

Open Access

## Molecular and Biological Characterization of an Isolate of *Cucumber mosaic virus* from *Glycine soja* by Generating its Infectious Full-genome cDNA Clones

Mi Sa Vo Phan<sup>1</sup>, Jang-Kyun Seo<sup>2</sup>, Hong-Soo Choi<sup>2</sup>, Su-Heon Lee<sup>3</sup> and Kook-Hyung Kim<sup>1,4\*</sup>

<sup>1</sup>Department of Agricultural Biotechnology and Plant Genomics and Breeding Institute, Seoul National University, Seoul 151-921, Korea

<sup>2</sup>Crop Protection Division, National Academy of Agricultural Science, Rural Development Administration, Suwon 441-707, Korea

<sup>3</sup>Department of Applied Biology, Kyungpook National University, Daegu 702-701, Korea

<sup>4</sup>Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Korea

(Received on February 11, 2014; Revised on March 17, 2014; Accepted on March 20, 2014)

**Molecular and biological characteristics of an isolate of *Cucumber mosaic virus* (CMV) from *Glycine soja* (wild soybean), named as CMV-209, was examined in this study. Comparison of nucleotide sequences and phylogenetic analyses of CMV-209 with the other CMV strains revealed that CMV-209 belonged to CMV subgroup I. However, CMV-209 showed some genetic distance from the CMV strains assigned to subgroup IA or subgroup IB. Infectious full-genome cDNA clones of CMV-209 were generated under the control of the *Cauliflower mosaic virus* 35S promoter. Infectivity of the CMV-209 clones was evaluated in *Nicotiana benthamiana* and various legume species. Our assays revealed that CMV-209 could systemically infect *Glycine soja* (wild soybean) and *Pisum sativum* (pea) as well as *N. benthamiana*, but not the other legume species.**

**Keywords :** *Cucumber mosaic virus*, sequence analysis, phylogenetic analyses, infectious clones, wild soybean

*Cucumber mosaic virus* (CMV) is a type member of the genus *Cucumovirus* and has a very broad host range (Palukaitis and García-Arenal, 2003). Its genome comprises three single-stranded plus-sense RNAs, named RNA 1–3 in order of decreasing size, which encode five viral proteins: 1a, 2a, 2b, 3a (movement protein; MP), and capsid protein (CP). Initially, CMV strains were classified into two main subgroups, I and II, based on converging biological, serological, physical criteria (Hu et al., 1995; Ilardi et al.,

1995; Wahyuni et al., 1992), nucleic acid hybridization (Owen and Palukaitis, 1988), and gene sequences (Owen et al., 1990; Sala and Bala, 1999). CMV strains were further divided into three subgroups including IA, IB, and II based upon phylogenetic analyses of a large number of the sequences of the CP gene as well as rearrangements in the 5' non-translated region (NTR) of RNA3 (Roossinck et al., 1999). However, additional analyses with full-genome sequences of CMV strains demonstrated that some strains were grouped in neither subgroup I nor subgroup II but in a new subgroup, named subgroup III (Jacquemond, 2012; Liu et al., 2009); while some strains were in between subgroups I and II (Chen et al., 2007; Jacquemond, 2012; Maoka et al., 2010), and the others were in an intermediate position between subgroups IA and IB (Jacquemond, 2012; Lin et al., 2004; Masuta et al., 2002). Thus, analysis of at least one part of each genomic segment was necessary for confident genotyping of CMV strains (Jacquemond, 2012).

Ever since the first report of CMV strain isolated from soybean (Koshimizu and Izuka, 1958), several additional CMV strains have been isolated (Hong et al., 2003; Senda et al., 2004; Takahashi et al., 1980). The firstly isolated CMV soybean strain induced mild mosaic and stunting on infected soybean plants and was transmissible through seeds with a high efficiency by aphids (Koshimizu and Izuka, 1963; Takahashi et al., 1980). Even though the biological characteristics of CMV soybean strain have been studied extensively, there was no available sequence information on CMV soybean strain genomes. A previous report analyzing the phylogenetic relationship of the nucleotide sequences of the 3a and CP genes of seven CMV soybean strains showed that these CMV soybean strains formed a distinct cluster separated from the other CMV strains (Hong et al., 2003).

\*Corresponding author.

Phone) +82-2-880-4677, FAX) +82-2-873-2317

E-mail) kookkim@snu.ac.kr

Construction of infectious cDNA clones corresponding to the genomes of RNA viruses is necessary since viral cDNA clones are essential sources for maintaining their original RNA genomes and for engineering viral vectors (Boyer and Haenni, 1994; Nagyová and Subr, 2007). In addition, easy manipulation of virus cDNA clones allows mutagenesis (deletions, insertions, substitutions) of specific regions of virus genomes and subsequent introduction of the mutated or wild type genomes into host plants for studies of virus life cycle, functions of viral proteins and their interactions with various host factors involved in virus replication, cell to cell movement, resistance responses (Bendahmane et al., 2002; Kim and Palukaitis, 1997; Paalme et al., 2004; Prüfer et al., 1995).

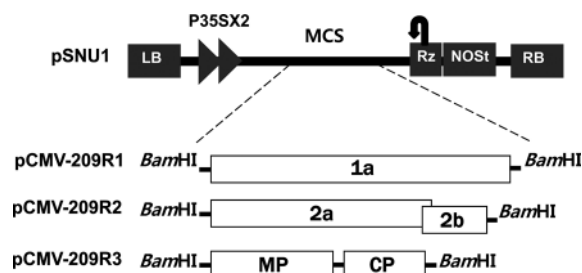
In this study, we obtained a CMV isolate from *Glycine soja* (wild soybean) and named CMV-209. Using this isolate, we generated infectious full-genome cDNA clones of CMV-209. Using the cloned CMV-209, its molecular characteristics were examined by determining complete nucleotide sequences of each genomic RNA segment and comparing to those of the other representative CMV strains. The biological characteristics were investigated by inoculating the cloned CMV-209 in various legume species as well as *Nicotiana benthamiana*.

