position 631 is responsible for the HR in cowpea (Karasawa *et al*., 1999). I could also show that one single amino acid exchange of RNA 2 at position 792 lead to a change in symptoms. Two explanations are proposed: Firstly, because RNA 2 encodes the RdRp, a part of the replicase complex, the exchange of amino acids in the 2a protein may affect the replication efficiency. This may lead to a reduced accumulation of virus to levels required at a critical time that are needed for the onset of symptom development. Secondly, it is likely that the exchange of an amino acid in the 2a protein reduced or abolished the interaction between 2a and/or the plant factors which also may reduce the replication efficiency.

Several mechanisms in the host plant may play a role in resistance-breaking. Dominant resistances are the result of active mechanisms. They correspond to incompatible interactions between viruses and hosts that usually result in a HR controlled by dominant resistance (R) genes (Fraser, 1990, 1992, 1999). Different viral proteins, including RdRps (Meshi *et al*. 1988; Hamamoto *et al*. 1997; Padgett *et al*. 1997; Erickson *et al*. 1999), MPs (Meshi *et al*. 1989; Weber *et al*. 1993; Weber and Pfitzner 1998) and CPs (Bendahmane *et al*. 1995; Berzal-Herranz *et al*. 1995) have been identified as avirulence determinants. The gene for gene theory may provide a good basis for explaining the resistance-breaking nature of isolate AN in VC246. The avirulence factor of isolate AN is clearly the 2a and/or 2b protein which was determined by using reassortants and a genetically modified virus. Although the exact mechanism is not clear yet it may involve direct interaction of viral and host proteins. On the other hand, the mutated 2b protein may escape the targeting of the

corresponding R protein of line VC246. Therefore, either VIGS or a combination of R-gene mediated resistance and VIGS may induce resistance-breaking.

In lines PBC370, PBC549 and PBC495 the virus was either not detectable or restricted to the vascular tissue. The genetic determinants of P3613 that support virus multiplication were mapped into RNA 1 and RNA 3 by reassortment studies. These lines seem to have a recessive or incompletely dominant resistance as discussed above. The recessive or incompletely dominant trait was thought to result from loss or the presence of a mutated version of factor(s) essential for virus multiplication and/or movement within the host plant. This type of resistance may allow spread of virus, but inhibit multiplication and/or symptom development (Fraser, 1990, 1992, 1999). All recessive virus resistances characterized to date agree with this model (Schaad and Carrington, 1996; Nicolas *et al*., 1997; Keller *et al*., 1998; Johansen *et al*., 2001; Ruffel *et al*., 2002; Yamanaka *et al*., 2001, 2002; Diaz *et al*., 2004). Therefore, the resistance of these lines may be the consequence of the lack or existence of a mutated version of a factor involved in the establishment of infection. These factors are necessary for direct or indirect interaction with viral proteins encoded on RNA 1 and 3. The phenomenon of virus restriction to the vascular tissue was also described by other researchers. Systemic spread of TSWV was blocked in a N-transgenic tobacco line and virus was restricted to the vascular tissue (Schwach *et al.,* 2004). A similar phenomenon was found for a CP gene-mediated resistance against *Potato mop-top virus* (PMTV) in transgenic *N. benthamiana* plants (Germundsson *et al*., 2002). In both cases the VIGS appeared to be involved although other mechanisms may also play a role. My reassortment studies indicate that an involvement of the gene silencing suppressor of CMV in chili for restriction of the virus in the veins can only be effective in combination with other viral proteins encoded on RNA 1 and 3.

During the last decade the pathogenic determinants of the CMV genome for some hosts, mainly tobacco, were mapped extensively with “artifical” (*in vitro* transcripts) reassortants (Palukaitis and Garcia-Arenal, 2003 and references herein). However, all pathogenic factor(s) of CMV in previously published experiments were determined using only two isolates. When mapping with different isolate combinations I could show that the pathogenic determinant is distinct for each isolate.

Six different isolates (P3613, P522, AN, KS44, RT68, PV0420) covering subgroup Ib and serotype II were used to determine pathogenic determinants. Selection of these isolates was based on their different biological behaviour in chili lines VC246, PBC370 as well as in *N. glutinosa*. Results showed that the pathogenic determinant of the CMV genome on chili and tobacco is dependent on the RNA segment combination. For example, the genetic determinant of isolate KS44 that leads to resistance-breaking in line VC246 was mapped to RNA 2 with reassortants of isolate P3613 and isolate KS44, but was mapped to RNA 1 with reassortants of isolate P522 and isolate KS44. Isolate PV0420, a serotype II isolate induces ring-like etching on *N. glutinosa* and *N. tabacum* cv. Samsun NN. This genetic trait was mapped to RNA 2 and/or RNA 3 with reassortants of RT68 and PV0420, whereas it was mapped to RNA 1 and/or RNA 2 with reassortants of AN and PV0420. RNA 3 was reported to be the determinant of ring-like etching when isolates Fny and LS were used for genetic mapping (Zhang *et al.,* 1994).

The above data suggested that the intricate interaction among segments of virus genomes may play an important and necessary role during the pathogenetic process. There may be a requirement for direct, coordinated interactions between these segments or corresponding proteins. In case of lack of coordinated functions, inoculation will lead to differences in the virulence. Moreover, the genome organisation of most plant viruses, which are typically densely packed with information, may provide a molecular basis for the interaction among genomic sequences and individual gene products. In the case of CMV, five known functional viral proteins are encoded in three segments and each protein plays a multifunctional role (Palukaitis *et al.*, 1992; Palukaitis and Garcia-Arenal, 2003). It is well known that some of the pathogenic processes need the involvement of more than one protein. For example, all of the CMV-encoded proteins appear to play some role in virus movement (Palukaitis *et al.*, 1992; Palukaitis and Garcia-Arenal, 2003; Salánki *et al.*, 1997). Similar coordinated interactions for pathogenicity among RNA segments of tripartite plant virus have been reported. Systemic infection levels with the *Brome mosaic virus* (BMV) in cowpea were influenced by all three genomic RNAs (De Jong and Ahlquist, 1995).

Interestingly, three different genetic determinants of the isolate Fny, which infects squash systemically, have been reported. Isolate Fny systemically infected zucchini squash (*Cucurbita pepo* cv. Black Beauty), and plants showed severe symptoms. Another isolate Sny infected zucchini squash, but induced mild symptoms. Reassortments between isolate Fny

and Sny were constructed. Assessment of symptoms induced by the reassortants indicated that RNA 1 is responsible for both the severity and the rapid induction of symptoms elicited by isolate Fny in zucchini squash (Roossinck and Palukaitis, 1990). However, reassortants of isolate Fny and isolate M, which could not infect zucchini squash systemically, revealed that CP of RNA 3 of isolate Fny is responsible for the systemic infection in zucchini squash (Shintaku and Palukaitis, 1990; Wong *et al.*, 1999). Moreover, the assessment of pathogenicity of reassortants of isolate Fny and isolate LK, which was not able to systemically infect zucchini squash, showed that RNA 2 of Fny is involved in the systemic infection in zucchini squash (Choi *et al.*, 2003). The three reports, indicating the involvement of three different RNA segements in symptom expression, support the conclusion attained in this study that biological behaviour is strongly isolate-dependent and that the significance of single amino acid changes should not be overestimated.

The compatibilities among the three segments of CMV may be important for the stability of virus particles, virus infection, movement and symptom expression. In the genera *Bunyavirus* and *Tospovirus* reassortments were restricted to viruses within the same serogroup, and even there incompatibilities were observed (Best, 1961, 1968; Pringle, 1996). This phenomenon was also observed in my experiments with several isolate combinations. Isolates AN and KS44 induced a systemic mosaic in VC246 and symptomless infection in PBC370, however, any of reassortants among them failed to infect VC246 and PBC370. Likewise, reassortants between isolates P522 and AN were not stable and passages from tobacco to other hosts failed. Furthermore, one reassortant of isolate RT68 and PV0420 failed to infect VC27a, a line susceptible to all CMV isolates.

The genetic exchange by natural reassortment of viral segments is an important source of genetic variation (Fraile *et al.*, 1997; Roossinck, 2002). Natural reassortants between subgroup Ia and Ib were recently reported (Fraile *et al.*, 1997; Lin *et al.*, 2004). However, no isolates with reassortant genomes derived from subgroups Ia and Ib were found in Asian isolates. Three reasons may be responsible for the absence of detectable reassortants in this study: (a) the sample size of this study is not high enough to detect reassortants; (b) reassortment events did occurr (between segment RNA 1 and RNA 2) but could not be detected with by RFLP analysis; (c) reassortment events did occur, but the resulting reassortants were not favored and were subsequently selected out. There are no reports of natural reassortants between serotype I and serotype II. There is only one report of reassortants of

both serotypes by *in vitro* transcripts (Zhang *et al.*, 1994). However, in this study a stable reassortant [reassortant $1_A2_A3_{pv}$, which contains RNA 1 and RNA 2 of isolate AN (serotype I) and RNA 3 of isolate PV0420 (serotype II)] was obtained by local lesion isolation, in some aspects which is a more natural generation of reassortants compared to *in vitro*-mixing of RNA transcripts.

An interesting "artifical" reassortant $1_P2_A3_A$ (RNA 1 from isolate P3613, RNA 2 and RNA 3 from isolate AN) was found in this study, which is a resistance-breaking reassortant in line VC246 and seems to be more competitive compared to the parental isolate AN. All local lesions revealed the $1_P2_A3_A$-combination to be present and no parental isolate was detected, when a mixture of isolate P3613 and AN was used for infection (data not shown). This is quite uncommon, because the pathogen's acquired ability to overcome disease resistance usually occurs at the cost of associated reduced fitness for survival (Jones, 2001; Leach *et al.*, 2001).

This is the first study of detailed characterization of Asian CMV isolates from chili. The phylogenetic analysis revealed the presence of more than the two reported subgroups, Ia, Ib and serotype II. Further studies with more isolates collected from different host plants and geographical niches of Asia are needed to confirm the presence of these additional subgroups. This will help to better understand the evolution and population structure of this virus in Asia. Optimization of the microarray technique for the differentiation of subgroups should lead to an effective tool for the detection of CMV, especially in large scale samples, and in epidemiological investigations. The study of genetic determinants of CMV isolates showed that interaction and/or compatibility among genome segments may play an important and necessary role in the process of plant-virus interaction. To get more insight in these mechanisms more reassortants of different isolate combinations have to be tested on several hosts. The study of the resistance mechanisms of four chili lines singled out a new line that can be used for breeding programs. However, further study by backcrossing with a susceptible line are necessary for a complete understanding of the resistance mechanism of line VC246. The characterization and isolation of the resistance gene(s) of line VC246 should facilitate the development of molecular markers and should provide the basis for molecular breeding.

5 Summary

Cucumber mosaic virus (CMV) is an economically important pathogen on chili in Asia. Due to the lack of commercially available resistant lines and a lack of epidemiological knowledge about the CMV population in Asia, the disease is difficult to manage. To improve this situation 38 CMV isolates originated from Asia, Europe and USA have been characterized and plant lines, originated from a worldwide germplasm collection of the AVRDC (Asian Vegetable Research and Development Centre, Taiwan), have been screened for resistance with those isolates.

The characterization of those isolates were carried out by symptomatology, serology, RT-PCR-RFLP, microarray and phylogenetic analysis using coat protein (CP), movement protein (MP) and the gene silencing suppressor genes (2b). The phylogenetic analyses with a large number of CP, MP or 2b sequences confirmed the previous classification of CMV isolates into serotypes I and II as well as subgroups Ia, Ib. However, the Asian isolates distributed into several subgroups within subgroup Ib with high bootstrap support. Serology and RT-PCR-RFLP could differentiate the serotypes I and II, a further differentiation into the subgroups Ia and Ib was best possible with the microarray technique. This technique was based on probes derived from variable parts of the CP gene. The results of the microarrays showed a clear differentiation of subgroups Ia, Ib and serotype II with oligonucleotide probes for the first time.

Symptomatology studies revealed four phenotypes on *N. glutinosa* and three phenotypes on four resistant chili lines PBC370, PBC549, PBC459 and VC246. However, it was almost not possible to establish a relationship of isolates between sequences and symptoms. The elementary resistance mechanism of two of those chili lines was investigated by grafting and reciprocal crosses and due to these results possible mechanisms of two resistant lines could be proposed. Line PBC370 may have recessive, partially dominant, or polygenic recessive inheritance, whereas the resistance gene involved in line VC246 is probably dominant. Therefore, VC246 may be a new, hopefully, true resistant line that can be used for breeding programs. Although the chili lines were resistant to most of the isolates, some of the isolates were able to overcome this resistance and caused severe symptoms. A reassortment system for mapping of the genetic traits of the CMV genome was established and used to clarify the determinants of chili resistance breaking by these isolates. The result revealed that RNA 2 of isolate AN was responsible for resistance breaking on line VC246. Further studies with reverse genetics confirmed this result and showed that a 1100 bp cassette of the 2a/2b overlapping region was responsible for resistance breaking. On the

basis of a comparison of the 2b gene of resistance breaking and non resistance breaking isolates, two single amino acids changes were selected to examine which might be responsible for this process. However, the results indicated that none of either amino acid is solely involved in resistance breaking of chili, but interestingly changed the symptoms on tobacco.

Moreover, the influence of the interaction and/or compatibility among CMV genome segments to the biological behaviour on tobacco and chili was investigated with a set of six biologically distinct isolates. These results indicated that interaction and/or compatibility among genome segments play an important role in the process of plant-virus interaction so that it was concluded that results of genetic mapping of determinants of CMV should be cautiously explained.

Taken together, the characterization of the population structure of CMV in Asia, the development of a detection method for the differentiation of serotypes and subgroups and the knowledge of the genetic background of some resistant chili lines as well as the biological properties of some resistance breaking isolates may give a basis to improve the Asian chili production in the future.

6 Bibliography

Aaziz R, Tepfer M (1999) Recombination in RNA viruses and in virus-resistant transgenic plants. *J Gen Virol* **80** :1339-1346

Abad J, Anastasio G, Fraile A, García-Arenal F (2000) A search for resistance to C*ucumber mosaic virus* in the genus *Lycopersicon*. *J Plant Pathol* **82**: 39-48

Adam G, Gaedigk K, Mundry K-W (1983) Alterations of a plant rhabdovirus during successive mechanical transfers. *Z Pflanzenkr Pflanzenschutz* **90**:28-35

Al Kaff NS, Covey SN, Kreike MM, Page AM, Pinder R, Dale PJ (1998) Transcriptional and posttranscriptional plant gene silencing in response to a pathogen. *Science* **279:** 2113–2115

Anandalakshmi R, Pruss GJ, Ge X, Marathe R, Mallory AC, Smith TH, Vance VB (1998) A viral suppressor of gene silencing in plants. *Proc Natl Acad Sci* USA **95**: 13079–13084

Anonymous (1998) Detection and biodiversity of *Cucumber mosaic cucumovirus*. Conclusions from a ringtest of European Union COST 823 (new technologies to improve phytodiagnosis). *J Plant Pathol* **80**: 133-149

Baulcombe D (1994) Replicase-mediated resistance: a novel type of virus resistance in transgenic plants. *Trends Microbiol* **2**: 60-63

Baulcombe D (1996) Mechanisms of pathogen-derived resistance to viruses in transgenic plants. *Plant Cell* **8**: 1833-1844

Baulcombe D (2002) Viral suppression of systemic silencing. *Trends Microbiol* **10**: 306-308

Beachy RN (1994) Introduction: transgenic resistance to plant viruses. *Semin Virol* **4**: 327-328

Beachy RN, Loesch-Fries S, Tumer NE (1990) Coat protein-mediated resistance against virus infection. *Annu Rev Phytopathol* **28**: 451-74

Ben Chaim A, Grube RC, Lapidot M, Jahn M, Paran I (2001) Identification of quantitative trait loci associated with resistance to *Cucumber mosaic virus* in *Capsicum annuum*. *Theor Appl Genet* **102**: 1213–1220

Bendahmane A, Kohm BA, Dedi C, Baucombe DC (1995) The coat protein of *Potato virus* X is a strain-specific elicitor of Rx1-mediated virus resistance in potato. *Plant J* **8**: 933-941

Berzal-Herranz A, de la Cruz A, Tenllado F, Diaz-Ruiz JR, Lopez L, Sanz AI, Vaquero C, Serra MT, Garcia-Luque I (1995) The Capsicum L3 gene-mediated resistance against the tobamoviruses is elicited by the coat protein. *Virology* **209**: 498-505

Best, R. J. (1961). Recombination experiments with strains A and E of tomato spotted wilt virus. Virology 15, 327–339.

Best RJ (1968) Tomato spotted wilt virus. In: *Adv Virus Res* (KM Smith and MA Lauffer, eds). Academic Press, New York, pp 65-145

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DANN. *Nucleic Acid Res* **7**: 1513-1523

Blackman ML, Boevink P, Santa Cruz S, Palukaitis P, Oparka KJ (1998) The movement protein of *Cucumber mosaic virus* traffics into sieve elements in minor veins of *Nicotiana clevelandii*. *Plant Cell* **10**: 525–537

Bonnet J, Fraile A, Sacristan S, Malpica JM, Garcia-Arenal F (2005) Role of recombination in the evolution of natural populations of *Cucumber mosaic virus*, a tripartite RNA plant virus. *Virology* **332**: 359-368

Boonham N, Walsh K, Preston S, North J, Smith P and Barker I. 2002. The detection of tuber necrotic isolates of *Potato virus* Y, and the accurate discrimination of PVY^O, PVY^N and PVY^C strains using RT-PCR. *J Virol Methods* **102**: 103-112

Boonham N, Walsh K, Smith P, Madagan K, Graham I, Barker I (2003) Detection of potato viruses using microarray technology: towards a generic method for plant viral disease diagnosis. *J Virol Methods* **108**: 181-187

Brigneti G, Voinnet O, Li WX, Ji LH, Ding SW, Baulcombe DC (1998) Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana benthamiana*. *EMBO J* **17**: 6739-6746

Bruyere A, Wantroba M, Flasinski S, Dzianott A, Bujarski JJ (2000) Frequent homologous recombination events between molecules of one RNA component in a multipartite RNA virus. *J Virol* **74**: 4214-4219

Cai WQ, Wang R, Zhao BY (1989) CMV specific monoclonal antibodies production by electrofusion. *Acta Microbiol Sinica* **29**: 444–451

Canto T, Palukaitis P (1999a) The hypersensitive response to *Cucumber mosaic virus* in *Chenopodium amaranticolor* requires virus movement outside the initially infected cell. *Virology* **265**: 74–82

Canto T, Palukaitis P (1999b) Are tubules generated by the 3a protein necessary for *Cucumber mosaic virus* movement? *Mol Plant Microbe Interact* **12**: 985–993

Canto T, Choi SK, Palukaitis P (2001) A subpopulation of *Cucumber mosaic virus* RNA 1 contains 3′ termini originating from RNAs 2 or 3. *J Gen Virol* **82**: 941-945

Canto T, Prior DA, Hellwald KH, Oparka KJ, Palukaitis P (1997) Characterization of *Cucumber mosaic virus*. IV. Movement protein and coat protein are both essential for cell-to-cell movement of cucumber mosaic virus. *Virology* **237**: 237-248

Caranta C, Palloix A, Lefebvre V, Daubeze AM (1997) QTLs for a component of partial resistance to cucumber mosaic virus in pepper: restriction of virus installation in host cells. *Theor Appl Genet* **94**: 431-438

Caranta CA, Pflieger S, Lefebvre V, Daubeze AM, Thabuis A, Palloix A (2002) QTLs involved in the restriction of *Cucumber mosaic virus* (CMV) long-distance movement in pepper. *Theor Appl Genet* **104**: 586–591

Carpenter C, Simon AE (1996) In vivo restoration of biologically active 3′ ends of virus-associated RNAs by nonhomologous RNA recombination and replacement of a terminal motif. J Virol 70: 478-486

Chen B, Francki RIB (1990) Cucumovirus transmission by the aphid *Myzus persicae* is determined solely by the viral coat protein. *J Gen Virol* **71**: 939–944

Choi SK, Choi JK, Park WM, Ryu KH (1999) RT-PCR detection and identification of three species of cucumoviruses with a genus-specific single pair of primers. J Virol Methods **83**:67-73

Choi SK, Choi JK, Ryu KH (2003) Involvement of RNA2 for systemic infection of *Cucumber mosaic virus* isolated from lily on zucchini squash. *Virus Res* **97**:1-6

Cogoni C, Macino G (1999) Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature* **399**: 166-169

Cogoni C, Irelan JT, Schumacher M, Schmidhauser T, Selker EU, Macino G (1996) Transgene silencing of the al-1 gene in vegetative cells of Neurospora is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. *EMBO J* **15**: 3153-3163

Covey SN, Al-Kaff NS, Langara A, Turner DS (1997) Plants combat infection by gene silencing. *Nature* **385**: 781–782

Crescenzi A, Barbarossa L, Gallitelli D, Martelli GP (1993) *Cucumber mosaic virus* populations in Italy under natural epidemic conditions and after a satellite-mediated protection test. *Plant Dis* **77**: 28–33