## Materials and Methods

**Virus source and construction of the full-length cDNA clones of CMV-209.** Wild soybean samples showing typical mosaic symptoms were collected in Suwon, Korea 2006. To find the samples infected with CMV, the collected samples were screened by RT-PCR using the CMV specific primers CMV R3 561 Fw and CMV R3 1473 Rv (Table 1). One (sample No. 209) of the wild soybean samples was found to be infected with CMV. Total RNA was extracted from the CMV-infected wild soybean (No. 209) using TRI

**Table 1.** Primers used for CMV-detection and construction of the infectious cDNA clones of CMV-209

Name	Sequence (5' → 3')
CMV R3 561 Fw	TTGGGAATCGTAAGCGGTGTTTTG
CMV R3 1473 Rv	TTACAACGTTCACTCCCCACAAAG
CMV 3 end <i>Bam</i> HI	HIGCGGATCCTGGTCTCCTTTTRGAGRCC
CMV R1R2 5 end <i>Bam</i> HI	GCGGATCCTTTATTTACAAGAGCG-TACGG
CMV R3 5 end <i>Bam</i> HI	GCGGATCCGTAATCTTACCACTGTGT-GTGTG
CMV detR2 F	CCCTGTGGTGATCCGAGTAA
CMV detR2 R	CCGTAAGCTGGATGGACAA



**Fig. 1.** Schematic representation of the T-DNA vector, pSNU1, and CMV-209 cDNA constructs. The pSNU1 vector contains, in sequential order, a left border of T-DNA (LB), a double 35S promoter (P35Sx2), multiple cloning site (MCS), a *cis*-cleaving ribozyme sequence (Rz), a NOS terminator (NOST), and a right border of T-DNA (RB). The constructs pCMV-209R1, pCMV-209R2 and pCMV-209R3 express RNA1, RNA2 and RNA2 of CMV strain 209, respectively. The restriction enzyme cleavage sites used to make the constructs are shown.

Reagent method (MRC, USA) and subjected to RT-PCR to amplify full-length cDNAs corresponding to CMV RNA1, RNA2, and RNA3 using the specific primer pairs that contain *Bam*HI restriction site (Table 1). The amplicons were digested with *Bam*HI and ligated with the modified binary vector pSNU1 (Park and Kim, 2006) which was digested with *Bam*HI. The resulting clones carrying the full-length cDNA of RNA1, RNA2 or RNA3 of CMV were named as pCMV-209R1, pCMV-209R2 and pCMV-209R3, respectively (Fig. 1).

**Infectivity test and screening host range for CMV-209.** All of the CMV constructs (i.e. pCMV-209R1, pCMV-209R2, and pCMV-209R3) were individually transformed into *Agrobacterium tumefaciens* strain GV2260 by electroporation and cultured as previously described (Park and Kim, 2006). The *Agrobacterium* transformants were grown at 30°C until  $OD_{600} = 0.5 \pm 0.05$ . A mixture of three *Agrobacterium* cultures harboring constructs pCMV-209R1, pCMV-209R2, and pCMV-209R3 in a 1:1:1 ratio to yield a total  $OD_{600}$  of  $0.17 \pm 0.02$  was used for co-agroinfiltration. This final *Agrobacterium* mixture was infiltrated into the abaxial surface of the *N. benthamiana* leaves (3 to 4 weeks old) by using a 1-ml syringe without a needle. pCMV-Fny clones (pCR1(+), pCR2(+), and pCR3(+); Seo et al., 2009) were also included in agroinfiltration as control. After 7 days, the upper noninoculated *N. benthamiana* leaves were subjected to RT-PCR using the primer pairs CMV detR2 F and CMV detR2 R (Table 1) to examine whether the infiltrated plants are systemically infected with CMV-209. As shown in Table 2, the sap extracts from the *N. benthamiana* plants

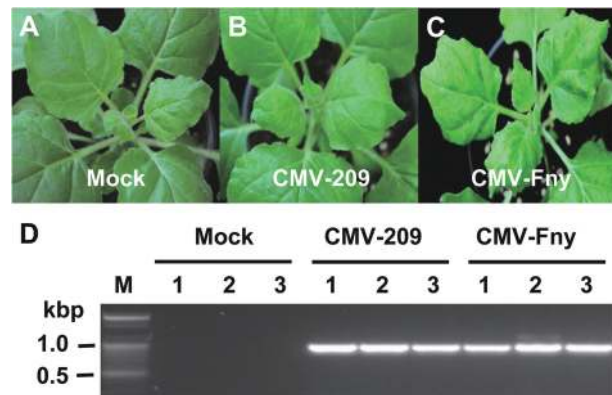
systemically infected with CMV-209 were used as viral inoculums for sap-inoculation of 14 legume cultivars including *Pisum sativum* (pea), *Vigna sinensis* (cowpea), *V. angularis* (cv. Chungbupat), *Phaseolus vulgaris* (Kidney bean), *Ph. vulgaris* (Pinto bean), *Glycine max* (cv. Somyong), *G. max* (cv. Pureunkong), *G. max* (cv. Sowon), *G. max* (cv. Danbaek), *G. max* (cv. L29), *G. max* (cv. Daewon), *G. max* (cv. V94-5152), *G. max* (cv. Huanggumkong), *G. max* (cv. Lee74), and *G. soja* (wild soybean).