Culver JN (1997) Viral avirulence genes. In: *Plant-Molecular Interactions*, vol 2. (G Stacey and NT Keen, eds.) Chapman and Hall, New York. pp 196-219

Daniels J, Campbell RN (1992) Characterization of *Cucumber mosaic virus* isolates from California. *Plant Dis* **76**: 1245-1250

Dawson WO, Hilf ME (1992) Host-range determinants of plant viruses. *Annu Rev Plant Physiol Plant Mol Biol* **43**: 527-555

De Jong W, Ahlquist P (1995) Host-specific alterations in viral RNA accumulation and infection spread in a *Brome mosaic virus* isolate with an expanded host range. *J Virol* **69**: 1485–1492

Diaz JA, Nieto C, Moriones E, Truniger V, Aranda MA (2004). Molecular characterization of a *Melon necrotic spot virus* strain that overcomes the resistance in melon and nonhost plants. *Mol Plant-Microbe Interact* **17**: 668-675

Ding SW, Anderson BJ, Haase HR, Symons RH (1994) New overlapping gene encoded by the *Cucumber mosaic virus* genome. *Virology* **198**: 593–601

Ding SW, LiW-X, Symons RH (1995a) A novel naturally occurring hybrid gene encoded by a plant RNA virus facilitates long distance virus movement. *EMBO J* **14**: 5762–5772

Ding B, Li Q, Nguyen L, Palukaitis P, Lucas WJ (1995b) *Cucumber mosaic virus* 3a protein potentiates cell-to-cell trafficking of CMV RNA in tobacco plants. *Virology* **207**: 345–353

Diveki Z, Salanki K, Balazs E (2004) The necrotic pathotype of the *Cucumber mosaic virus* (CMV) ns strain is solely determined by amino acid 461 of the 1a protein. *Mol Plant Microbe Interact* **17**: 837-845

Domingo E, Holland JJ (1997) RNA virus mutations and fitness for survival. *Annu Rev Microbiol* **51**: 151-178

Duprat A, Caranta C, Revers F, Menand B, Browning KS, Robaglia C (2002) The Arabidopsis eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. *Plant J* **32**: 927-934

Edwards MC, Gonsalves D, Provvidenti R (1983) Genetic analysis of C*ucumber mosaic virus* in relation to host resistance: Location of determinants for pathogenicity to certain legumes and *Lactuca saligna*. *Phytopathology* **73**: 269-273

Edwardson JR, Christie RG (1991). Cucumoviruses, In: CRC Handbook of viruses infecting legumes. CRC Press, Boca Raton, Florida. pp 293–319

Efron B, Halloran E, Holmes S (1996) Bootstrap confidence levels for phylogenetic trees. *Proc Natl Acad Sci* USA **93**:7 085-7090

Eigen M (1993) Viral quasispecies. *Sci Am* **269**: 42-49

Erickson FL, Holzberg S, Calderon-Urrea A, Handley V, Axtell M, Corr C, Baker B (1999) The helicase domain of the TMV replicase proteins induces the N-mediated defense response in tobacco. *Plant J* **18**: 67-75

Fire A (1999) RNA-triggered gene silencing. *Trends Genet* **15**: 358-363

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**: 806-811

Fraile A, Alonso-Prados JL, Aranda MA, Bernal JJ, Malpica JM, Garcia-Arenal F (1997) Genetic exchange by recombination or reassortment is infrequent in natural populations of a tripartite RNA plant virus. *J Virol* **71**: 934-940

Fraser RSS (1990) The genetics of resistance to plant-viruses. *Annu Rev Phytopathol* **28**: 179-200

Fraser RSS (1992) The genetics of plant-viruses interaction: implications for plant breeding. *Euphytica* **63**: 175-185

Fraser RSS (1999) Plant resistance to viruses. In: *Encyclopedia of Virology* (A Granoff and RG Webster, eds.) Academic Press, San Diego, California. pp 1300-1307

Gallitelli D (2000) The ecology of *Cucumber mosaic virus* and sustainable agriculture. *Virus Res* **71**: 9-21

Gal-On A, Kaplan I, Roossinck MJ, Palukaitis P (1994) The kinetics of infection of zucchini squash by *Cucumber mosaic virus* indicate a function for RNA 1 in virus movement. *Virology* **205**: 280-328

Gal-On A, Kaplan I, Palukaitis P (1995) Differential effects of satellite RNA on the accumulation of *Cucumber mosaic virus* RNAs and their encoded proteins in tobacco vs. zucchini squash with two strains of CMV helper virus. *Virology* **208**: 58–66

Garcia-Arenal F, Palukaitis P (1999) Structure and functional relationships of satellite RNAs of *Cucumber mosaic virus*. *Curr Top Microbiol Immunol* **239**: 37-63

Garcia-Arenal F, Escriu F, Aranda MA, Alonso-Prados JL, Malpica JM, Fraile A (2000) Molecular epidemiology of *Cucumber mosaic virus* and its satellite RNA. *Virus Res* **71**: 1-8

Gera A, Loebenstein G, Raccah B (1979) Protein coats of two strains of *Cucumber mosaic virus* affect transmission by *Aphis gossypii*. *Phytopathology* **69**: 396–399

Gerhold D, Rushmore T, Caskey CT (1999) DNA chips: promising toys have become powerful tools. *Trends Bioch Sci* **24**: 168-173

Germundsson A, Sandgren M, Barker H, Savenkov EI, Valkonen JPT (2002) Initial infection of roots and leaves reveals different resistance phenotypes associated with coat protein gene-mediated resistance to *Potato mop-top virus*. *J Gen Virol* **83**: 1201–1209

Graves MV, Roossinck MJ (1995) Characterization of defective RNAs derived from RNA 3 of the Fny strain of *Cucumber mosaic cucumovirus*. *J Virol* **69**: 4746–4751

Graves MV, Pogany J, Romero J (1996). Defective interfering RNAs and defective viruses associated with multipartite RNA viruses of plants. *Semin Virol* **7**: 399–408

Green SK (2000) CMV/ChiMV resistance identified and characterized. In: AVRDC Report 1999. Asian Vegetable Research and Development Center. Shanhua, Tainan, Taiwan. 19-20

Green SK (2001) CMV/ChiMV resistance identified and characterized. In: AVRDC Report 2000. Asian Vegetable Research and Development Center. Shanhua, Tainan, Taiwan. 21-22

Green SK, Kim JS (1991) Characterization and control of pepper viruses: A literature review. *AVRDC Tech Bull* No. 18, Shanhua, Taiwan.

Greene AE, Allison RF (1994) Recombination between viral RNA and transgenic plant transcripts. *Science* **263**: 1423-1425

Greenleaf WH (1986) Pepper breeding. In: Breeding Vegetable Crops. (MJ Bassett, ed.) AVI Publishing Co., Westport, USA. pp 69-134

Grube RC, Zhang Y, Murphy JF, Loaiza-Figueroa F, Lackney VK, Provvidenti R, Jahn MK (2000). New source of resistance to *Cucumber mosaic virus* in *Capsicum frutescens*. *Plant Dis* **84**: 885-891

Guo HS, Garcia JA (1997) Delayed resistance to *Plum pox potyvirus* mediated by a mutated RNA replicase gene: involvement of a gene silencing mechanism. *Mol Plant Microbe Interact* **10**:160–170

Habili N, Symons RH (1989) Evolutionary relationship between luteoviruses and other RNA plant viruses based on sequence motifs in their putative RNA polymerases and nucleic acid helicase. *Nucleic Acid Res* **17**: 9543-9555

Hadden JF, Black LL (1987) Comparison of virulence of tomato and pepper isolates of *Colletotrichum* spp. Abstracts of the Annual Meeting of The American Phytopathological Society, Southern Division, 1-4th February 1987, Nasheville, USA. *Phytopathology* **77**: 641

Hadden JH, Black LL (1989) Anthracnose of pepper caused by *Colletotrichum* spp. In: (SK Green, ed.) Tomato and Pepper Production in the Tropics. AVRDC, Taiwan. pp 189-199

Hamamoto H, Watanabe Y, Kamada H, Okada Y (1997) Amino acid changes in the putative replicase to *Tomato mosaic tobamovirus* that overcome resistance in Tm-1 tomato. *J Gen Virol* **78**: 461-464

Hanada K, Tochihara H (1980) Genetic analysis of C*ucumber mosaic, Peanut stunt* and C*hrysanthemum mild mottle viruses*. *Ann Phytopathol Soc Jpn* **46**: 159-168

Hasse A, Richter J, Rabenstein F (1989) Monoclonal antibodies for detection and serotyping of *Cucumber mosaic virus*. *J Phytopathol* **127**: 129-136

Hayes RJ, Buck KW (1990) Complete replication of a eukaryotic virus RNA *in vitro* by a purified RNA-dependent RNA polymerase. *Cell* **63**: 363-368

Heinze C, Letschert B, Hristova D, Yankulova M, Kauadjouor O, Willingmann P, Atanassov A, Adam G (2001) Variability of the N-protein and the intergenic region of the S RNA of *Tomato spotted wilt tospovirus* (TSWV). *New Microbiol* **24**: 175-187

Hellwald K-H, Palukaitis P (1995) Viral RNA as a potential target for two independent mechanisms of replicase-mediated resistance against *Cucumber mosaic virus*. *Cell* **83**: 937–946

Higuchi R, Krummel B, Saiki RK (1998) A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic acids Res* **16**: 7351-7367

Hillis DM, Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst Biol* **42**: 182-192

Hristova D, Yordanova A, Mavrodieva V (2002) Differentiation of Bulgarian isolates from *Cucumber mosaic virus* by serological methods. *J Phytopathol* **150**: 334–339

Hsu YH, Wu CW, Lin BY, Chen HY, Lee MF, Tsai Ch (1995) Complete genomic RNA sequences of *Cucumber mosaic virus* strain NT9 from Taiwan. *Arch Virol* **140**: 1841-1847

Hsu HT, Barzuna L, Bliss W, Perry KL (2000) Identification and subgrouping of *Cucumber mosaic virus* with mouse monoclonal antibodies.*Virology* **90**: 615–620

Hull R (2002) Matthew's Plant Virology. Academic Press, San Diego and London

Huttinga H (1996) Sensitivity of indexing procedures for viruses and viroids In: *Advances in Botanical Research,* vol 23 (SH Deboer, JH Andrews, IC Tommerup, eds.). Academic Press Ltd., UK. pp 59-71

Ishihama A, Barbier P (1994) Molecular anatomy of viral RNA-directed RNA polymerases. *Arch Virol* **134**: 235–258

Ji LH, Ding SW (2001) The suppressor of transgene RNA silencing encoded by *Cucumber mosaic virus* interferes with salicylic acid-mediated virus resistance. *Mol Plant Microbe Interact* **14**: 715-724

Johansen IE, Lund OS, Hjulsager CK, Laursen J (2001) Recessive resistance in *Pisum sativum* and Potyvirus pathotype resolved in a gene-for-cistron correspondence between host and virus. *J Virol* **75**: 6609-6614

Jones JDG (2001) Putting knowledge of plant disease resistance genes to work. *Curr Opin Plant Biol* **4**: 281–287

Kadare G, Haenni AL (1997) Virus-encoded RNA helicases. *J Virol* **71**: 2583–2590

Kane MD, Jatkoe TA, Stumpf CR, Lu J, Thomas JD, Madore SJ (2000) Assessment of the sensitivity and specificity of oligonucleotide (50 mer) microarrays. *Nucleic Acids Res* **15**: 4552-4557

Kaplan IB, Shintaku MH, Li Q, Zhang L, Marsh LE, Palukaitis P (1995) Complementation of virus movement in transgenic tobacco expressing the *Cucumber mosaic virus* 3a gene. *Virology* **209**: 188–199

Kaplan IB, Lee KC, Canto T, Wong SM, Palukaitis P (2004) Host-specific encapsidation of a defective RNA 3 of Cucumber mosaic virus. *J Gen Virol* **85**: 3757-3763

Karasawa A, Okada I, Akashi K, Chida Y, Hase S, Nakazawa-Nasu Y, Ito A, Ehara Y (1999) One amino acid change in *Cucumber mosaic virus* RNA polymerase determines virulent/avirulent phenotypes on cowpea. *Phytopathology* **89**: 1186-1192

Kasschau KD, Carrington JC (2001) Long-distance movement and replication maintenance functions correlate with silencing suppression activity of potyviral HC-Pro. *Virology* **285**: 71-81

Katomaeda M, Gao Q, Small PM (2001) Microarray analysis of pathogens and their interaction with hosts. *Cell Microbiol* **3**: 713-719

Kellam P (2001) Post-genomic virology: the impact of bioinformatics, microarrays and proteomics on investigating host and pathogen interactions. *Rev Med Virol* **11**: 313-329

Keller KE, Johansen IE, Martin RR, Hampton RO (1998) Potyvirus genome-linked protein (VPg) determines *Pea seed-borne mosaic virus* pathotype-specific virulence in *Pisum sativum*. *Mol Plant-Microbe Interact* **11**: 124-130

Kim C-H, Palukaitis P (1997) The plant defense response to *Cucumber mosaic virus* in cowpea is elicited by the viral polymerase gene and affects virus accumulation in single cells. *EMBO J* **16**: 4060-4068

Kimura, M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**: 111–120

King HC, Sinha AA (2001) Gene expression profile analysis by microarrays-promise and pitfalls. *J Am Med Assoc* **286**: 2280-2288

Kumar S, Tamura K, Jakobson IB, Nei M (2001) MEGA 2: molecular evolutionary genetics analysis. Version 2.1. *Bioinformatics* **17**:1244-1245

Lai MMC (1992) RNA recombination in animal and plant viruses. *Microbiol Rev* **56**: 61-79

Lapidot M, Paran I, Ben-Joseph R, Ben-Harush S, Pilowsky M, Cohen S, Shifriss C (1997) Tolerance to *Cucumber mosaic virus* in pepper: Development of advanced breeding lines and evaluation of virus level. *Plant Dis* **81**:185-188

Leach JE, Vera Cruz CM, Bai J, Leung H (2001) Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Annu Rev Phytopathol* **39**: 187–224

Lee GP, Min BE, Kim CS, Choi SH, Harn CH, Kim SU, Ryu KH (2003) Plant virus cDNA chip hybridization for detection and differentiation of four cucurbit-infecting Tobamoviruses. *J Virol Methods* **110**: 19-24

Lellis AD, Kasschau KD, Whitham SA, Carrington JC (2002) Loss-of-susceptibility mutants of *Arabidopsis thaliana* reveal an essential role for eIF(iso)4E during Potyvirus infection. *Curr Biol* **12**: 1046-1051

Letschert B, Adam G, Lesemann D-E, Willingmann P, Heinze C (2002) Detection and differentiation of serologically cross-reacting tobamoviruses of economical importance by RT-PCR and RT-PCR-RFLP. *J Virol Methods* **106**: 1-10

Li Q, Palukaitis P (1996) Comparison of the nucleic acid- and NTP-binding properties of the movement protein of *Cucumber mosaic cucumovirus* and *Tobacco mosaic tobamovirus*. *Virology* **216**: 71–79

Li WX, Ding SW (2001) Viral suppressors of RNA silencing. *Curr Opin Biotechnol* **12**:150-154

Li Q, Ryu KH, Palukaitis P (2001) *Cucumber mosaic virus*-plant interactions: identification of 3a protein sequences affecting infectivity, cell-to-cell movement, and long-distance movement. *Mol Plant Microbe Interact* **14**: 378-385

Li H, LiWX, Ding SW (2002) Induction and suppression of RNA silencing by an animal virus. *Science* **296**: 1319–1321

Lin HX, Rubio L, Smythe A, Jiminez M, Falk BW (2003) Genetic diversity and biological variation among California isolates of *Cucumber mosaic virus*. *J Gen Virol* **84**: 249-258

Lin HX, Rubio L, Smythe AB, Falk BW (2004) Molecular population genetics of *Cucumber mosaic virus* in California: evidence for founder effects and reassortment. *J Virol* **78**: 6666-6675

Lin NS, Hsu YH, Hsu HT (1990) Immunological detection of plant viruses and a mycoplasmalike organism by direct tissue blotting on nitrocellulose membranes. *Phytopathology* **80**: 824-828

Lindbo JA, Silva-Rosales L, Proebsting WM, Dougherty WG (1993) Induction of a highly specific antiviral state in transgenic plants: implications for regulation of gene expression and virus resistance. *Plant Cell* **5**: 1749–1759

Liu S, He X, Park G, Josefsson C, Perry KL (2002) A conserved capsid protein surface domain of *Cucumber mosaic virus* is essential for efficient aphid vector transmission. *J Virol* **76**: 9756-9762

Lopez MM, Bertolini E, Olmos A, Caruso P, Gorris MT, Llop P, Penyalver R, Cambra M (2003) Innovative tools for detection of plant pathogenic viruses and bacteria. *Int Microbiol* **6**: 233-243

Lot H, Kaper JM (1976) Further studies on the RNA component distribution among the nucleoproteins of *Cucumber mosaic virus*. *Virology* **74**: 223-226

Marathe R, Anandalakshmi R, Smith T, Pruss G, Vance V (2000) RNA viruses as inducers, suppressors and targets of posttranscriptional gene silencing. *Plant Mol Biol* **43**: 295-306

Marathe R, Guan Z, Anandalakshmi R, Zhao H, Dinesh-Kumar SP (2004) Study of *Arabidopsis thaliana* resistome in response to *Cucumber mosaic virus* infection using whole genome microarray. *Plant Mol Biol* **55**: 501-520

Marchoux G, Marrou J, Devergne JC, Quiot JB, Douine L, Lot H (1975). *Cucumber mosaic virus* hybrids constructed by exchanging RNA components. *Meded Fac Landbouwwet Rijksuniv Gent* **40**: 59-72

Marchuk D, Drumm M, Saulino A, Collins FS (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* **19**: 1154.

Marck C (1988) DNA Strider: a "C" program for the fast analysis of DNA and Protein sequences on the Apple Macintosh family computers. *Nucleic Acids Res* **16**: 1829-1836

Martelli, GP and Russo M. 1985. In "The plant viruses. Vol.1: Polyhedral Virions with Tripartite Genomes" (R.I.B.Francki, ed.) pp.163-205.Plenum, New York

Mastari J, Lapierre H, Dessens JT (1998) Asymmetrical distribution of *Barley yellow dwarf virus* PAV variants between host plant species. *Phytopathology* **88**: 818-821

McKinney HH (1935) Evidence of virus mutation in the common mosaic of tobacco. *J Agric Res* **51**: 951-981

Meshi T, Motoyoshi F, Adachi A, Watanabe Y, Takamstsu N, Okada Y (1988) Two concomitant base substitutions in the putative replicase genes of *Tobacco mosaic virus* confer the ability to overcome the effects of a tomato resistance gene, tm-1. *EMBO J* **7**: 1575-1581

Meshi T, Motoyoshi F, Maeda T, Yoshiwoka S, Watanabe H, Okada Y (1989). Mutations in the *Tobacco mosaic virus* 30-K protein gene overcome tm-2 resistance in tomato. *Plant Cell* **1**: 515-522

Miller ED, Hemenway C (1998) History of coat protein-mediated protection. *Methods Mol Biol* **81**: 25-38

Monma S, Sakata Y (1997) Screening of Capsicum lines for resistance to *Cucumber mosaic virus*. *J Jpn Soc Hortic Sci* **65**: 769-776

Moury B (2004) Differential Selection of Genes of *Cucumber Mosaic Virus* Subgroups. *Mol Biol Evol* **21**:1602–1611

Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell* **2**: 279-289

Nicaise V, German-Retana S, Sanjuán R, Dubrana M-P, Mazier M, Maisonneuve B, Candresse T, Caranta C, LeGall O (2003) The eukaryotic translation initiation factor 4E controls lettuce susceptibility to the Potyvirus *Lettuce mosaic virus*. *Plant Physiol* **132**: 1272-1282

Nicolas O, Dunnington SW, Gotow LF, Pirone TP, Hellmann GM (1997) Variations in VPg protein allow a potyvirus to overcome via gene resistance in tobacco. *Virology* **237**: 452-459

Nono-Womdim R, Gebre-Selassie K, Polloix A, Pochard E, Marchoux G (1993a) Study of multiplication of *Cucumber mosaic virus* in susceptible and resistant *Capsicum annuum* lines. *Ann Appl Biol* **122**: 49–56

Nono-Womdim R, Polloix A, Gebre-Selassie K, Marchoux G (1993b) Partial resistance of bell pepper to *Cucumber mosaic virus* movement within plants: field evaluation of its efficiency in southern France. *J Phytopathol* **137**: 125–132

O'Reilly EK, Kao CC (1998) Analysis of RNA-dependent RNA-polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. *Virology* **252**: 287–303

Padgett HS, Watanabe Y, Beachy RN (1997) Identification of the TMV replicase sequence that activates the N gene-mediated hypersensitive response. *Mol Plant-Microbe Interact* **10**: 709-715

Palukaitis P, Garcia-Arenal F (2003) Cucumoviruses. *Adv Virus Res* **62**: 241-323

Palukaitis P, Zaitlin M (1997) Replicase-mediated resistance to plant virus disease. *Adv Virus Res* **48**: 349-377

Palukaitis P, Roossinck MJ, Dietzgen RG, Francki RI (1992) *Cucumber mosaic virus*. *Adv Virus Res* **41**: 281-348