### Sequencing, comparative, and phylogenetic analyses.

The full-length cDNA sequences in the infectious clones of CMV-209 (pCMV-209R1, pCMV-209R2 and pCMV-209R3) were sequenced using the appropriate primers (the list of primers used for the sequencing is available on request). The nucleotide sequences and deduced amino acids sequences of CMV-209 were compared with those of other representative CMV strains retrieved from the GenBank database. Sequences of PSV-P, PSV-W and PSV-Mi as well as TAV-KC were used as out-groups (Table 3). Multiple sequence alignments of all the analyzed sequences were obtained using CLUSTALS (Thompson et al., 1997). The sequence similarity comparisons were performed using Lasergen Package. The phylogenetic analyses were carried out using the Neighbor-Joining method (Saitou and Nei, 1987). The confidence probability was estimated using the bootstrap test (1000 repetitions) (Dopazo, 1994; Felsenstein, 1985; Rzhetsky and Nei, 1992). The evolutionary distances were computed using the p-distance method (Nei and Kumar, 2000) and were the units of the number of nucleotide or amino acid differences per site. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

## Results and Discussion

**Infectivity and host range of CMV-209.** To verify the infectivity of the full-genome cDNA clone of CMV-209, we first infiltrated *N. benthamiana* plants with a mixture of agrobacterium cultures harboring pCMV-209R1, pCMV-209R2, and pCMV-209R3. All the infiltrated plants showed very mild or no visible symptom (Fig. 2B) while showing typical symptoms including mosaic, stunting, or distortion of the leaves from CMV-Fny infiltrated plants (Fig. 2C). Although there was no obvious visible symptom from CMV-209 infiltrated plants, progeny viruses were detected from the upper systemic leaves by RT-PCR using CMV-specific primers (Fig. 2D) demonstrating that the full-genome cDNA clones of CMV-209 are fully infec-



**Fig. 2.** Infectivity of pCMV-209. Mock (A) represents healthy plant. pCMV-209 induced latent symptom in *Nicotiana benthamiana* (B) while pCMV-Fny caused typical symptoms including mosaic, stunting, or distortion of leaves (C). Photographs were taken at 13 days post-inoculation. (D) Detection of progeny viruses in *N. benthamiana* plants inoculated with CMV strains. Total RNAs, isolated from upper uninoculated leaves of each *N. benthamiana* plant harvested at 7 days post-agroinfiltration, were analyzed by RT-PCR. M indicates a 100 bp DNA ladder (Bioneer). Lanes 1 to 3 represent results from three RNA samples extracted from three *N. benthamiana* plants.

tious.

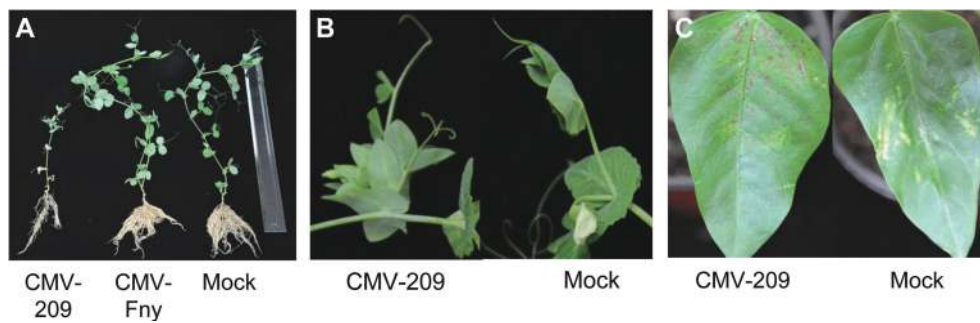
Next, since CMV-209 was isolated from wild soybean, the host range and infectivity of CMV-209 were examined in various legume species. To this end, twelve legume

**Table 2.** The infectivity of CMV-209 in various legume species<sup>a</sup>

Species	Symptoms <sup>b</sup> on the leaves	
	Inoculated	Upper
<i>Pisum sativum</i> (pea)	+/ns	+/st
<i>Vigna sinensis</i> (cowpea)	+/nll	–
<i>V. angularis</i> (cv. Chungbupat)	–	–
<i>Phaseolus vulgaris</i> (Kidney bean)	–	–
<i>Ph. Vulgaris</i> (Pinto bean)	–	–
<i>Glycine max</i> (cv. Somyong)	–	–
<i>G. max</i> (cv. Pureunkong)	–	–
<i>G. max</i> (cv. Sowon)	–	–
<i>G. max</i> (cv. Danbaek)	–	–
<i>G. max</i> (cv. L29)	–	–
<i>G. max</i> (cv. Daewon)	–	–
<i>G. max</i> (cv. V94-5152)	–	–
<i>G. max</i> (cv. Huanggumkong)	–	–
<i>G. max</i> (cv. Lee74)	–	–
<i>G. soja</i> (wild soybean)	+/ns	+/M, st

<sup>a</sup>Inoculation with the plant sap infected with CMV-209.

<sup>b</sup>Symbols: +, virus infection confirmed by RT-PCR; –, no virus infection confirmed by RT-PCR; ns, no symptom or latent infection; nll, necrotic local lesion; st, stunting; M, mosaic.



**Fig. 3.** Host range of CMV-209. CMV-209 induced severe stunting (A), rosetting (B) in *Pisum sativum*, and necrosis local lesions in inoculated *Vigna sinensis* leaves (C). CMV-Fny did not infect *Pisum sativum* (A). Mock represents healthy plant. CMV-209 and CMV-Fny represent plants infected by CMV-209 and CMV-Fny, respectively. Photographs were taken at 18 days (A, B) and 9 days (C) post-inoculation.

species were inoculated with the sap extracted from the *N. benthamiana* plants systemically infected by the CMV-209 infectious clone. Among the tested legume species, only *G. soja* (wild soybean) and *P. sativum* (pea) were systemically infected with CMV-209 (Table 2). The pea plants infected with CMV-209 showed severe stunting symptom (Table 2; Fig. 3A), rosetting of stem (Fig. 3B) and their roots were significantly shorter and fewer than those of healthy and CMV-Fny infected pea plants (Fig. 3A). CMV-209 inoculated *Vigna sinensis* (cowpea) showed necrotic spots on inoculated leaves (Fig. 3C).

Host adaptation is one of the main factors for sequence convergence or divergence of viruses. CMV-209 could systemically infect pea (*P. sativum*) and *N. benthamiana*, but not the other host plants tested in this study, suggesting that CMV-209 is host-selective strain (Table 2). CMV-209 seems to have adapted itself to infect pea because other typical CMV strain such as CMV-Fny could not infect pea (Fig. 3A). In fact, Hong et al. (2003) also reported that CMV soybean strains might have adapted to soybeans in their evolutionary history and that they had evolved faster than CMV subgroup IA.

Interestingly, CMV-209 also induced necrotic spots in inoculated leaves of *V. sinensis* as other CMV strains did (Fig. 3C). Although two amino acid residues of the 2a protein at positions 631 and 641 being Phe and Ala, respectively, were also conserved in CMV-209, which have been reported to be crucial for causing a local hypersensitive response in cowpea (Kim and Palukaitis, 1997), it remains to be determined whether these two conserved amino acid residues are actually responsible for causing necrotic local lesions. The difference of host range and/or infectivity between the CMV-209 and other reported CMV strains in tested host plants observed in this study might be due to sequence changes in CMV-209 RNA genome segments. In this regard, it is worth mentioning

that the RNA3 segment including 3a and CP proteins has been reported as a host range determinant and thus might be more responsible as host-interactive constraints than the other viral gene products (Hong et al., 2007; Hong et al., 2003; Roossinck, 2002; Ryu et al., 1998). We are currently conducting experiments with reassorted RNA segment using several CMV strains including Fny and 209 to further determine whether RNA3 segment is responsible for determining host range and/or infectivity.

**Table 3.** Sequences obtained from the GenBank database

Virus strains	Subgroup	Database accession number		
		RNA1	RNA2	RNA3
Fny	IA	D00356	D00355	D10538
Rs	IA	AJ511988	AJ517801	AJ517802
Y	IA	D12537	D12538	D12499
Legume	IA	D16403	D16406	D16405
Mf	IA	AJ276479	AJ276480	AJ276481
New Delhi	IB	GU111227	GU111228	GU111229
PI1	IB	AM183114	AM183115	AM183116
NT9	IB	D28778	D28779	D28780
Tfn	IB	Y16924	Y16925	Y16926
LS	II	AF416899	AF416900	AF127976
Trk7	II	AJ007933	AJ007934	L15336
LY	II	AF198101	AF198102	AF198103
Q	II	X02733	X00985	M21464
Bx	III	DQ399548	DQ399549	DQ399550
PHz	III	EU723568	EU723570	EU723569
PSV-P	I	EU570236	EU570237	EU570238
PSV-W	II	U33145	U33146	U31366
PSV-Mi	III	AY429431	AY429430	AY775057
TAV-KC	–	AJ320273	AJ320274	AJ237849

**Table 4.** Genome characteristics of CMV-209

RNAs	Components	Position	Nucleotide (base)	Protein (aa/kDa)
RNA1	5' UTR	1-94	94	–
	ORF 1a	95-3076	2982	993/111.365
	3' UTR	3077-3360	284	–
RNA2	5' UTR	1-77	77	–
	ORF 2a	78-2654	2577	858/97.043
	ORF 2b	2413-2745	333	110/12.846
	3' UTR	2746-3057	312	–
RNA3	5' UTR	1-109	109	–
	ORF 3a	110-949	840	279/30.595
	IR	950-1279	330	–
	ORF CP	1280-1936	657	218/24.136
	3' UTR	1937-2243	307	–

–, noncoding sequence

**Molecular characteristics of CMV-209.** The complete nucleotide (nt) sequence of CMV-209 was determined and deposited in the GeneBank database (Accession Nos: KJ400002, KJ400003, KJ400004). The sizes of the genomic RNAs varied for the different strains and species. The sizes

of RNAs 1, 2, and 3 of the strains in CMV subgroup I were in range of 3357–3365 nt, 3036–3060, and 2213–2220 nt, while in CMV subgroup II were 3389–3391, 3038–3053, 2197–2209 nt, respectively (Palukaitis and Garcia-Arenal, 2003). The length of RNA1, RNA2, and RNA3 of CMV-209 were 3,360 nt, 3,057 nt, and 2,243 nt, respectively, indicating that CMV-209 belongs to CMV subgroup I.

The coding part of the genome consists of five ORFs, i.e. 1a, 2a, 2b, MP, and CP. The first ORF encodes the largest protein (111.65 kDa; 993 aa) which possesses a putative methyltransferase domain and a helicase motif. The second ORF codes the RNA-dependent RNA polymerase (RdRp) protein (97.043 kDa; 858 aa). The third ORF encodes the smallest protein (12.846 kDa; 110 aa), which has function as a silencing suppressor. The fourth is a putative MP (30.595 kDa; 279 aa) and the fifth is the 24.136 kDa CP consisting of 218 aa.