Pearson WR (1990) Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol* **83**: 63-98

Pochard E, Daubèze AM (1989) Progressive construction of a polygenic resistance to *Cucumber mosaic virus* in the pepper. In: *Proc 7th EUCARPIA Meet Genet Breed Capsicum Eggplant.* Kragujevac, Yougoslavia. pp 187-192

Perry KL, Habili N, Dietzgen RG (1993) A varied population of *Cucumber mosaic virus* from peppers. *Plant Pathol* **42**: 806–810

Porta C, Devergne JC, Cardin L, Briand JP, van Regenmortel MHV (1989) Serotype specificity of monoclonal antibodies to *Cucumber mosaic virus*. *Arch Virol* **104**: 271-285

Pringle CR (1996) Genetics and genome segment reassortment. In*:* The Bunyaviridae (RM Elliott, ed.). Plenum, New York. pp. 189–226

Qiu W, Moyer JW (1999) Tomato spotted wilt *tospovirus* adapts to the TSWV N gene-derived resistance by genome reassortment. *Phytopathology* **89**: 575-582

Qiu WP, Geske SM, Hickey CM, Moyer JW (1998) *Tomato spotted wilt tospovirus* genome reassortment and genome segment-specific adaptation. *Virology* **244**: 186-194

Qu F, Morris TJ (2002) Efficient infection of *Nicotiana benthamiana* by *Tomato bushy stunt virus* is facilitated by the coat protein and maintained by p19 through suppression of gene silencing. *Mol Plant Microbe Interact* **15**: 193–202

Rao ALN, Francki RIB (1982) Distribution of determinants for symptom production and host range on the three RNA components of *Cucumber mosaic virus*. *J Gen Virol* **61**: 197–205

Ratcliff FG, Harrison BD, Baulcombe DC (1997) A similarity between viral defense and gene silencing in plants. *Science* **276**: 1558–1560

Ratcliff FG, MacFarlane SA, Baulcombe DC (1999) Gene silencing without DNA: RNA-mediated cross-protection between viruses. *Plant Cell* **11**: 1207–1215

Ribicki EP (1995) Family *Bromoviridae*. In: Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses (FA Murphy, CM Fauquet, DHL Bishop, SA Ghabrial, AW Jarvis, GP Martelli, MA Mayo, MD Summers, eds.). Springer-Verlag, Wien. pp 450–457

Rizzo TM, Palukaitis P (1990) Construction of full-length cDNA clones of *Cucumber mosaic virus* 1,2 and 3: generation of infectious RNA transcripts. *Mol Gen Genet* **222**: 249-256

Rizos H, Gunn LV, Pares RD, Gillings MR (1992) Differentiation of *Cucumber mosaic virus* isolates using the polymerase chain reaction. *J Gen Virol* **73**: 2099-2103

Rodriguez-Alvardo G, Kurath G, Dodds JA (1995) Heterogeneity in pepper isolates of *Cucumber mosaic virus*. *Plant Dis* **79**: 450-455

Romano N, Macino G (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Mol Microbiol* **6**: 3343-3353

Roossinck MJ (1997) Mechanisms of plant virus evolution. *Annu Rev Phytopathol* **35**: 191-209

Roossinck MJ (2002) Evolutionary history of *Cucumber mosaic virus* deduced by phylogenetic analyses. *J Virol* **76**: 3382-3387

Roossinck MJ, Palukaitis P (1990) Rapid induction and severity of symptoms in zucchini squash (*Cucurbita pepo*) map to RNA 1 of *Cucumber mosaic virus. Mol Plant-Microbe Interact* **3**: 188-192

Roossinck MJ, Zhang L, Hellwald KH (1999) Rearrangements in the 5' nontranslated region and phylogenetic analyses of *Cucumber mosaic virus* RNA 3 indicate radial evolution of three subgroups. *J Virol* **73**: 6752-6758

Roth BM, Pruss GJ, Vance VB (2004) Plant viral suppressors of RNA silencing. *Virus Res* **102**: 97-108

Rott ME, Jelkmann W (2001) Characterization and detection of several filamentous viruses of cherry: Adaptation of an alternative cloning method (DOP-PCR) and modification of an RNA extraction protocol. *Eur J Plant Pathol* **107**: 411-420

Ruffel S, Dussault MH, Palloix A, Moury B, Bendahmane A, Robaglia C, Caranta C (2002) A natural recessive resistance gene against P*otato virus* Y in pepper corresponds to eukaryotic initiation factor 4E(eIF4E). *Plant J* **32**: 1067-1075

Ruiz MT, Voinnet O, Baulcombe DC (1998) Initiation and maintenance of virus-induced gene silencing. *Plant Cell* **10**: 937–946

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**:406-425

Salánki K, Carrère I, Jacquemond M, Balázs E, Tepfer M (1997) Biological properties of pseudorecombinant strain created with *Cucumber mosaic virus* and *Tomato aspermy virus*. *J Virol* **71**: 3597–3602

Sambrook J, Fritsch EF, Maniatis F (1989) Molecular cloning: A Laboratory Manual. Cold Spring Harbour Laboratory Press, New York.

Sanford JC, Johnson SA (1985) The concept of pathogen-derived resistance. *J Theoret Biol* **113**: 395-405

Schaad MC, Carrington JC (1996) Suppression of long-distance movement of *Tobacco etch virus* in a non-susceptible host. J. Virol., 70:2556-2561

Schena M, Shalon D, Heller R, Chai A, Brown PO, Davies RW (1996) Parallel human genomic analysis: microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci* USA **93**: 10614-10619

Schena M, Heller RA, Theriault TP, Konrad K, Lachenmeier E, Davies RW (1998) Microarrays: biotechnologie's discovery platform for functional genomics. *Trends Biotechnol* **16**: 301-306

Schneider WL, Roossinck MJ (2000) Evolutionarily related Sindbis-like plant viruses maintain different levels of population diversity in a common host. *J Virol* **74**: 3130-3134

Schneider WL, Roossinck MJ (2001) Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions. *J Virol* **75**: 6566-6571

Schwach, F, Adam G, Heinze C (2004) Expression of a modified nucleocapsid-protein of *Tomato spotted wilt virus* (TSWV) confers resistance against TSWV and *Groundnut ringspot virus* (GRSV) by blocking systemic spread. *Mol Plant Pathol* **5**: 309-316

Shi BJ, Palukaitis P, Symons RH (2002) Differential virulence by strains of *Cucumber mosaic virus* is mediated by the 2b gene. *Mol Plant Microbe Interact* **15**: 947-955

Shi BJ, Miller J, Symons RH, Palukaitis P (2003) The 2b protein of cucumoviruses has a role in promoting the cell-to-cell movement of pseudorecombinant viruses. *Mol Plant Microbe Interact* **16**: 261-267

Shintaku M, Palukaitis P (1990) Genetic mapping of *Cucumber mosaic virus*. In: *Viral Genes and Plant Pathogenesis* (TP Pirone and JG Shaw, eds.) Springer, New York. pp 156-164

Simon AE, Nagy PD (1996) RNA recombination in *Turnip crinkle virus*: its role in formation of chimeric RNAs, multimers and in 39-end repair. *Semin Virol* **7**: 373–379

Singh J, Thakur MR (1979). Reaction of some pepper lines (*Capsicum annuum* L.) to *Cucumber mosaic virus*. *Indian J Micro Plant Pathol* **9**: 276

Smith PG, Villalon, B, Villa PL (1987) Horticultural classification of peppers grown in the United States. *HortScience* 22: 11-13

Soards AJ, Murphy AM, Palukaitis P, Carr JP (2002) Virulence and differential local and systemic spread of *Cucumber mosaic virus* in tobacco are affected by the CMV 2b protein. *Mol Plant Microbe Interact* **15**: 647-653

Suzuki M, Kuwata S, Kataoka J, Masuta C, Nitta N, Takanami Y (1991) Functional analysis of deletion mutants of *Cucumber mosaic virus* RNA3 using an *in vitro* transcription system. *Virology* **183**: 106–113

Suzuki K, Kuroda T, Miura Y, Murai J (2003) Screening and field trails of virus resistant sources in *Capsicum* spp. *Plant Dis* **87**: 779-783

Szittya G, Burgyan J (2001) *Cymbidium ringspot tombusvirus* coat protein coding sequence acts as an avirulent RNA. *J Virol* **75**: 2411-2420

Takeshita M, Suzuki M, Kuwata S, Takanami Y (1998) Involvement of *Cucumber mosaic cucumovirus* RNA2 and RNA3 in viral systemic spread in radish plant. *Arch Virol* **143**: 1109–1117

Takeshita M, Kikuhara K, Kuwata S, Furuya N, Takanami Y (2004) Competition between wild-type virus and a reassortant from subgroups I and II of CMV and activation of antiviral responses in cowpea. *Arch Virol* **149**: 1851–1857

Takken FLW, Joosten MHAJ (2000) Plant resistance genes: their structure, function and evolution. *Eur J Plant Pathol* **106**: 699–713

Tenllado F, Garcia-Luque I, Serra MT, Diaz-Ruiz JR (1994) Rapid detection and differentiation of *tobamoviruses* infecting L-resistant genotypes of pepper by RT-PCR and restriction analysis. *J Virol Methods* **47**: 165-174

Thein SL, Wallace RB (1986) The use of synthetic oligonucleotides as a specific hybridization probe in the diagnosis of genetic disorders. In: *Human Genetic Diseases: a practical approach* (KE Davis, ed.), IRL Press, Herndon, Virginia. pp 33-50

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-4680

Velasco L, Janssen D, Ruiz-Garcia L, Segundo E, Cuadrado IM (2002) The complete nucleotide sequence and development of a differential assay for a *Pepper mild mottle virus* (PMMoV) isolate that overcomes L3 resistance on pepper. *J Virol Methods* **106**: 135-140

Villalon B (1981) Breeding peppers to resist virus disease. *Plant Dis* **65**: 557-562

Voinnet O (2001) RNA silencing as a plant immune system against viruses. *Trends Genet* **17**: 449-459

Voinnet O, Pinto YM, Baulcombe DC (1999) Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proc Natl Acad Sci* USA **96**:14147-14152

Wallace RW (1997) DNA on a chip-serving up the genome for diagnostics and research. *Mol Med Today* **3**: 324-384

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen NP, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES (1998) Large scale identification, mapping and genotyping of single-nucleotide polymorphisms in the human genome. *Science* **280**: 1077-1082

Waterhouse PM, Wang MB, Lough T (2001) Gene silencing as an adaptive defence against viruses. *Nature* **411**: 834-842

Weber H, Pfitzner AJP (1998) Tm-$2^{(2)}$ resistance in tomato requires recognition of the carboxy terminus of the movement protein of *Tomato mosaic virus*. *Mol Plant-Microbe Interact* **11**: 498-503

Weber H, Schultze S, Pfitzner AJ (1993) Two amino acid substitutions in *the Tomato mosaic virus* 30-kilodalton movement protein confer the ability to overcome the Tm-$2^{(2)}$ resistance gene in the tomato. *J Virol* **67**: 6432-6438

White KA (1996) Formation and evolution of *Tombusvirus* defective interfering RNAs. *Semin Virol* **7**: 409–416

White KA, Morris TJ (1999) Defective and defective interfering RNAs of monopartite plus-stranded RNA plants viruses. *Curr Top Microbiol Immunol* **239**: 1–17

Wilson TMA (1993) Strategies to protect crop plants against viruses: pathogen-derived resistance

blossoms. *Proc Natl Acad Sci* USA **90**: 3134-3141

Wong S-M, Thio SS-C, Shintaku M, Palukaitis P (1999) The rate of cell-to-cell movement in squash of *Cucumber mosaic virus* is affected by sequences of the capsid protein. *Mol Plant-Microbe Interact* **12**: 628–632

Worobey M, Holmes EC (1999) Evolutionary aspects of recombination in RNA viruses. *J Gen Virol* **80**: 2535-2543

Xin HW, Ding SW (2003) Identification and molecular characterization of a naturally occurring RNA virus mutant defective in the initiation of host recovery. *Virology* **317**: 253–262

Xu P, Roossinck MJ (2000) *Cucumber mosaic virus* D satellite RNA-induced programmed cell death in tomato. *Plant Cell* **12**: 1079-1092

Yamanaka T, Ohta T, Takahashi M, Meshi T, Schmidt R, Dean C, Naito S, Ishikawa M (2000) TOM1, an Arabidopsis gene required for efficient multiplication of a *Tobamovirus*, encodes a putative transmembrane protein. *Proc Natl Acad Sci* USA **97**: 10107-10112

Yamanaka T, Imai T, Satoh R, Kawashima A, Takahashi M, Tomita K, Kubota M, Meshi T, Naito S, Ishikawa M (2002) Complete inhibition of tobamovirus multiplication by simultaneous mutations in two homologous host genes. *J Virol* **76**: 2491-2497

Ye ZH, Varner JE (1991) Tissue-specific expression of cell wall proteins in developing soybean tissues. *Plant Cell* **3**: 23-37

Yu C, Wu J, Zhou X (2005) Detection and subgrouping of *Cucumber mosaic virus* isolates by TAS-ELISA and immunocapture RT-PCR. *J Virol Methods* **123**: 155-161

Zaitlin M Hull R (1987) Plant virus-host interactions. *Annu Rev Plant Physiol* **38**: 291-315

Zhang L, Hanada K, Palukaitis P (1994) Mapping local and systemic symptom determinants of *Cucumber mosaic cucumovirus* in tobacco. *J Gen Virol* **75**: 3185–3191

Acknowledgements

I wish to express my deep gratitude to my advisor, Prof. Dr. Günter Adam, for the possibility to join his research group and for his useful discussions, valuable advise, encouragement and support during my stay as well as his proof-reading and suggestions on preparing this thesis.

My special thanks go to Dr. Cornelia Heinze for her direct daily supervision and the helpful advices not only in science but also in different aspects of life, and also for her useful comments, kind advice, elaborate discussions and the out-of-hours critical revision during preparing this manuscript. I have to say that she is one of my best friends rather than a supervisor. She was always patient, understanding, considerate and supportive during my whole study. I hope that I could do my best in the future to ensure that the time and effort she invested in me will not be wasted.

I would like to give my gratitude to Dr. Peter Willingmann, one of the deepest impression persons in my life. He is always friendly to me, trying to make life wonderful with lots of jokes and funs and trying to help me in my experiments and life. I also appreciated for his useful discussions, advices as well as critical and careful reading during my writing this thesis.

My sincere gratitude goes to Prof. Dr. Edgar Maiss for interest in my work and valuable suggestions and to Prof Dr. Reinhard Lieberei for his many help during last three years and useful discussions.

My sincere appreciation is extended to Dr. Sylvia K. Green, who initiated this project that led my study here being possible and provided the chili resistance lines, kindly shared information as well as agreed for correcting my manuscripts of this thesis.

Prof. Dr. Peter Palukaitis kindly provided the full-length clones and gave me encouragement. Some of the isolates and antibodies used in this study were kindly provided by Dr. Max Schöfelder and Dr. Stephan Winter. I would like to thank Dr. Jürg E. Frey for designing the microarry detection method. I appreciated the useful discussion and valuable suggestions from Prof. Dr. David Baucombe and Prof. Dr. Jari P. T. Valkonen.

Doing research in the plant virology group of Hamburg University has been an unforgettable and enjoyable experience. For the friendship, helpful comments and nice working atmosphere I am grateful to Dr. Klaus von Schwartzenberg, Marta Fernández-Núñez, Judith Mehrmann, Sigrid Geroge and Semra Ünsal. Special thanks go to Dr. Malgorzata Sadowska-Rybak and Dr. Frank Schwach for helping me "*how to survive in Hamburg*……". Dr. Frank Schwach also spent lots of time for discussing my results.

I would like to present honest thanks to my previous adviser Prof. Dr. Yong Liu for his support, for understanding my decision to come to Germany, and for continuous help and encouragement throughout my phD program; to Prof. Dr. Longping Yuan and Prof. Dr. Bida Gao for their help and continuous encouragement.

Many thanks go to my Chinese friends in Hamburg for their pleasant company while being abroad, particularly to Dr. Qing Wang and Dr. Yu Shi.

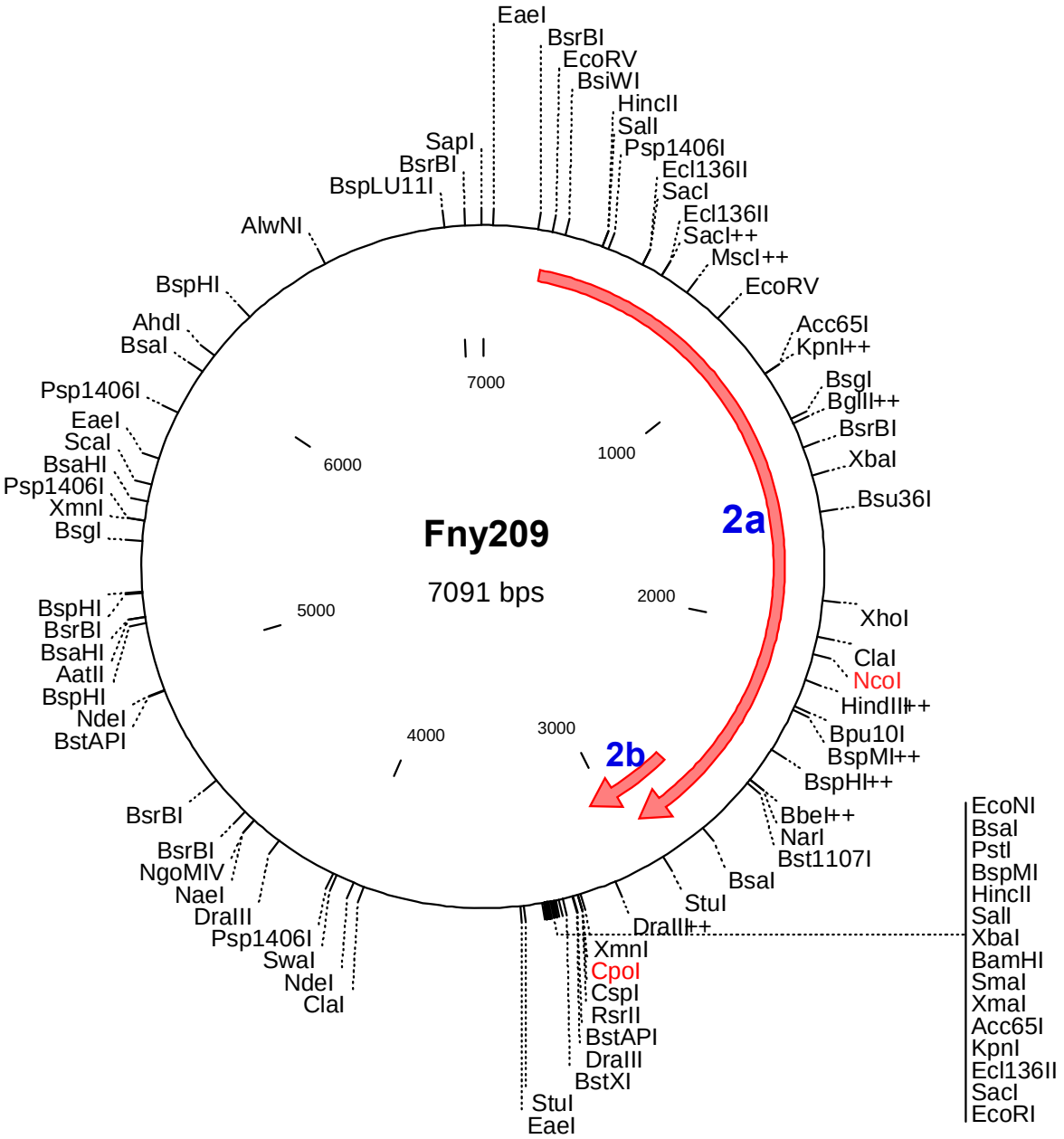
I wish to express my deepest gratitude to my wife Jing Dai and our coming baby. Thanks a million for standing by me regardless of what happened, supporting me, believing me and encouraging me. My great appreciation goes to my whole family for their endless love, supporting and understanding throughout these years, especially to my parents-in law.

Finally I wish to extend my sincere gratitude to the scholarships from GTZ (from June, 2002 to May, 2004) and from ICSC-World Laboratory (from June, 2004 to May, 2005).

The research presented in this dissertation was made possible through an international project “Development of locally-adapted, multiple disease-resistant, high-yielding chilli (*Capsicum annuum*) cultivars for targeted countries in Asia” (Project No.: 2001.7860.8-001.00/Contract No.: 81051899) funded by the Deutsche Gesellschaft fuer Technische Zusammenarbeit GmbH (GTZ, Germany) and a project “Ansätze zur Verbesserung der Produktion von Chili”(project No.07/2004) funded by the Vater und Sohn Eiselen-Stiftung (Ulm, Germany).

7 Appendix

*7.1 Map of the full length clone Fny 209**



*The 2a and 2b ORF are indicated with red arrows, respectively. The restriction enzyme sites (NcoI and CpoI) for exchanging the cassette between AN and this plasmid were indicated in red.