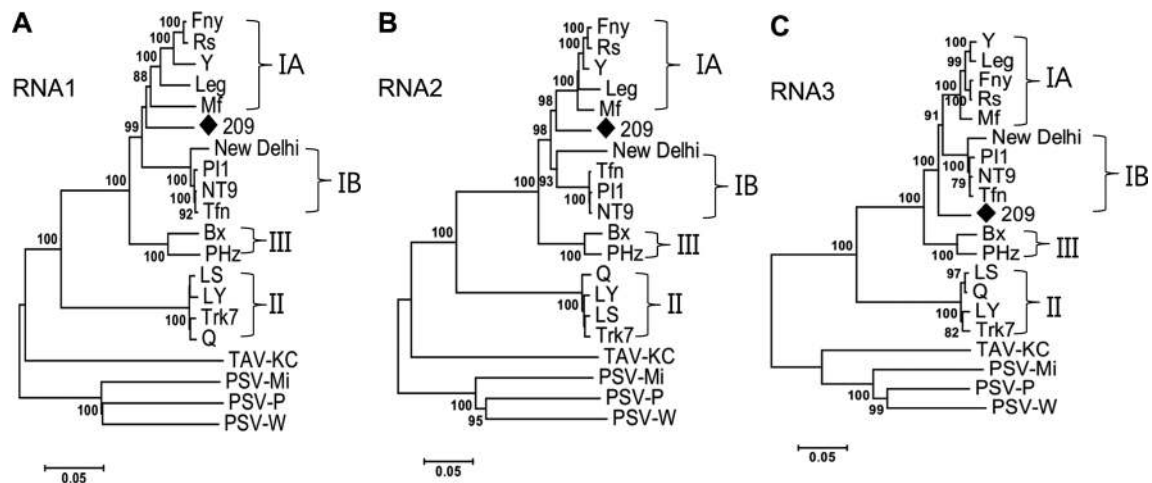
The nucleotide positions of the five ORFs (1a, 2a, 2b, 3a, and CP) for CMV-209 and length of which were presented in Table 4. The initiation codon for the 1a, 2a, 2b, 3b, and CP genes of CMV-209 were positioned at position 95, 78, 2413, 110, and 1280 compared with 95–98, 85–87, 2417–2419, 120–121, and 1257–1260 of subgroup IA and 95–

**Table 5.** Sequence identity between the RNAs, ORFs, and encoded proteins of the CMV-209 strain and those of other strains of CMV, PSV, and TAV

Virus strains	RNA1				RNA2				
	Complete	5' UTR	ORF 1a (nt/aa)	3' UTR	Complete	5' UTR	ORF 2a (nt/aa)	ORF 2b (nt/aa)	3' UTR
Fny (IA)	92.1	94.6	92.4/96.8	88.8	92.4	87.0	92.9/95.2	87.1/80.0	87.8
Rs (IA)	91.9	94.6	92.1/96.5	88.6	92.3	84.4	92.8/95.1	87.7/80.0	88.2
Y (IA)	91.2	91.4	91.5/95.1	87.5	92.8	84.4	93.4/95.6	87.7/80.0	87.8
Legume (IA)	92.1	94.6	92.4/96.4	88.9	91.5	85.5	92.3/94.6	84.0/66.4	86.0
Mf (IA)	91.6	93.5	91.6/94.7	90.3	92.3	87.0	92.9/94.9	87.4/78.2	86.8
New Delhi (IB)	89.6	91.4	89.4/94.0	91.1	90.0	83.1	90.7/91.4	81.8/72.5	88.7
PI1 (IB)	90.5	91.4	90.5/96.3	89.7	90.8	96.1	91.4/93.9	84.8/73.4	86.3
NT9 (IB)	90.6	91.4	90.6/96.2	90.4	90.9	96.1	91.4/93.9	85.2/74.3	87.0
Tfn (IB)	90.5	91.4	90.5/96.2	90.5	91.1	96.1	91.6/93.8	85.2/74.3	86.7
LS (II)	76.4	84.3	77.1/85.1	68.4	71.8	76.6	72.8/75.5	64.7/52.0	66.9
Trk7 (II)	76.4	84.3	77.1/84.6	68.8	71.8	76.6	72.8/75.3	63.4/52.0	66.3
LY (II)	76.3	83.1	77.0/84.4	68.5	71.8	76.6	72.9/75.1	65.0/52.0	65.5
Q (II)	76.4	84.6	77.1/84.4	67.6	72.0	76.6	73.0/74.3	–/–	65.6
Bx (III)	89.5	92.9	89.7/94.4	85.9	88.4	84.7	89.1/87.0	78.4/70.0	86.1
PHz (III)	89.5	96.8	89.7/94.5	84.8	88.3	84.7	89.0/89.7	78.7/70.9	87.1
PSV-P (I)	67.5	65.5	67.7/73.6	65.1	58.5	62.5	58.5/53.8	50.7/35.8	59.0
PSV-W (II)	67.6	62.7	68.2/73.5	62.9	58.4	64.8	58.9/53.9	–/–	–
PSV-Mi (III)	68.7	67.5	68.6/72.1	64.1	59.9	61.1	59.2/55.1	52.2/40.2	66.3
TAV-KC	68.7	72.8	68.5/73.5	69.9	61.5	68.8	60.7/58.8	43.9/28.7	68.8

**Table 5.** Sequence identity between the RNAs, ORFs, and encoded proteins of the CMV-209 strain and those of other strains of CMV, PSV, and TAV (continued)