7.2 Sequences

7.2.1 CP sequences

>**A1**
ATGGACAAATCTGAATCTACCAGTGCTGGTCGTAATCGCCGACGTCGTCCGCGTCGCGGTTCCCGCTCCGCTT
CCTCCTCCGCGGATGCTACATTTAGAGTCCTGTCGCAACAGCTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCTACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGTTCACATCTATT
ACCCTGAAGCCACCAAAAATAGACAAAGGGTCTTATTATGGCAAAAGGTTGTTACTTCCTGATTCAGTCACTG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACTGTGTG
GGTGACGGTCCGTAAAGTTCCTGCCTCCTCGGACCTGTCCGTTTCCGCCATCTCTGCTATGTTCGCGGACGGA
GCCTCACCAGTACTGGTTTATCAGTATGCCGCATCCGGAGTTCAAGCTAACAACAAATTGTTGTATGATCTTT
CGGTGGTGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGATGAGCTAGTACTTCATGTCGACATTGAGCACCAACGCATTCCCACATCTGGAGTGCTCCCAGTTTGA
ACTCGTGTTTTCCAGAACCCTCYCTCCATTTTCTGAGGCGGGAGCTGAGTTGGTAGTGTTACTATAAACTGCC
TGAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG
>**A4**
ATGGACAAATCTGGATCTACCAGTGCTGGTCGTAATCGCCGACGTCGTCCGCGTCGCGGTTCCCGCTCCGCTT
CCTCCTCCGCGGATGCTACATTTAGAGTCCTGTCGCAACAGCTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCTACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGTTCACATCTATT
ACCCTGAAGCCACCAAAAATAGACAAAGGGTCTTATTATGGTAAAAGGTTGTTACTTCCTGATTCAGACACTG
AGTTCGGTAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACTGTGTG
GGTGACGGTCCGTAAAGTTCCTGCCTCCTCGGACCTGTCCGTTTCCGCCATCTCTGCTGTGTTCGCGGACGGA
GCCTCACCAGTACTGGTTTATCGGTATGCCGCATCCGGAGTTCAAGCTAACAACAAATTGTTGTATGATCTTT
CGGTGATGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGATGAGCTAGTACTTCATGTCGACATTGAGCACCAACGCATTCCCACATCTGGAGTGCTCCCAGTTTGA
ACTCGTGTTTTCCAGAACCCTCCCTCCATTTTCTGAGGCGGGAGCTGAGTTGGCAGTGTTACTATAAACTGCC
TGAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG
>**AN**
ATGGACAAATCTGGATCACCCAGTGCTGGTCGTAATCGCCGACGTCGTCCGCGTCGCGGTTCCCGCTCCGCTT
CCTCCTCCGCGGATGCTACATTTAGAGTCCTGTCGCAACAGCTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCTACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGTTCACATCTATT
ACCCTGAAGCCACCAAAAATAGACAAAGGGTCTTATTATGGCAAAAGGTTGTTACTTCCTGATTCAGTCACTG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACTGTGTG
GGTGACGGTCCGTAAAGTTCCTGCCTCCTCGGACCTGTCCGTTTCCGCCATCTCTGCTATGTTCGCGGACGGA
GCCTCACCAGTACTGGTTTATCAGTATGCCGCATCCGGAGTTCAAGCTAACAACAAATTGTTGTATGATCTTT
CGGTGGTGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGATGAGCTAGTACTTCATGTCGACATTGAGCACCAACGCATTCCCACATCTGGAGTGCTCCCAGTTTGA
ACTCGTGTTTTCCAGAACCCTCCCTCCATTTTCTGAGGCGGGAGCTGAGTTGGTAGTGTTACTATAAACTGCC
TGAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG
>**BRB**
ATGGACAAATCTGGATCTCCCAGTGCCGGTCGTAATCGTCGACGTCGTCCGCGTCGCGGTTCCCGCTCAGCTT
CCTCCTCCGCTGATGCTAATTTTAGAGTCCTGTCGCAACAACTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCTACTATTAACCACCCAACCTTTGTAGGGAGTGAACGTTGTAAACCTGGGTACACGTTCACATCTATT
ACCCTGAAACCACCGAAAATAGACAAAGGGTCTTATTATGGTAAAAGGTTGTTGCTTCCTGATTCAGTCACTG
AGTTCGATAAGAAGCTTGTTTCGCGCGTTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGTAAAGTTCCTGCCTCCTCGGACTTATCCGTTACTGCCATCTCTGCTATGTTTGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTACGCCGCATCCGGAGTTCAAGCCAACAATAAACTATTGTACGATCTCT
CTGCGATGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATGCGCTCGA
GACTGACGAGCTGGTACTTCATGTCGACATCGAGCACCAACGTATTCCCACATCTGGAGTGCTCCCAGTTTGA
ACCAATGTTTTCCCAGAATCCTCCCTCCAGTTCGCTACGGCGGGAGCTGAGTTGGCAGTGTTGCTATAAACTG
TCTGAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG
>**Ca**
ATGGACAAATCTGAATCTACCAGTGCTGGTCGTAATCGTCGACGTCGTCCGCGTCGCGGTTCCCGCTCCGCTT
CCTCCTCCGCGGATGCTACTTTCAGAGTCTTGTCGCAGCAACTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCAACTATTAACCACCCAACCTTTGTGTGTAGTGAGCGCTGTAAACCTGGATACACGTTCACATCTATT
ACCCTGAAACCACCGAAAATAGACCGTGGGTCTTATTATGGTAAAAGGTTGTTACTACCTGATTCAGTCACGG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGGAAAGTTCCTGCCTCCTCGGACTTATCCGTTGCCGCTATCTCTGCTATGTTTGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCCGCATCTGGCGTCCAAGCCAACAACAAACTGTTGTATGATCTTT
CAGCGATGCGCGCTGATATCGGTGACATGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGACGAACTGGTACTTCATGTCGACATTGAGCACCAACGCATTCCCACATCAGGGGTGCTCCCAGTTTGA
ATCCGTGTTTCCAGAGCCCTCTCTCCGATTTCTGTGGCGAGAGCTGAGTTGGTAGTATTGCTTTAAACTGCCT
GAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG
>**CH1**
ATGGACAAATCTGGATCAACCAGTGCCGGTCGTAATCGTCGACGTCGTCCGCGTCGCGGTTCCCGCTCCGCCT
CCTCCTCCGCGGATGCTACATTTAGAGTCCTGTCGCAACAGCTTTCGCGACTTAATAAGACGTTAGCAGCTGG

TCGTCCTACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGTTCACCTCGATT
ACCCTGAAGCCACCAAAAATAGACAAGGGTTCTTATTATGGCAAAAGGTTGTTACTTCCTGATTCAGTCACTG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACTGTGTG
GGTGACGGTTCGTAAAGTTCCTGCCTCCTCGGACCTGTCCGTTTCCGCCATCTCTGCTATGTTCGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTACGCTGCATCTGGTGTTCAAGCCAACAACAAATTGTTGTATGATCTTT
CAGTGATGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGATGAGCTAGTACTTCATGTTGACATTGAGCACCAACGCATTCCCACGTCTGGGGTGCTCCCAGTTTGA
ACTCGTGTTTTCCAGGATCCTCCCTCCGTTTTCTGTGGCGGGAGCTGAGTTGGTAGTGTTGCTATAAACTACC
TGAAGTCACTAAACGCTTCGCGGTGAACGGGTTGTCCATCCAG

>**CN03**

ATGGACAAATCTGAATCTACCAGTGCTGGTCGTAACCGTCGGCGTCGTCCGCGTCGCGGTTCCCGCTCCGCCT
CCTCCTCCGCGGATGCTAACTTTAGAGTCTTGTCGCAGCAACTTTCGCGACTTAACAAGACGTTAGCTACTGG
TCGTCCAACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGTTCACTTCGATT
ACCCTAAAGCCACCAAAGATAGACAAAGGGTCTTATTATGGCAAAAGGTTGTTGCTTCCTGATTCAGTCACAG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGCAAAGTTCCTGCCTCCTCGGACTTATCCGTTACCGCCATCTCTGCTATGTTCGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCAGCATCCGGAGTCCAAGCCAACAATAAATTGTTGTATGATCTTT
CGGCGATGCGCGCTGATATTGGCGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGACGAATTGGTACTTCATGTCGACATTGAGCACCAACGCATTCCCACATCTGGGGTGCTCCCAGTTTGA
ACCTGTGTTTTCCAGAACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGTAGTGTTACTATAAaCTGCC
TGAAGTCACTAAACGCTATGCGGTGAACGGGTTGTCCATCCAG

>**CN04**

ATGGACAAATCTGAATCAACCAGTGCTGGTCGTAACCGTCGGCGTCGTCCGCGTCGCGGTTCCCGCTCCGCCT
CCTCCTCCGCGGATGCTAACTTTAGAGTCTTGTCGCAGCAACTTTCGCGACTTAACAAGACGTTAGCTACTGG
TCGTCCAACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGCTCACTTCGATT
ACCCTAAAGCCACCAAAGATAGACAAAGGGTCTTATTATGGCAAAAGGTTGTTGCTTCCTGATTCAGTCACAG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGCAAAGTTCCTGCCTCCTCGGACTTATCCGTTACCGCCATCTCTGCTATGTTCGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCAGCATCCGGAGTCCAAGCCAACAATAAATTGTTGTATGATCTTT
CGGCGATGCGCGCTGATATTGGCGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGACGAATTGGTACTTCATGTCGACATTGAGCACCAACGCATTCCCACATCTGGGGTGCTCCCAGTTTGA
ACCTGTGTTTTCCAGAACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGTAGTGTTACTATAAACTGCC
TGAAGTCACTAAACGCTATGCGGTGAACGGGTTGTCCATCCAG

>**CN07**

ATGGACAAATCTGGATCAACCAGTGCTGGTCTAACCGTCGGCGTCGTCCGCGTCGCGGTTCCCGCTCCGCCTC
CTCCTCCGCGGATGCTAACTTTAGAGTCTTGTCGCAGCAACTTTCGCGACTTAACAAGACGTTAGCTACTGGT
CGTCCAACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGTTCACTTCGATTA
CCCTAAAGCCACCAAAGATAGACAAAGGGTCTTATTATGGCAAAAGGTTGTTGCTTCCTGATTCAGTCACAGA
GTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTGG
GTGACAGTCCGCAAAGTTCCTGCCTCCTCGGACTTATCCGTTACCGCCATCTCTGCTATGTTCGCGGACGGAG
CCTCACCGGTACTGGTTTATCAGTATGCAGCATCCGGAGTCCAAGCCAACAATAAATTGTTGTATGATCTTTC
GGCGATGCGCGCTGATATTGGCGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGAG
ACGGACGAATTGGTACTTCATGTCGACATTGAGCACCAACGCATTCCCACATCTGGGGTGCTCCCAGTTTGAA
CCTGTGTTTTCCAGAACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGTAGTGTTACTATAAaCTGCCT
GAAGTCACTAAACGCTATGCGGTGAACGGGTTGTCCATCCAG

>**HN**

ATGGACAAATCTGAATCAACCAGTGCTGGTCGTAACCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCCC
CCTCCTCCGCGGATGCTAACTTTAGAGTCTTGTCGCAGCAACTTTCGCGACTCAACAAGACGTTGGCAACTGG
TCGTCCAACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGTTCACTTCGATT
ACCCTGAAGCCACCAAAAATAGACAAAGGGTCTTATTATGGCAAAAGGTTGTTGCTTCCTGATTCAGTCACAG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTCTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGCAAAGTTCCTGCCTCGTCGGACTTATCCGTTACCGCCATCTCTGCTATGTTTGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCAGCATCCGGAGTCCAAGCCAACAATAAATTGTTGTATGATCTTT
CGGCGATGCACGCTGATATTGGTGATATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGACGAATTAGTACTCCATGTCGACATTGAGCACCAACGCATTCCCACATCTGGGGTGCTCCCAGTTTGA
ACCTGTGTTTTCCAGAACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGTAGTGTTGCTATAAACTACC
TGAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG

>**KB8**

ATGGACAAATCTGGATCACCCAGTGCCGGTCGTAATCGTCGACGTCGTCCGCGTCGCGGTTCCCGCTCCGCCT
CCTCCTCCGCGGATGCTACATTTAGAGTCCTGTCGCAACAGCTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCTACTATTAACCACCCAACCCTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGTTCACCTCGATT
ACCCTGAAGCCACCGAAAATAGACAAAGGTTCTTATTATGGCAAAAGGTTGTTACTTCCTGATTCAGTCACTG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACGGTCCGTAAAGTTCCTGCCTCCTCGGACCTGTCCGTTTCCGCCATCTCTGCTATGTTTGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTACGCTGCATCTGGCGTTCAAGCCAACAACAAATTGTTGTATGATCTTT
CAGTGATGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA

GACGGATGAGCTAGTACTTCATGTTGACATTGAGCACCAACGCATTCCCACATCTGGGGTGCTCCCAGTTTGA
ACTCGTGTTTTCCAGGATCCTCCCTCCGTTTTCTGTGGCGGGAGCTGAGTTGGTAGTGCTGCTATAAACTACC
TGAAGTCACTAAACGCTTCGCGGT

>**KS44**

ATGGACAAATCTGGATCTACCAGTGCCGGTCGTAATCGTCGACGTCGTCCGCGTCGCGGTTCCCGCTCCGCCT
CCTCCTCCGCGGATGCTACATTTAGAGTCCTGTCGCAACAGCTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCTACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGTTCACCTCGATT
ACCCTGAAGCCACCGAAAATAGACAAAGGTTCTTATTATGGCAAAAGGTTGTTACTTCCTGATTCAGTCACTG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACGGTCCGTAAAGTTCCTGCCTCCTCGGACCTGTCCGTTTCCGCCATCTCTGCTATGTTTGCGGACGGG
GCCTCACCGGTACTGGTTTATCAGTACGCTGCATCTGGCGTTCAAGCCAACAACAAACTGTTGTATGATCTTT
CAGTGATGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGATGAACTAGTACTTCATGTAGACATTGAGCACCAACGCATTCCCACGTCCGGGGTGCTCCCAGTTTGA
ACTCGTGTTTTCCAGAACCCTCCCTCCGTTTTCTGTGGCGGGAGCTGAGTTGGTAGTGTTGCTATAAaCTACC
TGAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG

>**P522**

ATGGACAAATCTGAATCTACCAGTGTCGGTCGTAGCCGTCGGCGTCGTCCGCGTCGTGGTTCCCGCTCCGCTT
CCTCCTCCGCGGATGCTAACTTTAGAGTCTTGTCGCAGCAGCTTTCGCGACTTAACAAGACGTTAGCAACTGG
TCGTCCTACTATTAACCACCCAACCTTTGTGGGTAGTGAACGCTGTAAACCTGGATACACGTTTACCTCGATT
ACCCTGAGGCCACCAAAAATAGACAAAGGGTCTTATTATGGCAAAAGGTTGTTACTTCCTGATTCAGTCACAG
AGTTCGACAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGCAAAGTTCCTGCCTCGTCGGACTTATCCGTTACCGCCATTTCTTCTATGTTCGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCAGCATCTGGAGTCCAAGCCAACAATAAGTTATTGTATGATCTTT
CGGCGGTGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGACGAATTGGTACTTCATGTCGACATTGAGCACCAACGCATTCCCACATCTGGGGTGCTCCCAGTTTGA
ACCTGTGTTTTCCAGAACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGTAGTGTTGCTATAAACTACC
TGAAGTCACTAAACGCTATGCGGTGAACGGGTTGTCCATCCAG

>**P3613**

ATGGACAAATCTGGATCAACCAGTGTTGGTCGTAACCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCTT
CCTCCTCCGCGGATGCTAACTTTAGAGTCTTGTCGCAGCAGCTTTCGCGACTTAACAAGACGTTAGCAACTGG
TCGTCCTACTATTAACCACCCAACCTTTGTGGGTAGTGAACGTTGTAAACCTGGATACACGTTTACCTCGATT
GCCCTGAGGCCACCAAAAATAGACAAAGGGTCCTATTATGGCAAAAGGTTGTTACTTCCTGATTCAGTCACAG
AGTTCGACAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGCAAAGTTCCTGCCTCGTCGGACTTATCCGTTACCGCCATTTCTTCTATGTTCGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCAGCATCTGGAGTCCAAGCCAACAATAAATTATTGTATGATCTTT
CGGCGATGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTGCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGACGAATTGGTACTTCATGTCGACATTGAGCACCAACGTATTCCCACATCTGGGGTGCTCCCAGCTTGA
ACCTGTGTTTTCCAGAACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGTAGTGTTGCTATAAACTGCC
TGAAGTCACTAAACGCTATGCGGTGAACGGGTTGTCCATCCAG

>**Peets**

ATGGACAAATCTGGATCTCCCCGTGCTGGTCGTAACCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCTT
CCTCCTCCGCGGATGCTAACTTTAGAGTCCTGTCGCAGCAGCTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCAACCATTAACCACCCAACCTTTGTGGGTAGTGAGCGATGTAAACCTGGGTACACGTTCACATCTATC
ACCCTGAAGCCACCGAAAATAGACCGTGGGTCTTATTATGGTAAAAGGTTGCTACTTCCTGATTCGGTCACGG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGTAAAGTTCCTGCCTCCTCGGACTTATCCGTTGCCGCCATCCCTGCTATGTTCGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCCGCATCCGGAGTCCAAGCTAATAATAAATTGCTGTACGATCTTT
CGGCGATGCGCGCTGATATTGGCGACATGCGAAAGTACGCCGTTCTCGTGTATTCAAAAGACGATGCTCTCGA
GACGGATGAGCTGGTACTTCATGTCGACATTGAGCACCAACGCATTCCCACATCAGGGGTGCTCCCAGTTTGA
ACTCGTGTTTTCCCAGAATCCTCCCCCCGACTACTGTGGCGGGAGCTGAGTTGGTAGTGTTGCTATAAACTGC
CTGAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG

>**PR**

ATGGACAAATCTGGATCACCcAGTGCCGGTCGTAACCGTGACGTCGTCCGCGTCGCGGTTCCCGCTCAGCTTC
CTCCTCCGCTGATGCTAATTTTAGAGTCCTGTCGCAACAACTTTCGCGACTTAATAAGACGTTAGCAGCTGGT
CGTCCTACTATTAACCACCCAACCTTTGTAGGGAGTGAACGTTGTAAACCTGGGTACACGTTCACATCTATCA
CCCTGAAACCACCGAAAATAGACAAAGGGTCTTATTATGGTAAAAGGTTGTTGCTTCCTGATTCAGTCACTGA
GTTCGATAAGAAGCTTGTTTCGCGCGTTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTGG
GTGACAGTCCGTAAAGTTCCTGCCTCCTCGGACTTATCCGTTACCGCTATCTCTGCTATGTTCGCGGACGGAG
CCTCGCCGGTACTGGTTTATCAGTACGCCGCATCCGGAGTTCAAGCTAACAATAAATTATTATATGATCTCTC
GGCGATGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATGCGCTCGAG
ACTGACGAGCTGGTACTTCATGTTGACATCGAGCATCAACGTATTCCCACATCTGGAGTGCTCCCAGTTTGAA
CCAGTGTTTTCCCAGAATCCTCCCTCCGGTCTGTAGCGGGAGCTGAGTTGGCAGTGTTGCTATAAACTGTCTG
AAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG

>**PV0418**

ATGGACAAATCTGGATCTACCAATGCTAGTAGAGCCTCCCGGCGTCGTCGCCCGCGTAGAGGTTCTCGGTCCG
CTTCTGGTGCGGATGCAGGGTTGCGTGCTTTGACTCAGCAGATGCTGAAACTCAATAAAACCCTCGCCATTGG
TCGTCCCACCCTTAACCACCCAACCTTCGTGGGTAGTGAAAGCTGTAAACCCGGTTACACTTTCACATCTATT
ACCCTGAAACCGCCTGAAATTGAGAAAGGTTCATATTTTGGTAGGAGGTTGTCTTTGCCAGATTCAGTCACGG
ACTATGATAAGAAGCTTGTTTCGCGCATTCAAATCAGGATTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTTACAGTTCGGAAAGTACCTTCATCATCCGATCTTTCCGTCGCCGCCGTCTCTGCTATGTTTGGCGATGGT
AATTCACCGGTTTTGGTTTATCAGTATGCTGCGTCCGGAGTTCAGGCCAACAATAAGTTACTTTATGACCTGT