Virus strains	RNA3					
	Complete	5' UTR	ORF 3a (nt/aa)	IR	ORF 3b (nt/aa)	3' UTR
Fny (IA)	93.0	95.4	91.9/92.8	89.8	94.4/95.9	93.0
Rs (IA)	93.1	95.4	92.0/92.8	90.1	94.2/95.9	93.7
Y (IA)	92.4	89.7	91.5/93.2	90.1	93.6/94.0	93.0
Legume (IA)	92.8	92.5	92.1/93.2	90.1	93.9/95.0	92.3
Mf (IA)	93.0	93.5	92.6/92.8	89.1	94.1/95.4	93.0
New Delhi (IB)	91.7	91.6	89.1/83.2	87.8	94.2/96.3	92.3
PI1 (IB)	92.7	91.6	92.5/92.8	86.3	93.8/95.9	92.3
NT9 (IB)	93.4	93.5	92.9/93.2	88.5	94.4/96.3	92.3
Tfn (IB)	93.2	93.5	93.0/93.2	86.3	94.7/96.8	92.0
LS (II)	75.4	46.4	78.7/81.4	70.9	75.0/81.6	70.7
Trk7 (II)	74.9	46.4	77.3/82.1	69.7	74.3/80.2	71.0
LY (II)	75.2	45.2	78.9/81.7	70.9	75.0/82.0	70.7
Q (II)	75.2	47.0	78.5/81.4	70.9	75.5/82.5	70.7
Bx (III)	89.4	90.6	90.0/91.4	82.4	91.6/94.0	88.0
PHz (III)	89.6	89.4	90.6/91.4	83.9	91.2/94.0	86.9
PSV-P (I)	57.8	40.0	65.7/66.7	48.8	26.5/47.7	62.1
PSV-W (II)	56.1	37.2	45.8/61.5	48.8	52.2/39.7	62.6
PSV-Mi (III)	57.0	39.4	65.0/64.5	45.8	51.7/48.4	63.4
TAV-KC	58.0	40.0	67.2/67.5	58.2	53.3/43.1	67.7

**Fig. 4.** Phylogenetic trees of cucumoviruses obtained from analysis in MEGA 6 for each genomic segment (panels A, B, and C for RNA1, 2, and 3, respectively). Numbers at nodes indicate the percent occurrence of nodes in 1000 bootstrap resampling. Bootstrap values lower than 70 are not indicated. Roman numerals indicate respective CMV subgroups.

96, 79–80, 2414–2415, 122, and 1255–1261 of subgroup IB, respectively. The size (2577 nt) of CMV-209 ORF2a was slightly different from subgroups IA and IB, which was 2574 nt in length. The length of CMV-209 ORF 3a (840

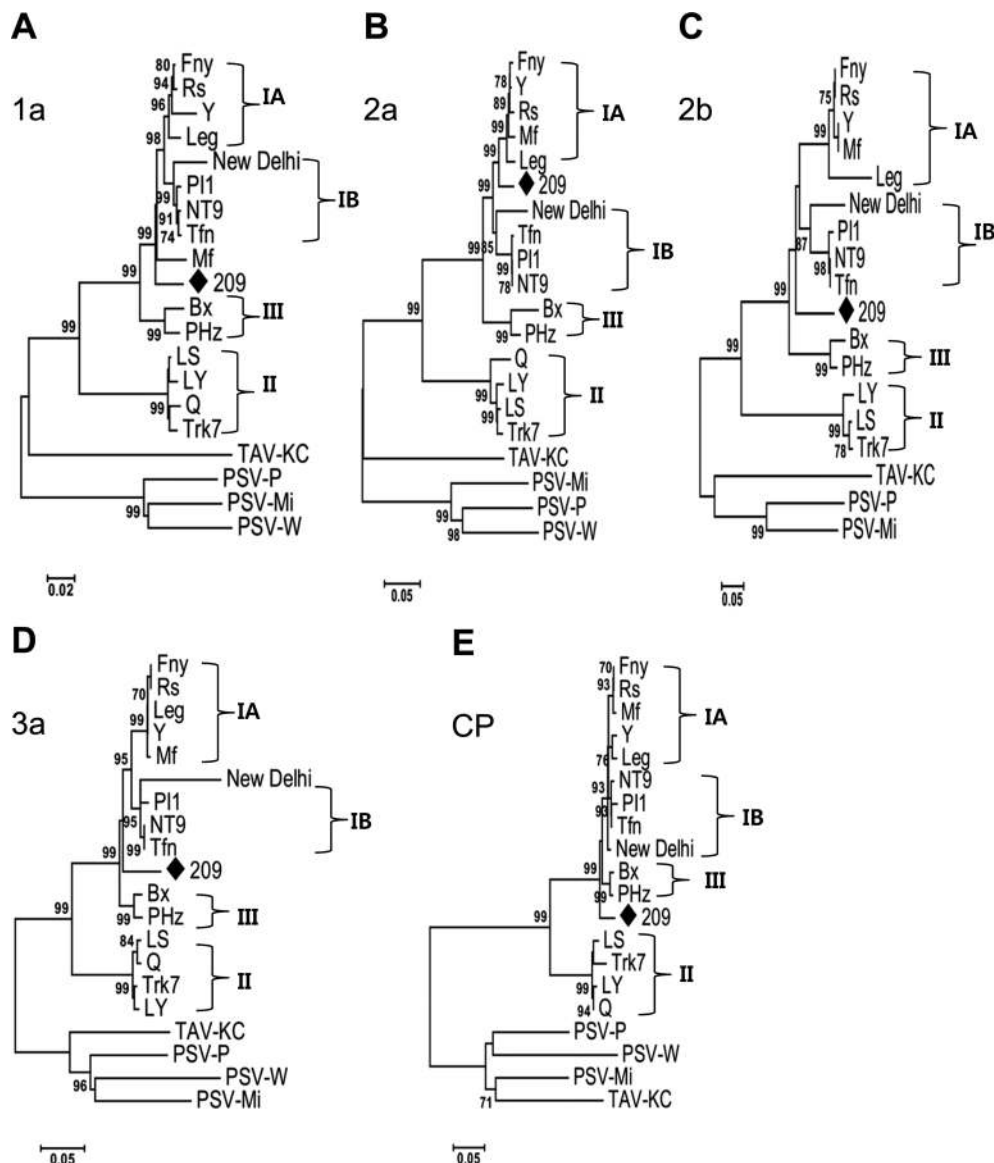
nt) were same as subgroup IA (840 nt) but slightly differed from that of subgroup IB (852 nt). ORFs 1a, 2b, and CP of CMV-209 were 2982 nt, 333 nt, 657 nt, respectively, in length as were in subgroups IA and IB (data not shown).