CCGAGATGCGTGCTGATATCGGCGACATGCGTAAGTACGCCGTCCTGGTTTACTCGAAAGACGATAAACTAGA
GAAGGACGAGATTGTGCTTCATGTCGACGTCGAGCATCAACGAATTCCTATCTCACGGATGCTCCCGACTTAG
TCCGTGTGTTTACCGGCGTCCGAAGACGTTAAACTACGCTCTCAATCGCGAGTGCTGAGTTGGTAGTATTGCT
CCAAACTGCCTGAAGTCCCTAAACGTGTTGTTGCGCGGGGAACGGGTTGTCCATCCAG
>**PV0420**
ATGGACAAATCTGGATCTACCGATGCTAGTAGAGCCTCCCGGCGTCGTCGCCCGCGTAGAGGTTCTCGGTCCG
CTTCTGGTGCGGATGCAGGGTTGCGTGCTTTGACTCAGCAGATGCTGAAACTCAATAAAACCCTCGCCATTGG
TCGTCCCACCCTTAACCACCCAACCTTCGTGGGTAGTGAAAGCTGTAAACCCGGTTACACTTTCACATCTATT
ACCCTGAAACCGCCTGAAATTGAGAAAGGTTCATATTTTGGTAGGAGGTTGTCTTTGCCAGATTCAGTCACGG
ACTATGATAAGAAGCTTGTTTCGCGCATTCAAATCAGGATTAAGCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTTACAGTTCGGAAAGTACCTTCATCATCCGATCTTTCCGTCGCCGCCATCTCTGCTATGTTTGGCGATGGT
AATTCACCGGTTTTGGTTTATCAGTATGCTGCGTCCGGAGTTCAGGCCAACAATAAGTTACTTTATGACCTGT
CCGAGATGCGTGCTGATATCGGCGACATGCGTAAGTACGCCGTCCTGGTTTACTCGAAAGACGATAAACTAGA
GAAGGACGAGATTGTGCTTCATGTCGACGTCGAGCATCAACGAATTCCTATCTCACGGATGCTCCCGACTTAG
TCCGTGTGTTTACCGGCGTCCGAAGACGTTAAACTACGCTCTCAATCGCGAGTGCTGAGTTGGTAGTATTGCT
CCAAACTGCCTGAAGTCCCTAAACGTGTTGTTGCGCGGGGAACGGGTTGTCCATCCAG
>**RT6**
ATGGACAAATCTGGATCAACCAGTGCCGGTCGTAATCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCGC
CCTCCTCCGCGGATGCTAACTTTAGAGTCTTGTCGCAGCAGCTTTCGCGACTTAACAAGACGTTAGCAGCTGG
TCGTCCAACTATCAACCACCCAACCTTTGTGGGTAGTGAACGCTGTAAACCTGGGTACACGTTCACATCTATT
ACCCTGAAGCCGCCGAAAATAGACCGCGGGTCTTATTATGGTAAAAGGTTGTTGCTACCTGATTCAGTCACGG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAACTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGTAAAGTTCCCGCCTCCTCGGACTTATCCGTTGCCGCTATCTCTGCTATGTTTGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCTGCATCCGGCGTTCAAGCCAACAACAAATTGTTGTACGGTCTTT
CGGCGATGCGCGCTGATATTGGCGACATGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGACGAGTTGGTACTTCATGTTGACATCGAGCACCAACGCATTCCCACATCTGGGGTGCTCCCAGTTTGA
ATCCGTGTTTTTCCCAGAATCCTCCCTCCGGTCCTGTGGCGGGAGCTGAGTTGGTAGTGTTACTATAAACTGC
CTAAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG
>**RT52**
ATGGACAAATCTGGATCTACCAGTGCTGGTCGTAACCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCCC
CCTCCTCCGCGGATGCTAACTTTAGAGTCTTGTCGCAGCAGCTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCAACTATTAACCACCCAACCCTTGTAGGGAGTGAACGCTGTAGACCTGGGTACGCGTTCACATCTATT
ACCCTAAAGCCACCAAAAATAGACCGTGGGTCTTATTACGGTAAAAGGTTGTTACTACCTGATTCAGTCACGG
AATATGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGTAAAGTTCCTGCCTCCTCGGACTTATCCGTTGCCGCCATCTCTGCTATGTTCGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCCGCATCTGGAGTCCAAGCCAACAACAAACTGTTGTATGATCTTT
CGGCGATGCGCGCTGATATAGGTGACATGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATACGCTCGA
GACGGACGAGCTAGTACTTCATGTTGACATCGAGCACCAACGCATTCCCACATCTGGAGTGCTCCCAGTCTGA
TTCCGTATTCCCAGAATCCTCCCTCCGATCTCTGTGGCGGGAGCTGAGTTGGCAGTTCTGCTATAAACTGTCT
GAAGTCACTAAACGTTTTTACGGTGAACGGGTTGTCCATCCAG
>**RT54**
ATGGACAAATCTGAATCACCCAGTGCTGGTCGTAACCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCTT
CCACCTCCGCGGATGCCAACTTTAGAGTCCTGTCGCAGCAACTTTCGCGACTTAATAAGACGTTGGCAGCTGG
TCGTCCTACCATTAACCACCCAACCTTTGTAGGGAGTGAACGTTGTAAACCTGGGTACACGTTCACATCTATT
ACCCTGAAACCACCGAAAATAGACCGCGGGTCTTATTATGGTAAAAGGTTGTTATTACCTGATTCAGTCACGG
AATTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGTAAAGTTCCTGCCTCCTCGGACCTGTCCGTCGCCGCCATCTCTGCTATGTTTGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCTGCATCCGGAGTCCAAGCCAACAATAAACTGTTGTACGATCTTT
CGGCGATGCGCGCTGATATTGGCGACATGCGCAAGTACGCCGTTCTCGTGTATTCAAAAGACGATGCACTCGA
GACGGATGAGCTAGTACTTCATGTCGACATCGAGCACCAACGCATTCCCACATCTGGGGTGCCCCCAGTTTGA
ATCCGTGTTTCCCAGAACCCTCCCTCCAGTCCTGAGGCGGGAGCTGAGTTGGCAGTGTTGCTATAAACTGTCT
GAAGTCACTAAACGCTATGCGGTGAACGGGTTGTCCATCCAG
>**RT55**
ATGGACAAATCTGGATCTACCAGTGCTGGTCGTAATCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCTC
CCTCCTCCGCGGATGCTACTTTTAGAGTCTTGTCGCAGCAGCTTTCGCGACTCAATAAGACATTAGCAGCTGG
TCGTCCAACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGCTGTAAACCTGGATACACGTTCACATCTATT
ACCCTGAAGCCACCGAAAATAGACCGTGGGTCTTATTATGGTAAAAGGTTGTTGCTACCTGATTCAGTCACGG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGTAAAGTTCCTGCCTCCTCGGACTTATCCGTTGCCGCTATCTCTGCTATGTTTGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCTGCATCCGGCGTTCAAGCCAACAACAAATTGTTGTATGATCTTT
CAGCGATGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGACGAGTTGGTACTTCATGTCGACATTGAGCACCAACGCATTCCCACATCTGGGGTGCTCCCAGTTTGA
ATCCGTGTTTTCCAGAACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGTAGTATTGCTTTAAACTGCC
TGAAGTCACTAACGCTTTGCGGTGAACGGGTTGTCCATCCAG
>**RT67**
ATGGACAAATCTGAATCTCCCAATGCTAGTAGAACCTCCCGGCGTCGTCGCCCGCGTAGAGGTTCTCGGTCCG
CTTCTGGTGCGGATGCAGGGTTGCGTGCTTtGACTCAGCAGATGCTGAGACTCAATAAACCCTCGCCATTGGT
CGTCCCACTCTTAACCACCCAACCTTCGTGGGTAGTGAAAGCTGTAAACCCGGTTACACTTTCACATCTATTA
CCCTGAAACCGCCTGAAATTGAGAAAGGTTCATATTTTGGTAGAAGGTTGTCTTTGCCAGATTCAGTCACGGA
CTATGATAAGAAGCTTGTTTCGCGCATTCAAATCAGGATTAATCCTTTGCCGAAATTTGATTCTACCGTGTGG

GTAACAGTTCGGAAAGTACCTTCATCATCCGATCTTTCCGTCGCCGCCATCTCTGCTATGTTTGGCGATGGTA
ATTCACCGGTTTTGGTTTATCAGTATGCTGCGTCCGGAGTTCAGGCCAACAATAAATTACTTTATGACCTGTC
CGAGGTGCGTGCTGATATCGGCGACATGCGTAAGTACGCCGTCCTGGTTTACTCGAAAGACGATAAACTAGAG
AAGGACGAGATTGTACTTCATGTCGACGTCGAGCATCAACGAATTCCTATCTCACGGATGCTCCCGACTTAGT
CCGTGTGTTTACCGGCGTCCGAAGACGTTAAACTACACTCTCGATCGCGAGTGCTGAGTTGGTAGTATTGCTC
CAAACTGCCTGAAGTCCCTAAACGTGTTGTTGCGCGGGGAACGGGTTGTCCATCCAG

>**RT68**

ATGGACAAATCTGGATCTCCCAATGCTAGTAGAACCTCCCGGCGTCGTCGCCCGCGTAGAGGTTCTCGGTCCG
CTTCTGGTGCGGATGCAGGGTTGCGTGCTTTGACTCAGCAGATGCTGAGACTCAATAAAACCCTCGCCATTGG
TCGTCCCACTCTTAACCACCCAACCTTCGTGGGTAGTGAAAGCTGTAAACCCGGTTACACTTTCACATCTGTT
ACCCTGAAACCGCCTGAAATTGAGAAAGGTTCATATTTTGGTAGAAGGTTGTCTTTGCCAGATTCAGTCACGG
ACTATGATAAGAAGCTTGTTTCGCGCATTCAAATCAGGATTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTTACAGTTCGGAAAGTACCTTCATCATCCGATCTTTCCGTCGCCGCCATCTCTGCTATGTTTGGTGATGGT
AATTCACCGGTTTTGGTTTATCAGTATGCTGCGTCCGGAGTTCAGGCCAACAATAAGTTACTTTATGACCTGT
CCGAGATGCGTGCTGATATCGGCGACATGCGTAAGTACGCCGTCCTGGTTTACTCGAAAGACGATAGACTAGA
GAAGGACGAGATTGTACTTCATGTCGACGTCGAGCATCAACGAATTCCTATCTCACGGATGCTCCcGACTTAG
TCCGTGTGTTTACCGGCGTCCGAAGACGTTAAACTACACTCTCAATCGCGAGTGCTGAGTTGGTAGTATTGCT
CCAAACTGCCTGAAGTCCCTAAACGTGTTGTTGCGCGGGGAACGGGTTGTCCATCCAG

>**RT88**

ATGGACAAATCTGGATCAACCAGTGCTGGTCGTAATCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCCC
CCTCCTCCGCGGATGCTAACTTTAGAGTCTTGTCGCAGCAGCTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCAACTATTAACCACCCAACCTTTGTAGGGAGTGAACGCTGTAGACCTGGGTACACGTTCACATCTATT
ACCCTAAAGCCACCAAAAATAGACCGTGGGTCTTATTACGGTAAAAGGTTGTTATTACCTGATTCAGTCACGG
AATATGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGTAAAGTTCCTGCCTCCTCGGACTTATCCGTTGCCGCCATCTCTGCTATGTTCGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCCGCATCTGGAGTTCAAGCCAACAACAAACTGTTGTATGATCTTT
CGGCGATGCGCGCTGATATAGGTGACATGAGAAAGTACGCCGTCCTCGTGTACTCAAAAGACGATGCGCTCGA
GACGGACGAGCTAGTACTTCATGTTGACATCGAGCACCAACGCATTCCCACGTCTGGGGTGCTCCCAGTCTGA
TTTCGTGTCCCCAGAATCCTCCCTCCGATCTCTGTGGCGGGAGCTGAGTTGGCAGTTCTGCTATAAACTGTCT
GAAGTCACTAAACGTTTTACGGTGAACGGGTTGTCCATCCAG

>**RT138**

ATGGACAAATCTGGATCACCCAGTGCTGGTCGTAACCGTCGTCGTCGTCCGCGTCGTGGTTCCCGCTCCGCTT
CCTCCTCCGCGGATGCTACGTTTAGAGTCTTGTCGCAACAGCTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCTACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGCTGTAGACCTGGATACACGTTCACATCGATT
ACCCTGAAGCCACCAAAAATAGACCGTGGGTCTTATTATGGTAAAAGGTTGTCACTTCCTGATTCAGTCACTG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACGGTCCGTAAAGTTCCTGCCTCCTCGGACCTGTCCGTTGCCGCCATCTCTGCTATGTTCGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTACGCCGCATCCGGAGTCCAAGCTAACAACAAACTGTTGCATGATCTTT
CGGCGATGCGCGCTGATATTGGCGACATGAGAAAGTACGCCGTACTCGTGTATTCAAAAGACGATGCACTCGA
GACGGATGAACTAGTACTTCATGCCGACATCGAGCACCAACGTATTCCCACATCTGGGGTGCTCCCAGTTTGA
ACTCGTGTTTTCCAGAACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGCAGTGTTGCTATAAACTGTC
TGAAGTCACTAAACGCTTTGCGCTGAACGGGTTGTCCATCCAG

>**RT144**

ATGGACAAATCTGGATCACCCAGCGCCGGTCGTAACCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCTC
CCTCCTCCGCGGATGCTAACTTTAGAGTCCTGTCGCAGCAACTTTCGCGACTTAATAAGACGTTAGCAGCTGG
TCGTCCTACCATTAACCACCCAACCTTTGTAGGGAGTGAACGTTGTAAACCTGGATACACGTTCACATCTATT
ACCCTAAAGCCACCAAAAATAGACCGCGGGTCTTATTATGGTAAAAGGTTGTTATTACCTGATTCAGTTACAG
AATTCGACAAGAAACTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGTAAAGTTCCTGCCTCCTCGGACTTGTCCGTCGCCGCCATCTCTGCTATGTTTGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCTGCATCCGGAGTCCAAGCCAACAATAAATTGTTGTATGATCTTT
CGGCGATGCGCGCTGATATTGGTGACATGCGGAAGTACGCCGTACTCGTGTATTCAAAAGACGATGCACTCGA
GACGGATGAGCTAGTGCTTCATGTCGACATCGAGCACCAACGTATTCCTACATCTGGGGTGCTCCCAGTTTGA
ATCCGTGTTTCCCAGGACCCTCCCTCCAGTTTCTGAGGCGGGAGCTGAGTTGGCAGTGTTGCTATAAACTGTC
TGAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG

>**TR15**

ATGGACAAATCTGGATCTACCAGTGCTGGTCGTAACCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCTT
CCTCCTCCGCGGATGCTACTTTCAGAGTCTTGTCGCAGCAGCTTTCGCGACTTAATAAGACATTAGCAGCTGG
TCGTCCAACTATTAACCACCCAACCTTTGTGGGTAGTGAGCGTTGTAAACCTGGATACACGTTCACATCTATT
ACCCTGAAACCACCGAAAATAGACCGTGGGTCTTATTATGGTAAAAGGTTGTTGCTACCTGATTCAGTCACGG
AGTTCGATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTG
GGTGACAGTCCGGAAAGTTCCTGCCTCCTCGGACTTATCCGTTGCCGCCATCTCTGCTATGTTTGCGGACGGA
GCCTCACCGGTACTGGTTTATCAGTATGCTGCATCTGGCGTCCAAACCAACAACAAATTGTTGTATGATCTCT
CAGCGATGCGCGCTGATATTGGTGACATGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATGCGCTCGA
GACGGACGAGTTGGTACTTCATGTCGACATCGAGCACCAACGCATTCCCACATCAGGGGTGCTCCCAGTTTGA

ATCCGTGTTTTCCAGGACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGCAGTATTGCTTTAAACTGTC
TGAAGTCACTAAACGCTTTGCGGTGAACGGGTTGTCCATCCAG

7.2.2 MP sequences

>**A1**
ATgGCTTTCCAAGGTACCAGTAGGACtTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAGAAGA
TATTATTtAGCCCTGAAGCCATTAAGAAAtTGGCTACTGAGTGTGACCTAGGCCGGCATCACTGGATGCGCGC
TGATAACGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATTGCTTCCTTCTTTAAGTCT
GGATATGATGTCGGTGAATTGAGCTCTAAAGGATATATGAGCGTCCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTCTCTACTGAAGCTGAGGGATCTTTGAGAATTTACCTGGCTGATCTAGGCGATAAGGAGTTATCTCCTAT
AGATGGGCAATGCGTTTCATTACATAACCATGATCTTCCCGCTTTAGTGTCTTTCCAGCCAACGTATGATTGC
CCCATGGAAAGTGTTGAGAATCGTAAGCGGTGTTTTGCAGTTGTAATCGAGAGACATGGTTATATTGGGTATA
CCGGCACCACAGCAAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATAT
CGCTGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTTGCTCGC
CTGTTGAAGTCGCAATTGAATAACATAGAATCTTCGCGATACGTTTTAACGGATTCGAAGATTAATCAGAATG
CGCGCAGTGAGTCCGAGGATAATTTGAATGTTGAGAGCCCTCCCTTCGCAATCGGGAGTCCCTCCGCGTCCCG
CTCCGAAGCCTTCAGACCGCAGGTGGTTAACGGT
>**A4**
ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAGAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAATTGGCTACTGAGTGTGGCCTAGGCCGGCATCACTGGATGCGCGC
TGATAGCGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATTGCTTCCTTCTTTAAGTCT
GGATATGATGTCGGTGAATTGAGCTCTAAAGGATATATGAGCGTCCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTCTCTACTGAAGCTGAGGGATCTTTGAGAATTTACCTGGCTGATCTAGGCGATAAGGAGTTATCTCCTAT
AGATGGGCAATGCGTTTCATTACATAACCATGATCTTCCCGCTTTAGTGTCTTTCCAGCCAACGTATGATTGC
CCCATGGAAAGTGTTGGGAATCGTAAGCGGTGTTTTGCAGTTGTAATCGAAAGACATGGTTATATTGGGTATA
CCGGCACCACAGCAAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAATTACACTCATAT
CGCTGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAaCGAAACCGTCAGCCGTTGCTCGC
CTGTTGAAGTCGCAATTGAATAATATAGAATCTTCGCGATACGTTTTAACGGATTCGAAGATTAATCAGAATG
CGCGCAGTGAGTCCGAGGATGATTTGAAtGTTGAGAGCCCTCCCTTCGCAATCGGGAGTCCTTCCGCGTCCCG
CTCCGAAGCCTTCAGACCGCAGGTGGTTAACGGt
>**AN**
ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAGAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAATTGGCTACTGAGTGTGACCTAGGCCGGCATCACTGGATGCGCGC
TGATAACGCTATTTCAGTCCGGCCCCTCGTTCCCGACGTAACCCACGGTCGTATTGCTTCCTTCTTTAAGTCT
GGATATGATGTCGGTGAATTGAGCTCTAAAGGATATATGAGCGTCCCTCAAGTGTTGTGTGCTGCTACTCGAA
CAGTCTCTACTGAAGCTGAGGGATCTTTGAGAATTTACCTGGCTGATCTAGGCGATAAGGAGTTATCTCCTAT
AGATGGGCAATGCGTTTCATTACATAACCATGATCTTCCCGCTTTAGTGTCTTTCCAGCCAACGTATGATTGC
CCCATGGAAAGTGTTGGGAATCGTAAGCGGTGTTTTGCAGTTGTAATCGAAAGACATGGTTATATTGGGTATA
CCGGCACCACAGCAAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATAT
CGCTGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTTGCTCGC
CTGTTGAAGTCGCAATTGAATAACATAGAATCTTCGCGATACGTTTTAACGGATTCGAAGATTAATCAGAATG
CGCGCAGTGAGTCCGAGGATAATTTGAATGTTGAGAGCCCTCCCTTCGCAATCGGGAGTCCTTCCGCGTCCCG
CTCCGAAGCCTTCAGACCGCAGGTGGTTAACGGT
>**BRB**
ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAGAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATTGCTTCCTTCTTTAAATCT
GGATATGATGTCGGTGAATTGAGCTCTAAAGGATATATGAGCGTTCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTTTCTACCGATGCTGAGGGATCTTTGAGAATTTACTTGGCTGATCTAGGTGATAAGGAGTTGTCTCCTAT
AGATGGACAATGCGTTTCATTACACAACCATGATCTTCCCGCTTTGGTGTCTTTCCAGCCGACCTATGATTGT
CCGATGGAAACTGTTGGGAATCGCAAGCGGTGTTTTGCAGTCGTTATCGAAAGACATGGTTATATTGGGTATA
CCGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATAT
CGCAGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTGTTGAAGTCGCAATTGAACAACATAGAATCTTCGCAATACGTTTTAACGCATTCGAAGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGATTTAAATGTTGAGAGCCCTCCCATCGCAACCGGGAGTTCTTCCGCGTCCCGCTC
CGAAACCTTTAGACCGCAGGTGGTTAACGGT
>**CA**
ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGAACTTCAAAAGA
TACTATTTAGCCCTGATGCCATTAAGAAGATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGTGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCTTCTTTCTTTAAGTCT
GGATATGATGCTGGTGAATTGAGTTCCAAAGGATACATGAGTGTTCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTTTCCACCGATGCTGAAGGATCTTTGAGAATTTACCTAGCTGATTTAGGCGACAAGGAGTTATCTCCTAT

AGATGGACAATGCGTCACATTACATAACCATGATCTGCCCGCTTTGGTGTCTTTCCAGCCGACGTATGATTGT
CCTATGGAAACTGTTGGGAATCGTAAGCGGTGTTTTGCTGTTGTTATTGAAAGACATGGTTATATTGGGTATA
CCGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGATTTTCTTCTAAGAATAACAACTACACTCACAT
CGCAGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACAAAACCGTCAGCTGTTGCTCGC
CTGTTGAAGTTGCAATTGAACAATATAGAATCCTCGCAATACGTTTTAACGAATGCGAAGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGAACTACATGTTGAGAGCCCTCCCATCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAGCTTTCAGACCGCAGGTGGTTAACGGT

>**CH1**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAGAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGTGC
TGATAATGCGATATCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCTTCTTTCTTTAAGTCT
GGATATGATGTCGGTGAATTGTGCTCTAAAGGATATATGAGCGTTCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTTTCTACCGATGCTGAGGGATCTTTGAGAATTTACCTGGCTGATCTAGGCGACAAGGAGTTGTCTCCTAT
AGATGGGCAATGCGTTTCGTTACATAACCATGATCTCCCCGCTTTAGTGTCTTTCCAACCGACGTATGACTGT
CCTATGGAAAGTGTTGGGAACCGTAAGCGGTGTTTTGCTGTCGTTATCGAGAGACATGGTTATATTGGGTATA
CCGGCACCACAGCAAGCGTGTGTAGTAATTGGCAAGCACGATTTTCCTCTAAGAATAACAACTACACTCATAT
CGCTGCTGGGAAGACTCTAGTACTGCCATTCAATAGATTAGCTGAGCAAACGAAACCGTCAGCCGTTGCTCGT
CTGTTGAAGTCGCAATTGAATAACATAGAATCTTCGCAATACGTTTTAACGGATTCGAAGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGAATTAAATGTTGAAAGCCCTCCCGTCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAACCTTTAGACCGCAGGTGGTTAACGGT

>**CN03**

ATGGCTTTCCAAGGTACCAGCAGGACTTTAACTCAACAgTCCTCAGCGGCTACGTCTGAGAACTTCAAAAGAT
ATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTGCTGAGTGTGACCTAGGCCGGCATCATTGGATGCGAGCT
GATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGCATTGTTTCTTTCTTTAAGTCTG
GGTATGATGCTGGTGAATTGTGCTCTAAAGGATACATGAGCGTCCCTCAAGTGTTGTGTGCTGTCACTCGAAC
AGTTTCTACCGATGCTGAAGGGTCTTTGAGAATTTACCTAGCTGACTTAGGCGACAAGGAACTATCTCCCATA
GATAAGCAATGTGTTACATTACATAACCATGATCTCCCCGCTTTAGTGTCTTTCCAACCGACGTATGATTGTC
CTATGGAGACAGTTGGGAATCGGAAGCGGTGTTTTGCTGTCGTTATCGAGAGACATGGTTACGTTGGGTATAC
CGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTCTCTTCTAAGAATAACAACTACACTCATATC
GCAGCTGGGAAGACTCTAGTACTGCCATTCAATAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGCC
TATTGAAGTCGCAACTGAATAACATAGAAGCTTCGCAATACGTTTTGACGGATTCGAAGATCAATCAGAATGC
GCGCAGTGAGTCCGAGGAATTAAATGTTGAGAGCCCTCCCGCCGCAATCGGGAGTTCTTCCGCGTCCcGCTTC
GaAAcCTTCAGAcCGCAGGTGGTTAACGGT

>**CN04**

ATGGCTTTCCAAGGTACCAGCAGGACTTTAACTCAACAGTCCTCAGCGGcTACGTCTGACGAACTTCAAAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTGCTGAGTGTGACCTAGGCCGGCATCATTGGATGCGAGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGCATTGTTTCTTTCTTTAAGTCT
GGGTATGATGCTGGTGAATTGTGCTCTAAAGGATACATGAGCGTCCCTCAAGTGTTGTGTGCTGTCACTCGAA
CAGTTTCTACCGATGCTGAAGGGTCTTTGAGAATTTACCTAGCTGACTTAGGCGACAAGGAACTATCTCCCAT
AGATAAGCAATGTGTTACATTACATAACCATGATCTCCCCGCTTTAGTGTCTTTCCAACCGACGTATGATTGT
CCTATGGAGACAGTTGGGAATCGGAAGCGGTGTTTTGCTGTCGTTATCGAGAGACATGGTTACGTTGGGTATA
CCGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTCTCTTCTAAGAATAACAACTACACTCATAT
CGCAGCTGGGAAGACTCTAGTACTGCCATTCAATAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTATTGAAGTCGCAACTGAATAACATAGAAGCTTCGCAATACGTTTTGACGGATTCGAAGATCAATCAGAATG
CGCGCAGTGAGTCCGAGGAATTAAaTGTTGAGAGCCCTCCCGCCGCAATCGGGAGTTCTTCCGCGTCCcGCTT
CGAAaCCTtCAGACcGCAGGTGGtTAACGGTA

>**CN07**

ATGGCTTTCCAAGgTACCAGCAGGACtTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGAACTTCAAAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTGCTGAGTGTGACCTAGGCCGGCATCATTGGATGCGAGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGCATTGTTTCTTTCTTTAAGTCT
GGGTATGATGCTGGTGAATTGTGCTCTAAAGGATACATGAGCGTCCCTCAAGTGTTGTGTGCTGTCACTCGAA
CAGTTTCTACCGATGCTGAAGGGTCTTTGAGAATTTACCTAGCTGACTTAGGCGACAAGGAACTATCTCCCAT
AGATAAGCAATGTGTTACATTACATAACCATGATCTCCCCGCTTTAGTGTCTTTCCAACCGACGTATGATTGT
CCTATGGAGACAGTTGGGAATCGGAAGCGGTGTTTTGCTGTCGTTATCGAGAGACATGGTTACGTTGGGTATA
CCGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTCTCTTCTAAGAATAACAACTACACTCATAT
CGCAGCTGGGAAGACTCTAGTACTGCCATTCAATAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTATTGAAGTCGCAACTGAATAACATAGAAGCTTCGCAATACGTTTTGACGGATTCGAAGATCAATCAGAATG
CGCGCAGTGAGTCCGAGGAATTAAATGTTGAGAGCCCTCCCGCCGCAATCGGGAGTTCTTCCGCGTCCCGCTT
CGAAACCTTCAGACCGCAGGTGGTTAACGGT

>**HN**

ATGGCTTTCCAAGGTACCAGCAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGAACTTCAAAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTGCTGAGTGTGACCTAGGCCGGCATCATTGGATGCGAGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGCATTGTTTCTTTCTTTAAGTCT
GGGTATGATGCTGGTGAATTGTGCTCTAAAGGATATATGAGCGTCCCTCAAGTGTTGTGTGCTGTTACTCGAA

CAGTTTCCACCGACGCTGAAGGTTCTCTGAGAATTTACCTAGCTGACTTAGGCGACAAGGAACTATCTCCCAT
AGATAAACAATGCGTCACATTACATAACCATGATCTCCCCGCTTTAGTGTCTTTCCAACCGACGTATGATTGT
CCTATGGAGACAGTTGGGAATCGGAAGCGGTGTTTTGCTGTCGTTATCGAGAGACATGGTTACGTTGGGTATA
CCGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATAT
CGCAGCTGGGAAGACTCTAGTACTGCCATTCAATAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTATTGAAGTCGCAACTGAATAACATAGAAGCTTCGCAATACGTTTTGACGGATTCGAAGATCAATCAGAATG
CGCGCAGTGAGTCCGAGGAATTAAATGTTGAGAGCCCTCCCGCCGCAATCGGGAGTTCTTCCGCGTCCCGCTT
CGAAACCTTCAGACCGCAGGTGGTTAACGGT

>**KB8**

ATgGCTTTCCAAgGTACCAGTAGGACTTtAACTCAACAgTCCTCAGCGGCTACGTCCGACGATCTTCAGAAGA
TATtGtTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGTGC
TGATAATGCGATATCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCTTCTTTCTTTAAGTCT
GGATATGTTGTCGGTGAATTGTGCTCCAAAGGATATATGAGCGTTCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTTTCTACCGATGCTGAAGGATCTTTGAGAATTTACCTGGCTGATCTAGGCGACAAGGAGTTGTCTCCTAT
AGATGGACAATGCGTTTCGTTACATAACCATGATCTCCCCGCTTTAGTGTCTTTCCAACCGACGTATGATTGT
CCTATGGAAAGTGTTGGGAACCGTAAGCGGTGTTTTGCTGTCGTTATCGAGAGACATGGTTATATTGGGTATA
CCGGCACCACAGCAAGCGTGTGTAGTAATTGGCAAGCACGATTTTCCTCTAAGAATAATAACTACACTCATAT
CGCTGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCGGCCGTTGCTCGT
CTGTTGAAGTCGCAATTGAATAACATAGAATCTTCGCAATACGTTTTAACGGATTCGAAGATTAATCAAAaTG
CGCGCAGTGAGTCCGAGGATCTGAATGTTGAAAGCCCTCCCGTCGCAATCGGGAGTTCTTCCGCGTCCcGCTC
CGaAAcCTTCAGAcCGCAGGTGGTTAaCGGT

>**KRS**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAGAAGA
TCTTATTTAGCCCTGAAGCCATTAAGAAGATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCTTCCTTCTTCAAGTCT
GGATATGATGTCGGTGAATTGTGCTCTAAAGGATACATGAGCGTTCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTGTCCACTGATGCTGAGGGATCTTTGAGAATTTACCTGGCTGATCTAGGCGATAAGGAGTTATCTCCTAT
AGATGGACAATGCGTTTCATTACATAACCATGATCTTCCCGCTTTGGTGTCTTTCCAGCCGACGTATGACTGT
CCCATGGAGACTGTTGGGAATCGTAAGCGGTGTTTTGCAGTCGTTATCGAAAGACATGGTTATATTGGGTATA
CCGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATAT
CGCAGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CGGTTGAAGTCGCAATTGAATAACATAGAATCTTCGCAATACGTTTTAACGAATTCGAAGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGATTTAAATGTTGAAAGCCCTCCCATCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAGCCTTCAGACCGCAGGTGGTTAACGGT

>**KS44**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAGAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGTGC
TGATAATGCGATATCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCTTCTTTCTTTAAGTCT
GGATATGATGTCGGTGAATTGTGCTCTAAAGGATATATGAGCGTTCCTCAAGTATTGTGTGCTGTCACTCGAA
CAGTTTCTACCGATGCTGAGGGATCTTTGAGAATTTACCTGGCTGATCTAGGCGACAAGGAGTTGTCTCCTAT
AGATGGGCAATGCGTTTCATTACATAATCATGATCTCCCCGCTTTAGTGTCTTTCCAACCGACGTATGACTGT
CCTATGGAAAGTGTTGGGAACCGTAAGCGGTGTTTTGCTGTCGTTATCGAGAGACATGGTTATATTGGGTATA
CCGGCACCACAGCAAGCGTGTGTAGTAATTGGCAAGCACGATTTTCCTCTAAGAATAACAACTACACTCATAT
CGCTGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTTGCTCGT
CTGTTGAAGTCGCAATTGAATAACATAGGATCTTCGCAATACGTTTTAACGGATTCGAAGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGATTTGAATGTTGAAAGCCCTCCCGTCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAACCTTTAGACCGCAGGTGGTTAACGGT

>**P522**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAAAAGA
TTTTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCATCTTTCTTTAAGTCT
GGATATGATGTCGGTGAATTATGCTCTAAAGGATATATGAGCGTCCCTCAGGTGTTGTGTGCTGTTACTCGAA
CAGTTTCTGCCGATGCTGAAGGGTCTCTGAGAATTTACCTGGCTGATTTAGGCGACAAGGAGTTATCTCCTAT
AGATGGACAATGCGTTTCGCTGCATAACCATGACCTCCCCGCTTTAGTTTCTTTCCAACCGACGTATGATTGT
CCGATGGAAATCGTTGGGAATCGTAAGCGGTGTTTTGCTGTCGTTATCGAAAGGCATGGTTATATTGGGTATA
CCGGCACCACGGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATAT
CGCAGCCGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTGTTGAAGTCGCAATTGAACAACATAGAATCTTCGCAATACGTTTTAACGAATTCGAGGATTAATCAAAaTG
CGCGCAGTGAGTCCGAGGATCTAAATGTTGAGAGCCcTCCcACCGCAaTCGGgAGTTCTTCcGCGTCCcGCTC
CGAAaCcTTCAGACcGCAGGTGGTtAaCGGT

>**P3613**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAAAAAA
TCTTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCATCTTTCTTTAAGTCT

GGATATGATGTCGGTGAATTATGCTCTAAAGGATATATGAGCGTCCCTCAGGTGTTGTGTGCTGTTACTCGAA
CAGTTTCTGCCGATGCTGAAGGGTCTCTGAGAATTTACCTGGCTGATTTAGGCGACAAGGAGTTATCTCCTAT
AGATGGACAATGCGTTTCGCTACATAACCATGACCTTCCCGCTTTAGTTTCTTTCCAACCGACGTATGATTGT
CCGATGGAAATTGTTGGGAATCGTAAGCGGTGTTTTGCTGTCGTTATCGAAAGGCATGGTTATATTGGGTATT
CCGGCACCACGGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATAT
CGCAGCCGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTGTTGAAGTCGCAATTGAACAACATAGAATCTTCGCAATACGTTTTAACGAATTCGAAGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGATCTAAATGTTGAGAGCCCTCCCACCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAACCTTCAGACCGCAGGTGGTTAACGGT

>**P3613(DI)**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAAAAAA
TCTTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCATCTTTCGTTATCGAA
AGGCATGGTTATATTGGGTATACCGGCACCACGGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTA
AGAATAACAACTACACTCATACCGCAGCCGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAAC
GAAACCGTCAGCCGTCGCTCGCCTGTTGAAGTCGCAATTGAACAACATAGAATCTTCGCAATACGTTTTAACG
AATTCGAAGATTAATCAAAATGCGCGCAGTGAGTCCGAGGATCTAAATGTTGAGAGCCCTCCCACCGCAATCG
GGAGTTCTTCCGCGTCCCGCTCCGAAACCTTCAGACCGCAGGTGGTTAACGGT

>**Peets**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAAAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTGCTGAGTGTAACCTAGGTCGGCATCATTGGATGCGAGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATTGCTTCCTTCTTTAAGTCT
GGATATGATGCTGGTGAATTGTGCTCTAAAGGATATATGAGCGTCCCTCAAGTGTTATGTGCTGTTACTCGAA
CAGTTTCTACCGACGCTGAAGGGTCTCTGAGAATTTACCTAGCTGATTTAGGCGACAAGGAGTTACCTCCTAT
AGATGGGCAATGCGTTTCACTACATAACCATGATCTTCCCGCTTTGGTGTCTTTCCAGCCGACGTATGATTGT
CCTATGGAAACTGTTGGGAATCGCAAGCGGTGTTTTGCAGTCGTTATCGAAAGACATGGTTATGTTGGGTATA
CCGGCACCACAGCTAGCGCGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATAT
CGCGGCTGGGAAGACTCTAGTACTGCCATTCACCAGATTAGCTGAGCAAACGAAACCGCCAGCCGTCGCTCGC
CTGTTGAAGTCGCAATTGAACAACATAGAATCTTCGCAATACGTTTTAACGAATTCGAAGATCAATCAAAATG
CGCGCAGTGAGTCCGAGGAAAATGTTGAGAGCCCTCCTATCGTAATCGGGAGTTCTTCCGCGTCCCGCTCCGA
ATCCTTTAGACCGCAGGTGGTTAACGGT

>**PR**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCCGACGATCTTCAGAAAA
TCTTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATCGCTTCCTTCTTTAAGTCT
GGATATGATGTCGGTGAATTGAGCTCTAAAGGATATATGAGCGTTCCTCAAGTATTGTGTGCTGTTACTCGAA
CAGTTTCTACCGATGCTGAGGGATCTTTGAGAATTTACTTGGCGGATCTAGGCGATAAGGAGTTGTCCCCTAT
AGATGGACAATGCGTTTCATTACACAACCATGATCTTCCCGCTTTGGTGTCTTTCCAACCGACGTATGATTGT
CCGATGGAAATTGTTGGGAACCGCAAGCGGTGTTTTGCAGTCGTTATCGAAAGACATGGTTATATTGGGTATA
CCGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATGT
CGCAGCTGGGAAGACTCTAGTACTGCCGTTCAACAGATTAGCTGAGCAAACTAAACCGTCAGCCGTCGCTCGC
CTGTTGAAGTCGCAATTGAACAACATAGAATCTTCGCAATACGTTTTAACGCATTCGAAGATTAATCaAAATG
CGCGCAGTGAGTCCGAGGATTTAAATGTTGAAAGCCCTCCCATCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAGCCTTCAGAcCGCAGGTGGTTAaCGGT

>**PW**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAAAAGA
TCTTATTTAGCCCTGAAGCCATTAAGAAATTGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGAGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGTTTCCTTCTTTAAATCT
GGATATGACGTCGGTGAATTGTGTTCTAAAGGATATATGAGCGTCCCTCAGGTGTTGTGTGCCGTTACTCGAA
CGGTCTCTACCGAGGCTGAGGGATCTTTGAGAATTTACCTGGCTGATCTAGGTGACAAGGAGTTATCTCCCAT
AGATGGACAATGCGTGTCATTACACAACCATGACCTTCCCGCTTTAGTGTCTTTCCAACCGACGTATGATTGT
CCTATGGAAGCAGTTGGGAATCGTAAGCGGTGTTTTGCAGTCGTTATCGAAAGGCATGGTTATATTGGGTATA
CCGGCACCACAGCAAGCGTGTGTAGTAATTGGCAAGCACGGTTCTCATCCAAGAATAACAACTACACTCATAT
CGCTGCTGGGAAGACTCTTGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTGTTGAAGTCGCAATTGAATAGCATAGAATCTTCGCGATACGTTTTAACGGATTCGAAGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGATTTAAATGTTGAAAGCCCTCCCATCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAACCTTCAGACCGCAGGTGGTTAACGGT

>**RS30**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAGAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGTGC
TGATAATGCGATATCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCTTCTTTCTTTAAGTCT
GGATATGATGTCGGTGAATTGTGCTCTAAAGGATATATGAGCGTTCCTCAAGTATTGTGTGCTGTCACTCGAA
CAGTTTCTACCGATGCTGAGGGATCTTTGAGAATTTACCTGGCTGATCTAGGCGACAAGGAGTTGTCTCCTAT
AGATGGGCAATGCGTTTCATTACATAATCATGATCTCCCCGCTTTAGTGTCTTTCCAACCGACGTATGACTGT

CCTATGGAAAGTGTTGGGAACCGTAAGCGGTGTTTTGCTGTCGTTATCGAGAGACATGGTTATATTGGGTATA
CCGGCACCACAGCAAGCGTGTGTAGTAATTGGCAAGCACGATTTTCCTCTAAGAATAACAACTACACTCATAT
CGCTGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTTGCTCAT
CTGTTGAAGTCGCAATTGAATAACATAGGATCTTCGCAATACGTTTTAACGGATTCGAAGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGATTTGAATGTTGAAAGCCCTCCCGTCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAACCTTTAGACCGCAGGTGGTTAACGGT

>**RT6**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGAACTTCAAAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAGATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCTTCTTTCTTTAAGTCT
GGATATGATGCTGGTGAATTAAGCTCTAAAGGATATATGAGCGTTCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTTTCCACCGACGCTGAGGGATCTTTGAGAATTTACCTAGCTGATTTAGGCGACAAGGAGTTATCTCCTAT
AGATGGACAGTGCGTAACATTACATAACCATGATCTACCCGCTTTAGTGTCTTTCCAACCGACGTATGATTGT
CCTATGGAAACTGTTGGGAATCGTAAGCGGTGTTTTGCTGTTGTTATCGAAAGACATGGTTATATTGGGTATA
CCGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCACAT
CGCAGCTGGGAAGACTCTAGTACTGCCGTTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTGTTGAAGTCGCAATTGAACAACATAGAATCTTCGCAATACGTTTTAACGAATGCGAAGATCAATCAGAATG
CGCGCAGTGAGTCCGAGGAATTAAATGTTGAGAGCCCTCCCATCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAACCTTCAGACCGCAGGTGGTTAACGGT

>**RT52(DI)**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAAAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATTGCTTCTGGGAAGACTCTA
GTACTGCCTTTCAACAGATTAGCTGAGCAAACAAAACCGTCAGCTGTCTCTCGCCTGTTGAAGTCGCAATTGA
ACAACATTGAATCTTCGCAATATTTGTTAACGAATGCGAAGATTAATCAAAATGCGCGCAGTGAGTCCGAGGA
TTTAAATGTTGAGAGCCCTCCCGCCGCAATCGGGAGTTCTTCCGCGTCCCGCTCCGAAGCCTTCAGACCGCAG
GTGGTTAACGGT

>**rt54**

ATGGCTTTCCAAGGTACCAGCAGGACTTTAACCCAACAGTCCTCAGCGGCTACGTCTGACGAACTTCAAAAGA
TATTATTTAGCCCTGACGCCATTAAGAAAATGGCTGCTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATTGCTTCCTTCTTTAAGTCT
GGATATGATGCTGGTGAATTGTGCTCTAAAGGATACATGAGTGTCCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTTTCTACCGATGCTGAAGGGTCTTTGAGAATTTACTTAGCTGATCTAGGCGACAAGGAGCTATCTCCCTT
AGATAAGCAGTGCGTCACATTACATAACCATGATCTTCCCGCTTTAGTGTCTTTCCAACCGACGTATGATTGT
CCCATGGAGACAGTTGGGAATCGGAAGCGGTGTTTTGCTGTTGTTATCGAAAGACATGGTTATATTGGGTATA
CCGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATAT
CGCAGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTGTTGAAGTCGCAATTGAACAATATAGAATCTTCGCAATACGTCTTGACGGATTCGAAGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGAATTAAATGTTGAGAGCCCTCCCGCCGCAATCGGGAGTTCTTCCGCGTCCCGCAT
CGAATCCTTTAGACCGCAGGTGGTTAACGGT