Further information of the other components of RNAs was indicated in detail in Table 4.

**Comparative analysis of the CMV-209 genome sequence.** According to the International Committee on Virus taxonomy (ICTV) criteria for discriminating species from the genus *Cucumovirus*, subgroups ordinarily have at least 65% sequence identicalness. The nt sequence similarity for the complete genomic RNAs among the CMV-209 and the representative CMV strains was ranged from 71.8 % to 93.4% while the similarity among CMV-209 and other cucumoviruses such as PSV and TAV was in the range

of 56.1–68.7% (Table 5). The genome sequence of CMV-209 had highest similarity with that of CMV-Fny (Table 5).

In CMV, within-subgroup nt homologies of subgroup II and III were very high (more than 98% and 94%, respectively). The nt sequence of CMV-209 had the sequence identity to subgroup II and III only around 76% and 89.5%, respectively (data not shown). In the case of subgroup I, the sequence identities of RNA1, 2, and 3 within subgroup IA were more than 92.7, 95.8, and 97.1%, respectively, and were more than 97.8, 91.2 and 97.4%, respectively, within subgroup IB. The sequence identities of RNA1, 2, and 3 between sister group, subgroup IA and IB, were in the



**Fig. 5.** Phylogenetic trees of cucumoviruses obtained from analysis in MEGA 6 for each ORF amino acid sequences. Numbers at nodes indicate the percent occurrence of nodes in 1000 bootstrap resampling. Bootstrap values lower than 70 are not indicated. Roman numerals indicate respective CMV subgroups.

range of 89.8–92.2, 89.1–91.5, and 91.8–93.9%, respectively (data not shown).

Genomic segment comparison of CMV-209 with the strains in subgroup I, including subgroup IA and IB, showed their identity levels of 89.6–91.2, 90.0–92.4, and 91.7–93.4% in RNA1, 2, and 3 sequences, respectively (Table 5). The 3'UTR of CMV has information required for minus strand synthesis of the viral RNA during replication and is highly conserved within subgroup. However, the 3'UTR nt sequences of CMV-209RNA1, RNA2, and RNA3 showed the similarities of 87.5–91.1, 86.0–88.7, and 92.0–93.7%, respectively, when compared to those of subgroup IA and IB (Table 5). Taken together, our analyses suggested that CMV-209 could belong to a distinct sister group from subgroup IA and IB.

**Phylogenetic analysis.** First, phylogenetic trees for each CMV-209 genomic segment were constructed, in which PSV and TAV were used as out groups. The trees revealed obviously that CMV-209 was clustered in subgroup I. For all three RNAs, a first branching distinguished from clear clustering of group II strains, a second one separated into subgroup III, and the next branching of the trees was precisely divided into subgroups IA, IB, and CMV-209. In the case of RNAs 1 and 2, CMV-209 shared same branching points to subgroup IA (Fig. 4A, B), suggesting that these two RNAs of CMV-209 had closer relationship to subgroup IA than subgroup IB. Although RNA3 of CMV-209 had the same progenitor with subgroup IA and IB, it formed a distinct branch before separation of branches of subgroup IA and IB (Fig. 4C).

Next, phylogenetic analyses of 1a, 2a, 2b, 3a, and CP genes based on amino acid sequences were established. The 2a and 3a protein trees had identical patterns with the full length RNA 2 and 3 nt sequence trees, respectively (Fig. 5B, D). The 1a and 2b protein trees presented a complete segregation of CMV-209 out of subgroups IA and IB (Fig. 5A, C), whereas 1a and 2b ORF nt sequence trees (data not shown) and genomic RNA 1 and 2 trees showed CMV-209 close to subgroup IA. Both CP aa sequence tree and CP nt sequence tree (data not shown) exhibited that CMV-209 was located in a new branch separated from subgroups IA and IB; however, CP protein tree gave a clearer separation. This may suggest that the difference of the biological characteristics among CMV-209 and subgroups IA and IB appeared at amino acid level. In general, amino acid phylogenetic analyses of CMV-209 and the CMV strains had unquestionably placed CMV-209 in another subgroup of subgroup I.

In summary, our analyses provide molecular evidences

that CMV-209 may belongs to a distinct subgroup within CMV subgroup I and that the adaptation of CMV-209 in pea (*P. sativum*) could be due to aa changes of RNA 3 segment based upon sequence analyses. Further study will test whether RNA3 is indeed responsible for adapting CMV-209 in pea by inoculating swapped RNA segments with the other CMV strains and mutant clones which contain site-directed mutations.

## Acknowledgments

This research was supported in part by grants from the Vegetable Breeding Research Center (No. 710001-03) through the Agriculture Research Center program from the Ministry for Food, Agriculture, Forestry and Fisheries; the Next-Generation BioGreen 21 Program (No. PJ00819801), Rural Development Administration (RDA); the Agenda Program (PJ00922904) funded by the RDA, Republic of Korea. MSVP was supported by a research fellowship from the Brain Korea 21 Plus Project.