>**RT55**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGAACTTCAAAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAGATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGTGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATTGCTTCCTTCTTTAAGTCT
GGATATGATGCTGGTGAATTGAGCTCTAAAGGATATATGAGCGTTCCTCAGGTGTTGTGTGCTGTTACTCGAA
CAGTTTCCACCGATGCTGAAGGATCTTTGAGAATTTACCTAGCTGATTTAGGCGACAAGGAGTTATCTCCTAT
AGATGGACGATGCGTCACATTACATAACCATGATCTTCCCGCTTTGGTGTCTTTCCAGCCGACGTATGATTGT
CCTATGGAAATTGTTGGGAATCGTAAGCGGTGTTTTGCTGTCGTCATCGAAAGACATGGTTACATTGGGTATA
CCGGTACCACAGCTAGCGTGTGTAGTAATTGGCAAGTACGATTTTCTTCTAAGAATAACAACTACACTCATAT
CGCAGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACAAAACCGTCAGCTGTCGCTCGC
CTGTTGAAGTCGCAATTGAACAACATAGAATCTTCGCAATACGTTTTAACGAATGCGAAGATCAATCAAAATG
CGCGCAGTGAGTCCGAGGAATTAAATGTTGAGAGCCCTCCCATCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAGCCTTCAGACCGCAGGTGGTTAACGGT

>**RT67**

ATGGCTTTCCAAGGTACCAGTAGGACGTTTACTCAACAGTCCTCGGCGGCGTCGTCTGACGACTTACAGAAGA
TATTATTCAGCCCCGATGCCATCAAGAAGATGGCTACTGAGTGTGACCTAGGTCGACATCATTGGATGCGCGC
GGATAACGCCATCTCTGTCAGACCTCTCGTTCCCCAAGTAACCAGTAACAATTTATTGTCTTTCTTTAAATCT
GGGTATGATGCCGGTGAATTGCGCTCTAAAGGCTATATGAGCGTTCCTCAAGTGCTGTGTGCCGTTACCAGGA
CGGTTTCTACGGATGCTGAGGGTTCTTTGAAAATTTATTTGGCTGACCTAGGTGATAAAGAATTATCCCCAAT
TGATGGGCAATGTGTTACTTTACATAATCATGAGCTCCCTGCTTTGATATCTTTCCAACCTACCTACGATTGC
CCCATGGAATTAGTTGGCAATCGGCATCGATGTTTTGCGGTAGTCGTTGAGAGACATGGTTATATTGGTTACG
GTGGTACCACTGCTAGCGTGTGTAGTAACTGGCAAGCTCAGTTTTCTTCAAAGAATAATAATTACACACACGC
CGCTGCTGGTAAGACTCTTGTGTTGCCTTACAACAGATTAGCTGAGCATTCGAAACCGTCAGCCGTTGCTCGC
CTGTTGAAGTCGCAGTTAAACAACGTTAGCTCATCGCGCTATCTTTTGCCGAACGTTGCTCTTAATCAAAATG

CGTCTGGGCACGAGTCCGAGATTTTAAACGAAAGCCCTCCCATCGCTATAGGGAGTCCGTCCGCGTCCCGTAA
CAATAGCTTCAGATCGCAGGTGGTTAACGGT

>**RT68**

ATGGCTTTCCAAGGTACCAGTAGGACGTTAACTCAACAGTCCTCGGCGGCGTCGTCTGACGACCTACAGAAGA
TATTATTCAGCCCCGATGCCATCAAGAAGATGGCTACTGAGTGTGACCTAGGTCGACATCATTGGATGCGCGC
GGATAACGCCATCTCTGTCAGACCTCTCGTTCCCCAAGTAACCAGTAACAATTTATTGTCTTTCTTTAAATCT
GGGTATGATGCCGGTGAATTGCGCTCTAAAGGCTATATGAGCGTTCCTCAAGTGCTGTGTGCCGTTACCAGGA
CGGTTTCTACGGATGCTGAGGGTTCTTTGAAAATTTATTTGGCTGACCTAGGTGATAAAGAATTATCCCCAAT
TGATGGGCAATGTGTTACTTTACATAATCATGAGCTCCCTGCTTTGATATCTTTCCAACCTACCTACGATTGC
CCCATGGAATTAGTTGGCAATCGGCATCGATGTTTCGCGGTAGTCGTTGAGAGACATGGTTATATTGGTTACG
GTGGTACCACTGCTAGCGTGTGTAGTAACTGGCAAGCTCAGTTTTCTTCAAAGAATAATAATTACACACACGC
TGCTGCTGGTAAGACTCTTGTGTTGCCTTACAACAGATTAGCTGAGCATTCGAAACCGTCAGCCGTTGCTCGC
CTGTTGAAGTCGCAGTTAAACAACGTTAGCTCATCGCGCTATCTTTTGCCGAACGTTGCTCTTAATCAAAATG
CGTCTGGGCACGAGTCCGAGATTTTAAACGAAAGCCCTCCCATCGCTATAGGGAGTCCGTCCGCGTCCCGTAA
CAATAGCTTCAGATCGCAGGTGGTTAACGGT

>**RT88**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAAAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCCATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATTGCTTCCTTCTTTAAGTCT
GGATATGATGTTGGTGAATTATGCTCAAAAGGATACATGAGCGTCCCTCAAGTGTTATGTGCTGTCACTCGAA
CCGTTTCCACTGATGCTGAAGGGTCTTTGAGAATTTACTTAGCTGATTTAGGCGACAAGGAGTTATCTCCCAT
AGATGGGCAATGCGTTTCGTTACATAACCATGATCTTCCCGCTTTGGTGTCTTTCCAACCGACGTATGATTGT
CCTATGGAAACAGTTGGGAATCGTAAGCGGTGTTTTGCTGTCGTTATTGAAAGACATGGTTACATTGGGTATA
CCGGTACCACAGCTAGCGTGTGTAGTAATTGGCAAGCAAGGTTTTCATCTAAGAATAACAACTACACTCATAT
CGCAGCTGGGAAGACTCTAGTACTGCCTTTCAACAGATTAGCTGAGCAAACAAAACCGTCAGCTGTTGCTCGC
CTGTTGAAGTCGCAATTGAACAACATTGAATCTTCGCAATATTTGTTAACGAATGCGAAGATTAATCAAAaTG
CGCGCAGTGAGTCCGAGGATTTAAaTGTTGAGAGcCCTcCCGCCGCAATCGGGAGTTCTTCCGCGTCCcGCTC
CGAAGcCTTCAGAcCGCAGGTGGTTAaCGGT

>**RT138**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAAAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATAGTCGTATTGCTTCCTTCCTTAAGTCT
GGATATGATGTCGGTGAATTGTGCTCTAAAGGCTATATGAGCGTCCCTCAAGTGTTATGTGCTGCTACTCGAA
CGGTTTCTACCGATGCTGAGGGATCTTTGAGAATTTACCTGGCTGATCTAGGCGATAAGGAGTTATCTCCTAT
AGATGGACAATGCGTATCATTGCATAACCATGATCTTCCCGCTTTAGTGTCATTCCAACCGACGTATGACTGT
CCTATGGAGACTGTTGGGAATCGTAAGCGGTGTTTTGCTGTCGTTATCGAAAGGCATGGTTATATTGGGTACA
CCGGCACCACAGCAAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAACAACAACTACACTCATAT
CGCTGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTGTTGAAGTCGCAATTGAATAACATAGAATCTTCGCAATACGTTTTGACGGATTCGAAGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGATTTGAATGTTGAAAGCCCTCCCATCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAACCTTCAGACCGCAGGTGGTTAACGGT

>**RT144**

ATGGCTTTCCAAGGTACCAGCAGAACTTTAACCCAACAGTCATCAGCGGCTACGTCTGACGAACTTCAAAAGA
TATTGTTCAGCCCTGATGCCATTAAGAAAATGGCTGCTGAGTGTGACCTAGGCCGGCATCATTGGATGCGTGC
TGATAATGCTATCTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATTGCTTCCTTCTTTAAGTCT
GGATATGATGCTGGTGAATTGTGCTCTAAAGGATACATGAGCGTCCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTTTCTACTGATGCTGAAGGGTCTCTGAGAATTTACCTAGCTGATTTAGGCGATAAGGAGTTGTCTCCTAT
AGATAGGCAATGCGTCACATTACATAACCATGATCTTCCCGCTTTGGTGTCTTTCCAACCGACATACGACTGT
CCCATGGAGATAGTTGGGAATCGGAAGCGGTGCTTTGCTGTTGTCATCGAAAAACATGGTTATGTTGGGTATA
CTGGCACCACGGCTAGCGTGTGTAGTAACTGGCAAGCACGGTTTTCTTCTAAGAACAACAACTACACTCATAT
CGCAGCCGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTCGCTCGC
CTGTTGAAGTCGCAACTGAATAACATAGAATCTTCGCAATACGTCTTGACGGATTCGAAGATTAATCAGAATA
CGCGCGGTGAGTCCGAGGAATTAAATGTTGAAAGCCCTCCTGTCGCAATCGGGAGTTCTTCCGCGTCCCGCTT
CGAATCCTTCAGGCCGCAGGTGGTTAACGGT

>**TR15**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGAACTTCAAAAGA
TATTATTTAGCCCTGAAGCCATTAAGAAGATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGCGC
TGATAATGCTATTTCAGTCCGGCCCCTCGTTCCCGAAGTAACCCACGGTCGTATTGCTTCTTTCTTTAAGTCT
GGATATGATGCTGGTGAATTGAGCTCCAAAGGATACATGAGCGTCCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTCTCCACCGATGCTGAAGGATCTTTGAGAATTTACCTAGCTGATCTGGGCGACAAGGAGTTGTCTCCTAT
AGATAGACAATGCGTCACATTACATAACCATGATCTGCCCGCTTTAGTGTCTTTCCAGCCGACGTATGATTGT
CCTATGGAAATTGTTGGGAATCGTAAGCGGTGTTTTGCTGTTGTTATCGAAAGACATGGTTATATTGGGTATA
CCGGCACCACAGCTAGCGTGTGTAGTAATTGGCAAGCACGGTTTTCTTCTAAGAATAACAACTACACTCATAT
CGCAGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACTAAACCGTCAGCTGTTGCTCGC

CTGTTGAAGTCGCAATTGAACAACATAGAATCTTCGCAATACGTTTTAACGGATGCGAGGATTAATCAAAATG
CGCGCAGTGAGTCCGAGGAATTAAATGTTGAAAGCCCTCCCGCCGCAATCGGGAGTTCTTCCGCGTCCCGCTC
CGAAGCTTTCAGACCGCAGGTGGTTAACGGT

>**UB50**

ATGGCTTTCCAAGGTACCAGTAGGACTTTAACTCAACAGTCCTCAGCGGCTACGTCTGACGATCTTCAGAAAA
TCTTATTTAGCCCTGACGCCATCAAGAAAATGGCTACTGAGTGTGACCTAGGCCGGCATCATTGGATGCGTGC
TGATAATGCGATATCAGTCCGGCCCCTCGTTCCCGAAGTAACCCATGGTCGTATTGCTTCCTTCTTTAAGTCT
GGATATGATGTCGGCGAATTGTGCTCTAAAGGATATATGAGCGTTCCTCAAGTGTTGTGTGCTGTTACTCGAA
CAGTTTCTACCGATGCTGAGGGATCTTTGAGAATTTACCTGGCTGATCTAGGCGACAAGGAGTTGTCTCCTAT
AGATGGGCAATGCGTTTCATTACATAACTATGATCTCCCCGCTTTAGTGTCTTTCCAACCGACGTATGACTGT
CCTATGGAAAGTGTTGGGAACCGTAAGCGGTGTTTTGCAGTCGTCATCGAGAGACATGGTTATATTGGGTATA
CCGGCACCACAGCAAGCGTGCGTAGTAATTGGAATGCACGATTTTCCTCTAAGAATAACAACTACACTCATAT
CGCTGCTGGGAAGACTCTAGTACTGCCATTCAACAGATTAGCTGAGCAAACGAAACCGTCAGCCGTTGCTCGC
CTGTTGAaGTCGCAATTGAATAACATAGAATCTTCGCAATACGTTTTAACGGATTCGAAGATTAGTCAAAGTG
CGCGCAGTGAGTCCGAGGATTTAAaTGTTGaAAGCCCTCCCGTCGCAGTCGGGAGTtCTTCcGCGTCCcGCTC
cGAAaCCTTtAGACcGCAGgTGGTTAaCGGT

7.2.3 2b sequences

>**AN**

GATGAATTCTTGTTTGCTCACTTCATGAGCTTTGTTGATCGATTGAAGTTTTTGGACAGAATGACTCAGTCTT
GTATCGATCAACTTTCACTCTTTTTCGAGTTGAAATACAGGAAGTCAGGGGCCGAGGCTGCTTTAATGTTAGG
CGCCTTTAAGAAATATACCGCTAATTTCCAATCCTATAAAGAGCTCTACTATTCAGATCGTCGTCAGTGCGAA
TTGATCAATTCGTTTAGTTGTGTGGAGTTAAGGATTGAGCGTTCGATTTCTACTAAGCAGCGAAAGAAGAAAG
ATGGAATTGAACGAAGGCGCAGTGACAAACGTCGAACTCCAGCTGGCTCGTATGATGGAGGTGAAGAGACAGA
GACGAAGGTCTCACAAGAAGAATCGACGGGAACGAGGTCACAAAAGTCCCAGCGAGAGGGCGCGTTCAAATCT
CAGACTGTTCCGTTTTCTACCGTTTTATCAGATAGACGGTTCGGAGCTGATAGAGATGCACCACCGTGCGCGC
GCGGTGGAATTGTCCGAGTCTGAGGCCCCTTGTTTTCCATTATCAGCGGAAGAAGACCATGATTTTGACGATA
CGGATTGGTTCGCTGGTAATGAATGGGCGGAAGGTGTGTTTTGAATCTCCCCCTTCCTTTTCTCCCGCCATTT
CTGAGGCGGGGCTGAGTTGGCAGTATTGCTCTAAACTGTCTGAAGTCACTAAACGCTTTGCGGTGAACGGGTT
GTCCATCC

>**CN04**

GATGAATTCTTGTTTGCTCACTTCATGAGCTTTGTTGATCGATTGAAGTTTTTAGACCGAATGACTCAGTCTT
GCATCGATCAGCTTTCACTCTTCTTCGAGTTGAAATACAGGAAGTCCGGGGCGGAGGCTGCCTTAATGCTAGG
CGCCTTTAAGAAATATACCGCTAATTTCCAGTCCTACAAAGAACTCTATTATTCAGATCGTCGTCAGTGCGAA
TTGATCAATTCGTTTAGTTGTGTAGAGTTGAGGATTGAGCGTTCGAGTTCCATTAAGCAGCGAAAGAAGAAAA
ATGGAATTGAACGAAGGCGCGATGACAAACGTCGAACTCCAACTGGCCCGCATGGTGGAGGCGAAGAGACAGA
GACGAAGGTCTCACAAGAAGAATCGACGGGAACGATGTCACAAAAGTCCCAGCGAGAGGGCGCGTTCAAATCT
CAGACTATTCCGCTTCCTACCGTTCTATCAGGTAGATGGTTCGGAACTGATAGAGATGTACCGCCACGTGAAC
GTGGCGGAATTGTCCGGGTCTGAGGCCCCTTGTTTTACGTTGCCAGCGGAAGATGACCATGATTTCGACGATA
CAGATTGGTTCGCTGGTAACGAATGGGCGGAAGGAGCGTTCTGAACCCCTCCTTCCTTTCTCCCTCCAGTTTT
GAGGCGGGAGCTGAGTTGGCAGTATTGCTATAAACTGTCTGAAGTCACTAAACGTTTTATCGGTGAACGGGTT
GTCCATCC

>**KB8**

GATGAATTCCTGTTTGCTCACTTCATGAGCTTTGTCGATCGATTGAAGTTTCTAGACAGAATGACTCAGTCTT
GCATCGATCAACTTTCACTTTTTCGAGTTGAAATATAGGAAGTCTGGGGCTGAGGCTGCTTTAATATTAGGCG
CCTTTAAGAAATATACCGCTAATTTCCAATCCTATAAAGAACTCTACTATTCAGATCGTCGTCAGTGCGAATT
GATCAATTCGTTTAGTTGTGCAGAGTTGAGGATTGAGCGTTCGATTTCCACTAAACAGCGAAAGAAGAAAGAT
GGAATTGAACGAAGGCGCAGTGACAAACGTCGAACTCCAGCTGGCTCGTATGATGGAGGTGAGGAGACAAAGA
CGAAAGTCTCACAAGAAGAATCGACGGGAACGAGGTCACAAAAGTCCCAGCGAGAGAGCGCGTTCAAATCTCA
GACTATTCCGATTTTTACCGTTTTATCAGATAGATGGTTCGGAACTGATAGAGATGTACCGCCACGCGAATGT
GGTGGAATTGTCCGAGTCTGAGGCTCCTCAGTTTGCGTTACCAGCGGAAGAAGACCATGATTTTGACGACACA
GATTGGTTCGCTGGTAATGAATGGGCGGAAGGTGCGTTTTGAATCTCCCCTTCCTTTTCTCCCTCCAGTTTTC
TGAGGCGGGAGCTGAGTTGGCAGTGCTGCTACAAACTGTCTGAAGTCACTAAACGCTTTGCGGTGAAGCGGGT
TGTCCATCC

>**KS44**

GATGAATTCCTGTTTGCTCACTTCATGAGCTTAGTCGATCGATTGAAGTTTCTGGATAGAATGACTCAGTCCT
GCATCGATCAACTTTCACTCTTTTTCGAGTTGAAATATAGGAAGTCTGGGGCTGAGGCTGCTTTAATATTAGG
CGCCTTTAAGAAATATACCGCTAATTTCCAATCCTATAAAGAACTCTACTATTCAGATCGTCGTCAGTGCGAA
TTGATCAATTCGTTTAGTTGTGCAGAGTTGAGGATTGAGCGTTCGATTTCCACTAAACAGCGAAAGAAGAAAG
ATGGAATTGAACGAAGGCGCAATGACAAACGTCGAACTCCAGCTGGCTCGCATGATGGAGGTGAGGAGACAAA
GACGAAAGTCTCACAAGAAGAATCGACGGGAACGAGGTCACAAAAGTCCCAGCGAGAGAGCGCGTTCAAATCT
CAGGCTATTCCGATTTTTACCGTTTTATCAGATAGATGGTTCGGAACTGATAGAGATGTACCACCACGCGAGT

GTGGTGGAATTGTCCGAGTCTGAGGCTCCTCGGTTTACGTTACCAGCGGAAGAAGACCATGATTTTGACGACA
CAGATTGGTTCGCTGGTAATGAATGGGCGGAAGGTGCGTTTTGAATCTCCCCTTCCTTTTTCTCCCTCCAGTT
TTCTGAGGCGGGAGCTGAGTTGGCAGTACTGCTACAAACTGTCTGAAGTCACTAAACGCTATGCGGTGAACGG
GTTGTCCATCC

>**P3613**

GATGAATTCCTGTTTGCTCACTTCATGAGCTTTGTTGATCGATTGAAGTTTTTGGATCGAATGACTCAGTCTT
GTATCGATCAGCTTTCACTCTTCTTCGAGTTGAAATACAGGAAGTCTGGGGCGGAGGCTGCCTTAATGTTAGG
CGCCTTTAAGAAATATACCGCTAATTTTCAGTCCTACAAAGAACTCTATTATTCAGATCGTCGTCAGTGCGAA
TTGATCAATTCGTTTAGTTGCGTAGAGTTGAGGATTGAGCGTTCGAGTTTCATTAAGCAGCGAAAGAAGAAGA
AAGATGGAATTGAACGCAGGCGCAATGACAAGCGTCGAACTCCAACTAGCCCGCATGGTGGAGGCGAAGAGAC
AGAGACGAAGATCTCACAAGAAGAATCGACGGGAACGATGTCACAAAAGTCCCAGCGAGAGGGCGCGTTCAAA
TCTCAGACTGTTCCGCTTCCTACCGTTCTTTCAAGTAGATGGTTTGGAACTGATAGAGATGTACCGCCACGCG
AGCGTGGCGGAATTGTCCGAGTCTGAGGCCCCTTGTTTTCCGTTGCCAGCGGAAGATGACCATGATTTCGACG
ATACAGATTGGTTCGCTGGTAACGAGTGGGCGGAAGGAGCATTCTGAACCCCTCCTTCCTTTCTCCCCCAGTT
TTCTGAGGCGGGAGCTGAGTTGGCAGTATTGCTATAAACTGTCTGAAGTCACTAAACGTTTTACGGTGAACGG
GTTGTCCATCC