## References

- Bendahmane, M., Szécsi, J., Chen, I., Berg, R. H. and Beachy, R. N. 2002. Characterization of mutant *Tobacco mosaic virus* coat protein that interferes with virus cell-to-cell movement. *Proc. Natl. Acad. Sci. USA* 99:3645–3650.
- Boyer, J. C. and Haenni, A. L. 1994. Infectious transcripts and cDNA clones of RNA viruses. *Virology* 198:415–426.
- Dopazo, J. 1994. Estimating errors and confidence intervals for branch lengths in phylogenetic trees by a bootstrap approach. *J. Mol. Evol.* 38:300–304.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791.
- Hong, J., Ohnishi, S., Masuta, C., Choi, J. and Ryu, K. 2007. Infection of soybean by *Cucumber mosaic virus* as determined by viral movement protein. *Arch. Virol.* 152:321–328.
- Hong, J. S., Masuta, C., Nakano, M., Abe, J. and Uyeda, I. 2003. Adaptation of *Cucumber mosaic virus* soybean strains (SSVs) to cultivated and wild soybeans. *Theor. Appl. Genet.* 107:49–53.
- Hu, J., Li, H., Barry, K., Wang, M. and Jordan, R. 1995. Comparison of dot blot, ELISA, and RT-PCR assays for detection of two *Cucumber mosaic virus* isolates infecting banana in Hawaii. *Plant Dis.* 79:902–906.
- Ilardi, V., Mazzei, M., Loreti, S., Tomassoli, L. and Barba, M. 1995. Biomolecular and serological methods to identify strains of cucumber mosaic cucumovirus on tomato 1. *EPPO Bulletin* 25:321–327.
- Jacquemond, M. 2012. Cucumber mosaic virus. *Adv. Virus Res.* 84:439–504.
- Kim, C. H. and 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.
- Koshimizu, Y. and Iizuka, N. 1963. Studies on soybean virus diseases in Japan. *Bull. Tohoku Natl. Agric. Exp. Stn.* 27:1–103.
- Lin, H. X., Rubio, L., Smythe, A. B. and Falk, B. W. 2004. Molecular population genetics of *Cucumber mosaic virus* in California: evidence for founder effects and reassortment. *J. Virol.* 78:6666–6675.
- Liu, Y. Y., Yu, S., Lan, Y. F., Zhang, C. L., Hou, S. S., Li, X. D., Chen, X. and Zhu, X. 2009. Molecular variability of five *Cucumber mosaic virus* isolates from China. *Acta Virol.* 53:89–97.
- Masuta, C., Seshimo, Y., Mukohara, M., Jung, H. J., Shigenori, U., Ryu, K. H. and Choi, J. K. 2002. Evolutionary characterization of two lily isolates of *Cucumber mosaic virus* isolated in Japan and Korea. *J. Gen. Plant Pathol.* 68:163–168.
- Nagyová, A. and Subr, Z. 2007. Infectious full-length clones of plant viruses and their use for construction of viral vectors. *Acta Virol.* 51:223–237.
- Nei, M. and Kumar, S. 2000. Molecular evolution and phylogenetics. Oxford University Press, UK. 333 pp.
- Owen, J. and Palukaitis, P. 1988. Characterization of *Cucumber mosaic virus* I. Molecular heterogeneity mapping of RNA 3 in eight CMV strains. *Virology* 166:495–502.
- Owen, J., Shintaku, M., Aeschleman, P., Tahar, S. B. and Palukaitis, P. 1990. Nucleotide sequence and evolutionary relationships of *Cucumber mosaic virus* (CMV) strains: CMV RNA 3. *J. Gen. Virol.* 71:2243–2249.
- Paalme, V., Gammelgård, E., Järvekülg, L. and Valkonen, J. P. 2004. *In vitro* recombinants of two nearly identical potyviral isolates express novel virulence and symptom phenotypes in plants. *J. Gen. Virol.* 85:739–747.
- Palukaitis, P. and García-Arenal, F. 2003. Cucumoviruses. *Adv. Virus Res.* 62:241–323.
- Park, S.-H. and Kim, K.-H. 2006. Agroinfiltration-based *Potato virus X* replicons to dissect the requirements of viral infection. *Plant Pathol. J.* 22:386–390.
- Prüfer, D., Wipf-Scheibel, C., Richards, K., Guilley, H., Lecoq, H. and Jonard, G. 1995. Synthesis of a full-length infectious cDNA clone of *Cucurbit aphid-borne yellows virus* and its use in gene exchange experiments with structural proteins from other luteoviruses. *Virology* 214:150–158.
- Roossinck, M. J. 2002. Evolutionary history of *Cucumber mosaic virus* deduced by phylogenetic analyses. *J. Virol.* 76:3382–3387.
- Roossinck, M. J., Zhang, L. and Hellwald, K. H. 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.
- Ryu, K. H., Kim, C. H. and Palukaitis, P. 1998. The coat protein of *Cucumber mosaic virus* is a host range determinant for infection of maize. *Mol. Plant-Microbe Interact.* 11:351–357.
- Rzhetsky, A. and Nei, M. 1992. A simple method for estimating and testing minimum-evolution trees. *Mol. Biol. Evol.* 9:945–967.
- Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406–425.
- Sala, K. and Bala, E. 1999. Molecular evidence for the existence of two distinct subgroups in cucumber mosaic cucumovirus. *Virus Genes* 18:221–227.
- Senda, M., Masuta, C., Ohnishi, S., Goto, K., Kasai, A., Sano, T., Hong, J. S. and MacFarlane, S. 2004. Patterning of virus-infected *Glycine max* seed coat is associated with suppression of endogenous silencing of chalcone synthase genes. *Plant Cell* 16:807–818.
- Seo, J.-K., Kwon, S.-J., Choi, H.-S. and Kim, K.-H. 2009. Evidence for alternate states of *Cucumber mosaic virus* replicase assembly in positive- and negative-strand RNA synthesis. *Virology* 383:248–260.
- Takahashi, K., Tanaka, T., Iida, W. and Tsuda, Y. 1980. Studies on virus diseases and causal viruses of soybean in Japan. *Bull. Tohoku Nat. Agric. Exp. Stn.* 62:1–130.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A. and Kumar, S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30:2725–2729.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. 1997. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25:4876–4882.
- Wahyuni, W., Dietzgen, R., Hanada, K. and Francki, R. 1992. Serological and biological variation between and within subgroup I and II strains of *Cucumber mosaic virus*. *Plant Pathol.* 41:282–297.