>**P522**

GATGAATTCTTGTTTGCTCACTTCATGAGCTTTGTTGATCGATTGAAGTTTTTGGACCGAATGACTCAGTCTT
GTATCGATCAGCTTTCACTCTTTTTTGAGTTGAAATACAGGAAGTCTGGGGCTGAGGCTGCTTTAATGTTAGG
CGCCTTTAAGAAGTATACCGCTAATTTTCAATCCTATAAGGAACTCTACTATTCAGATCGTCGTCAGTGCGGA
TTGATCAATTCGTTTAGTTGCGTAGAGTTGAGGATTGAGCGTTCGAGTTCCACTAAGCAGCGAAAGAAGGAAG
ATGGAATTGAACGAAGGCGCAGTGACAAACGTCGAACTCCAACTGGCTCGTATGGTGGAGGTGAAAAGACAGA
GACGAAGGTCTCACATGAAGAATCGACGGGAACGTGGTCACAAAAGTCCCAGCGAGAGGGCGCGTTCAAATCT
CAGATTGTTCCGCTTCCTACCATTTTATCAGGTAGATGGTTCGGAACTGATAGAGCTCCACCATGTGAACATG
GTGGAATTATCCGAATCTGAGGCCCCTCGTTTATCGTTACTGGCGGAAGAAGACCATGATTTTGACGATACGG
ATTGGTTCGCTGGTAACGAGTGGGCGGAAGGGTCGTTTTGAACCTCCCCTTTCCTTTCTCCCTCCGGTTTTCT
GAGGCGGGAGCTGAGTTGGCAGTATTGCTATAAACTGTCTGAAGTCACTAAACGTTTTTCGGTGAACGGGTTG
TCCATCC

>**PV0418**

GATGAATTCTTGTTTGCTCACTTTATGAGTTTTGTTGACAGGCTGAAGTTTATGGGCAGGATGTCGCAATCCT
GTATTGATCAATTGTCCATTTTCTTCGAGTTGAAATACAAGAAGTCCGGGAACGAAGCCGCCCAAGTTTTGGG
CGCGTTTAAGAAGTACACCGCTAATTTTAACGCCTATAAGGAGCTTTATTACTCTGACCGTCAACAATGTGAC
TTGGTCAATACGTTTTGTGTATCTGAGTTCCAGGTGATTCGTCGCACTATCGTTAAGAAGAAGAAGAATGGAT
GTGTTGACAGTAGTGGTGTCGACCGCCGACCTCCACCTAGCCCATTTGCAGGAGGTGAAACGTCGAAGACGAA
GGTCTCACGTCAGAAACCGGCGAGCGAGGGGTTACAAAAGTCCCAGCGAGAGAGCGCGATCTATAGCGAGACT
TTTCCAGATGTTACCATTCCACGGAGTAGATCCCGCGGATTGGTTTCCTGATGTCGTTCGCTCTCCGTCCGTT
ACCAGCCTTGTTTCTTATGAATCTTTTGATGATACTGATTGGTTTGCTGGTAACGAATGGGCCGAAGGGTCGT
TTTGATTTCCGACCCTTCGTCGTCCGAAGACGTTAAACTACACTCTCTTTATTGCGAGTGCTGAGTTGGTAGT
TTGCTCTAAACTATCTGAAGTCGCTAAATCCATTACTGGTTGCGAACGGGTTGTCCATCC

>**PV0420**

GATGAATTCCTGTTTGCTCACTTTATGAGTTTTGTTGACAGGCTGAAGTTTATGGGCAGGATGTCGCAATCCT
GTATTGATCAATTGTCCATTTTCTTCGAGTTGAAATACAAGAAGTCCGGGAACGAAGCCGCCCAAGTTTTGGG
CGCGTTTAAGAAGTACACCGCTAATTTTAACGCCTATAAGGAGCTTTATTACTCTGACCGTCAACAATGTGAC
TTGGTCAATACGTTTTGTGTATCTGAGTTCCAGGTGATTCGTCTCACTATCGTTAAGAAGAAGAAGAATGGAT
GTGTTGACAGTAGTGGTGTCGACCGCCGACCTCCACCTAGCCCATTTGCAGGAGGTGAAACGTCGAAGACGAA
GGTCTCACGTCAGAAACCGGCGAGCGAGGGGTTACAAAAGTCCCAGCGAGAGAGCGCGATCTATAGCGAGACT
TTTCCAGATGTTACCATTCCACGGAGTAGATCCCGCGGATTGGTTTCCTGATGTCGTTCGCTCTCCGTCCGTT
ACCAGCCTTGTTTCTTATGAATCTTTTGATGATACTGATTGGTTTGCTGGTAACGAATGGGCCGAAGGGTCGT
TTTGATTTCCGACCCTTCGTCGTCCGAAGACGTTAAACTACACTCTCTTTATTGCGAGTGCTGAGTTGGTAGT
TTGCTCTAAACTATCTGAAGTCGCTAAATCCATTACTGGTTGCGAACGGGTTGTCCATCC

7.3 Genomic markers among six isolates

Table 1: The genomic markers among six CMV isolates

	P3613			P522			KS44			AN			PV0420		
	[1]RNA1	RNA2	RNA3	RNA1	RNA2	RNA3	RNA1	RNA2	RNA3	RNA1	RNA2	RNA3	RNA1	RNA2	RNA3
P522	MboI	RsaI	AclI	/			/			/			/		
KS44	EcoRV	MboI	AclI	MboI	RsaI	AsuII									
AN	MboI	RsaI	AclI	MboI	PvuII	AsuII	AluI	RsaI	HindII						
PV0420	MboI	RsaI	AclI	MboI	RsaI	EcoRV	MboI	MboI	EcoRV	MboI	RsaI	EcoRI			
RT68	/	/	/	/	/	/	/	/	/	/	/	/	AluI	[2]/	VspI

[1]RNA 1, RNA 2 and RNA 3 stand for three different genome segments of CMV. The corresponding genomic marker in different genome segment of different virus combinations were indicated in the table. For instance, the genomic marker of RNA 2 between isolates P3613 and AN was Rsa I, which resulted in 520 bp and 230 bp fragments with P3613-RNA2 PCR-product and 736 bp fragment with AN-RNA2 PCR-product (Table 2).

[2] RNA 2 segment between PV0420 and RT68 was differentiated with RT-PCR at the annealing temperature of 61℃. This could be amplified a 736 bp fragment only with isolate PV0420.

Table 2: RFLP analysis of six CMV isolates*

Enzyme / isolate	RNA1 (bp)			RNA2 (bp)			CP (bp)						MP (bp)
	MboI	TasI	AluI	MboI	RsaI	HinfI	AclI	HindIII	HindII	EcoRV	EcoRI	VspI	AsuII
P3613	1124	/	/	400+150	520+230	/	187+585	/	/		/	/	/
P522	580+540	/	/	/	736	310+150	772	/	/	772	/	/	834
KS44	1000+80	267	540	520+150	500+180	/	772	458+300	732	/	/	/	712+122
AN	800+300	/	800+140	/	736	300+160	772		580+170	/	772	/	712+122
PV0420	350+250	440+230	/	350+250	560+270	/	772	460+300	/	470+267	590+170	788	/
RT68	/	/		/	/	/	/	/	/	/	/	332+456	/

*The RFLP analysis patterns for genomic markers of RNA1, RNA2 and RNA3 among 6 isolates. Those genomic markers were determined either from randomly selected restriction enzyme or from sequences analysis if the sequences of corresponding isolates were available. All of selected markers were reconfirmed by PCR-RFLP (2.2.14) with selected enzymes. The primers for PCR reaction are as follows: primers 5′-RNA1 and 3′-RNA1 are for RNA 1 amplification, primers 5′-RNA2 and 3′-RNA2 for RNA2 amplification, primers 5′-MP and 3′-MP for the MP of RNA 3 amplification and primers 5′-CP and 3′-CP for the CP of RNA 3 amplification. The band less than 100 bp is not shown on the table.

7.4 Statistical analysis of reassortants among six isolates

7.4.1 Reassortants of AN and P3613

Table 1: The six possible genome combinations between the isolates P3613 and AN and their frequency

Genotypes[1]	Name	observed frequency[2]	expected frequency	X^2	siginificance
$1_p2_p3_p$	P3613	4	4.75	0.11842	
$1_p2_A3_A$	1R1	2	4.75	1.59221	
$1_A2_p3_A$	1R2	12	4.75	11.06579	
$1_p2_p3_A$	1R3	14	4.75	18.01316	
$1_p2_A3_p$		0	4.75	4.75000	
$1_A2_A3_p$		0	4.75	4.75000	
$1_A2_p3_p$		0	4.75	4.75000	
$1_A2_A3_A$	AN	6	4.75	0.32895	
		38	38	45.36842	*P<0.000000

[1]Individual segments of CMV strains P3613 and AN were designated as 1, 2 or 3. For instance, $1_p2_A3_A$ contains RNA1 of P3613 and RNA2 plus RNA3 of AN. [2]Some reassortants contained an asymmetric complement of genome segments that appeared as one or more segments contributed by both parents with the remaining segments contributed by only one parent (e.g. $1_A2_A3_{A+p}$) were not shown on the table.

From fifty local lesions thirty-eight progeny isolates were detected to be reassortants, ten were parental isolates and twelve contained both parental strains in at least one genome segment. In RNA 1 both isolates, P3613 and AN, contributed almost equally, while in RNA 2 P3613 contributed to 77 % and in RNA 3 to 11 % to the progeny isolate. The frequency of detection of the parental and all possible reassortants genotypes revealed to be non random by X^2 analysis.

7.4.2 Reassortants of P3613 and KS44

From twenty-three local lesions thirteen progeny isolates were detected to be reassortants, eight were parental isolates and two contained both parental strains with at least one genome segment. In RNA 1 KS44 isolate contributed to 62 %, while in RNA 2 to 48 % and in RNA 3 to 96 % to the progeny isolate, respectively. The frequency of

detection of the parental and all possible reassortants genotypes revealed to be non random by X^2 analysis.

Table 2: The six possible genome combinations between the isolates P3613 and KS44 and their frequency

Genotypes[1]	Name	observed frequency[2]	expected frequency	X^2	siginificance
$1_p2_p3_p$	P3613	1	2.625	1.00595	
$1_p2_k3_k$	2R1	3	2.625	0.05357	
$1_p2_p3_k$	2R2	4	2.625	0.72024	
$1_k2_p3_k$	2R3	6	2.625	4.33929	
$1_p2_k3_p$		0	2.625	2.62500	
$1_k2_k3_p$		0	2.625	2.62500	
$1_k2_p3_p$		0	2.625	2.62500	
$1_k2_k3_k$	KS44	7	2.625	7.29167	
		21	21	21.28572	*P<0.003377

[1]Individual RNAs of CMV strains P3613 and KS44 were designated as 1, 2 or 3. For instance, $1_p2_k3_k$ contains RNA1 of P3613-CMV and RNA2 plus RNA3 of KS44-CMV. [2]some reassortments contained an asymmetric complement of genome segments that appeared as one or more segments contributed by both parents with the remaining segments contributed by only one parent (e.g. $1_p2_p3_{p+k}$) was not shown on the table.

7.4.3 Reassortants of P522 and KS44

Table 3: The six possible genome combinations between the isolates P522 and KS44 and their frequency

Genotypes[1]	Name	observed frequency[2]	expected frequency	X^2	siginificance
$1_p2_p3_p$	P522	0	5.5	5.5000	
$1_p2_k3_k$	3R1	14	5.5	13.1364	
$1_p2_p3_k$	3R2	1	5.5	3.6818	
$1_k2_p3_k$		0	5.5	5.5000	
$1_p2_k3_p$		0	5.5	5.5000	
$1_k2_k3_p$		0	5.5	5.5000	
$1_k2_p3_p$		0	5.5	5.5000	
$1_k2_k3_k$	KS44	29	5.5	100.4091	
		44	44	144.7273	*P<0.000000

[1]Individual RNAs of CMV strains P522 and KS44 were designated as 1, 2 or 3. For instance, $1_p2_k3_k$ contains RNA1 of P522-CMV and RNA2 plus RNA3 of KS44-CMV. [2]some reassortments contained an asymmetric complement of genome segments that appeared as one or more segments contributed by both parents with the remaining segments contributed by only one parent (e.g. $1_p2_p3_{p+k}$) was not shown on the table.

From forty-eight local lesions fifteen progeny isolates were detected to be reassortants between isolates P522 and KS44, twenty-nine were parental isolates and four contained both parental strains with at least one genome segment. In RNA 2 and RNA 3 isolate KS44 contributed almost all, while in RNA 1 contributed to 66 % to the progeny isolate. The frequency of detection of the parental and all possible reassortants genotypes revealed to be non random by X^2 analysis.

7.4.4 Reassortants of AN and KS44

Table 4: The six possible genome combinations between the isolates AN and KS44 and their frequency

Genotypes[1]	Name	observed frequency[2]	expected frequency	X^2	siginificance
$1_A2_A3_A$	AN	14	4.625	19.00338	
$1_A2_k3_k$	4R1	0	4.625	4.62500	
$1_A2_A3_k$	4R2	0	4.625	4.62500	
$1_k2_A3_k$	4R3	2	4.625	1.48986	
$1_A2_k3_A$		8	4.625	2.46284	
$1_k2_k3_A$		4	4.625	0.08446	
$1_k2_A3_A$		0	4.625	4.62500	
$1_k2_k3_k$	KS44	9	4.625	4.13581	
		37	37	41.05405	*P<0.000001

[1]Individual RNAs of CMV strains AN and KS44 were designated as 1, 2 or 3. For instance, $1_A2_k3_k$ contains RNA1 of AN-CMV and RNA2 plus RNA3 of KS44-CMV. [2]some reassortments contained an asymmetric complement of genome segments that appeared as one or more segments contributed by both parents with the remaining segments contributed by only one parent (e.g. $1_A2_A3_{A+k}$) was not shown on the table.

From forty-five local lesions derived from simultaneous inoculation of the two resistant breaking isolates KS44 and AN fourteen progeny isolates were detected to be reassortants, twenty-three were parental isolates and eight contained both parental strains with at least one genome segment. In RNA 1 AN contributed to 59 %, in RNA 2 to 43 % and in RNA 3 to 81 % to the progeny isolate. The frequency of detection of the parental and all possible reassortants genotypes revealed to be non random by X^2 analysis.

7.4.5 Reassortants of P522 and P3613

Table 5: The six possible genome combinations between the isolates P522 and P3613 and their frequency

Genotypes[1]	Name	observed frequency[2]	expected frequency	X^2	siginificance
$1_{p5}2_{p5}3_{p5}$	P522	3	3.0	0.00000	
$1_{p5}2_{p3}3_{p3}$	5R1	1	3.0	1.33333	
$1_{p5}2_{p5}3_{p3}$	5R2	0	3.0	3.00000	
$1_{p3}2_{p5}3_{p3}$	5R3	0	3.0	3.00000	
$1_{p5}2_{p3}3_{p5}$		5	3.0	1.33333	
$1_{p3}2_{p3}3_{p5}$		7	3.0	5.33333	
$1_{p3}2_{p5}3_{p5}$		0	3.0	3.00000	
$1_{p3}2_{p3}3_{p3}$	P3613	8	3.0	8.33333	
		24	24	25.33333	*P<0.000664

[1]Individual RNAs of CMV strains P522 and P3613 were designated as 1, 2 or 3. For instance, $1_{p3}2_{p5}3_{p5}$ contains RNA1 of P3613-CMV and RNA2 plus RNA3 of P522-CMV. [2]some reassortments contained an asymmetric complement of genome segments that appeared as one or more segments contributed by both parents with the remaining segments contributed by only one parent (e.g. $1_{p3}2_{p3}3_{p3+p5}$) was not shown on the table.

From thirty-two local lesions derived from simultaneous inoculation of the two non-resistant breaking isolates P522 and P3613 thirteen progeny isolates were detected to be reassortants, eleven were parental isolates and eight contained both parental strains with at least one genome segment. In RNA 1 P522 isolate contributed to 38 %, while in RNA 2 to 13 % and in RNA 3 to 63 % to the progeny isolate. The frequency of detection of the parental and all possible reassortants genotypes revealed to be non random by X^2 analysis.

7.4.6 Reassortants of P522 and AN

From twenty-five local lesions derived from simultaneous inoculation of the resistant breaking isolates AN and the non resistant breaking isolate P522 seven progeny isolates were detected to be reassortants, seventeen were parental isolates and one contained both parental strains with at least one genome segment. In RNA 2 and RNA3 AN isolate contributed to 96 %, while in RNA 1 to 79 % to the progeny isolate. The frequency of detection of the parental and all possible reassortants genotypes revealed to be non random by X^2 analysis.

Table 6: The six possible genome combinations between the isolates P522 and AN and their frequency

Genotypes[1]	Name	observed frequency[2]	expected frequency	X^2	siginificance
$1_p2_p3_p$	P522	0	3.0	3.00000	
$1_p2_A3_A$	6R1	4	3.0	0.33333	
$1_p2_p3_A$		0	3.0	3.00000	
$1_A2_p3_A$	6R2	1	3.0	1.33333	
$1_p2_A3_p$	6R3	1	3.0	1.33333	
$1_A2_A3_p$	6R4	1	3.0	1.33333	
$1_A2_p3_p$		0	3.0	3.00000	
$1_A2_A3_A$	AN	17	3.0	65.33334	
		24	24	78.66666	*P<0.000000

[1]Individual RNAs of CMV strains P522 and AN were designated as 1, 2 or 3. For instance, $1_p2_A3_A$ contains RNA1 of P522-CMV and RNA2 plus RNA3 of AN-CMV. [2]some reassortments contained an asymmetric complement of genome segments that appeared as one or more segments contributed by both parents with the remaining segments contributed by only one parent (e.g. $1_p2_p3_{p+A}$) was not shown on the table.

7.4.7 Reassortants of RT68 and PV0420

Table 7: The six possible genome combinations between the isolates RT68 and PV0420 and their frequency

Genotypes[1]	Name	observed frequency[2]	expected frequency	X^2	siginificance
$1_{pv}2_{pv}3_{pv}$	PV0420	3	3.375	0.0417	
$1_{pv}2_R3_R$	7R1	3	3.375	0.0417	
$1_{pv}2_{pv}3_R$		0	3.375	3.3750	
$1_R2_{pv}3_R$		0	3.375	3.3750	
$1_{pv}2_R3_{pv}$		0	3.375	3.3750	
$1_R2_R3_{pv}$		0	3.375	3.3750	
$1_R2_{pv}3_{pv}$		0	3.375	3.3750	
$1_R2_R3_R$	RT68	21	3.375	92.0417	
		27	27	109.0000	*P<0.000000

[1]Individual RNAs of CMV strains RT68 and PV0420 were designated as 1, 2 or 3. For instance, $1_{pv}2_R3_R$ contains RNA1 of PV0420-CMV and RNA2 plus RNA3 of RT68-CMV. [2]some reassortments contained an asymmetric complement of genome segments that appeared as one or more segments contributed by both parents with the remaining segments contributed by only one parent (e.g. $1_R2_R3_{R+pv}$) was not shown on the table.

From thirty-five local lesions derived from simultaneous inoculation three progeny isolates were detected to be reassortants, twenty-seven were parental isolates and six contained both parental strains with at least one genome segment. In RNA 1 RT68

isolate contributed to 78 %, while in RNA 2 and RNA 3 to 89 % to the progeny isolate. The frequency of detection of the parental and all possible reassortants genotypes revealed to be non random by X^2 analysis.

7.4.8 Reassortants of AN and PV0420

From thirty local lesions two progeny isolates were detected to be reassortants, eleven were parental isolates and seventeen contained both parental strains with at least one genome segment. In RNA 1 and RNA 2 AN isolate contributed to 77 %, while in RNA 3 to 62 % to the progeny isolate. The frequency of detection of the parental and all possible reassortants genotypes revealed to be non random by X^2 analysis.

Table 8: The six possible genome combinations between the isolates AN and PV0420 and their frequency

Genotypes[1]	Name	observed frequency[2]	expected frequency	X^2	siginificance
$1_{pv}2_{pv}3_{pv}$	PV0420	6	3.25	2.32692	
$1_{pv}2_A3_A$	8R1	0	3.25	3.25000	
$1_{pv}2_{pv}3_A$		0	3.25	3.25000	
$1_A2_{pv}3_A$		0	3.25	3.25000	
$1_{pv}2_A3_{pv}$		0	3.25	3.25000	
$1_A2_A3_{pv}$		4	3.25	0.17308	
$1_A2_{pv}3_{pv}$		0	3.25	3.25000	
$1_A2_A3_A$	AN	16	3.25	50.01923	
		26	26	68.76923	*P<0.000000

[1]Individual RNAs of CMV strains AN and PV0420 were designated as 1, 2 or 3. For instance, $1_{pv}2_A3_A$ contains RNA1 of PV0420-CMV and RNA2 plus RNA3 of AN-CMV. [2]some reassortments contained an asymmetric complement of genome segments that appeared as one or more segments contributed by both parents with the remaining segments contributed by only one parent (e.g. $1_A2_A3_{A+pv}$) was not shown on the table.