

Complete Genome Sequences and Evolutionary Analysis of *Cucurbit aphid-borne yellows virus* Isolates from Melon in Korea

Hae-Ryun Kwak^{1†}, Hee Ju Lee^{2†}, Eun-A Kim¹, Jang-Kyun Seo³, Chang-Seok Kim¹, Sang Gyu Lee², Jeong-Soo Kim⁴, Hong-Soo Choi¹, and Mikyeong Kim^{1*}

¹Crop Protection Division, National Institute of Agricultural Science, Wanju 55365, Korea

²Vegetable Research Division, National Institute of Horticultural and Herbal Science, Wanju 55365, Korea

³Graduate school of International Agricultural Technology, Seoul National University, Pyeongchang 25354, Korea

⁴Department of Plant Medicine, Andong National University, Andong 36729, Korea

(Received on March 19, 2018; Accepted on July 24, 2018)

Complete genome sequences of 22 isolates of *Cucurbit aphid-borne yellows virus* (CABYV), collected from melon plants showing yellowing symptom in Korea during the years 2013-2014, were determined and compared with previously reported CABYV genome sequences. The complete genomes were found to be 5,680-5,684 nucleotides in length and to encode six open reading frames (ORFs) that are separated into two regions by a non-coding internal region (IR) of 199 nucleotides. Their genomic organization is typical of the genus *Polerovirus*. Based on phylogenetic analyses of complete nucleotide (nt) sequences, CABYV isolates were divided into four groups: Asian, Mediterranean, Taiwanese, and R groups. The Korean CABYV isolates clustered with the Asian group with > 94% nt sequence identity. In contrast, the Korean CABYV isolates shared 87-89% sequence identities with the Mediterranean group, 88% with the Taiwanese group, 81-84% with the CABYV-R group, and 72% with another *polerovirus*, *M.* Recombination analyses identified 24 recombination events (12 different recombination types) in the analyzed CABYV population. In the Korean CABYV isolates, four recombination types were detected from eight isolates. Two recombination types were detected in the IR and

P3-P5 regions, respectively, which have been reported as hotspots for recombination of CABYV. This result suggests that recombination is an important evolutionary force in the genetic diversification of CABYV populations.

Keywords : CABYV, genetic analysis, melon, virus evolution

Handling Associate Editor : Ju, Ho-Jong

Cucurbit aphid-borne yellows virus (CABYV) is an important pathogen that causes yellowing symptoms in cucurbit crops worldwide. CABYV is a member of the genus *Polerovirus* in the family *Luteoviridae* (D'Arcy et al., 2005; Mayo and d'Arcy, 1999). It is transmitted by aphids, mostly *Aphis gossypii* and *Myzus persicae*, with the infection rate of 48-70% in a persistent manner (Kassem et al., 2013; Lecoq et al., 1992). Typical symptoms of CABYV include yellowing and thickening of lower and older leaves. CABYV has been shown to reduce yield by up to 50% in melons (Lecoq et al., 1992). CABYV was first described in melon and cucumber plants in 1992 in France and has since been detected in numerous cucurbit crops in many other countries (Abou-Jawdah et al., 1997; Al Saleh et al., 2015; Bananej et al., 2006; Juárez et al., 2004, 2005; Lecoq et al., 1992; Lemaire et al., 1993; Mnari Hattab et al., 2005; Omar and Bagdady, 2012; Orfanidou et al., 2014; Svoboda et al., 2011; Tomassoli and Meneghini, 2007; Xiang et al., 2008a; Yardımcı and Özgönen, 2007).

CABYV has a single-stranded positive sense RNA genome of approximately 5.7 kb nucleotides (nt) comprising six open reading frames (ORFs) that are separated into two

[†]These authors contributed equally to this work.

*Corresponding author.

Phone) +82-63-238-3301, FAX) +82-63-238-3838

E-mail) mkim00@korea.kr

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

regions by a non-coding internal region (IR) of about 200 nt (D'Arcy et al., 2005). The 5'-proximal ORFs (ORF 0, 1, and 2) are translated from genomic RNA and yield the proteins P0, P1, and the ribosomal frameshift protein P1-P2. The 3'-proximal ORFs (ORF 3, 4, and 5) are translated from subgenomic RNA and yield the proteins P3, P4, and readthrough protein P3-P5. P0 protein is a suppressor of post-transcriptional gene silencing (PTGS) (Pfeffer et al., 2002). P1 has regions of amino acid sequence similarity with serine proteases and genome-linked viral proteins (VPgs) of other polioviruses, and P1-P2 has amino acid motifs typical of RNA-dependent RNA polymerases (RdRP) (Guilley et al., 1994; Mayo and Miller, 1999). P3 is a coat protein (CP); P4 is a movement protein (MP); and P3-P5 is involved in transmission by aphids (Mayo and Miller, 1999).

To date, complete genome sequences of 11 CABYV isolates from France, Spain, China, Japan, and Taiwan have been reported (Kassem et al., 2013; Knierim et al., 2013; Lecoq et al., 1992; Xiang et al., 2008b), as well as partial sequences of ~100 CABYV isolates from several other countries. Based on phylogenetic analysis, CABYV isolates have been divided into two subgroups: the Asian and Mediterranean groups (Shang et al., 2009).

In Korea, CABYV was first detected by next-generation sequencing (NGS) in melons showing yellowing symptoms in 2014, and it was confirmed that the leaf yellowing symptom of melon is not merely a physiological disorder but a viral disease caused by CABYV (Lee et al., 2015).

In this study, we determined the complete genome sequences of 22 CABYV isolates collected from melons showing yellowing symptoms in melon-producing areas during 2013-2014. We analyzed the molecular characteristics and genetic structure of Korean isolates of CABYV in comparison with those of previously reported isolates using a range of methods to understand the evolutionary relationships among isolates.

Materials and Methods

Survey and virus isolates. A survey of CABYV infecting melon was carried out in seven melon-producing areas of Korea during the years 2013-2014 (Fig. 1). We collected 308 samples of melon (*Cucumis melo* L.) leaves that showed yellowing and mosaic symptoms (Table 1). Samples were maintained at -70°C until analysis of CABYV by reverse transcription-polymerase chain reaction (RT-PCR).

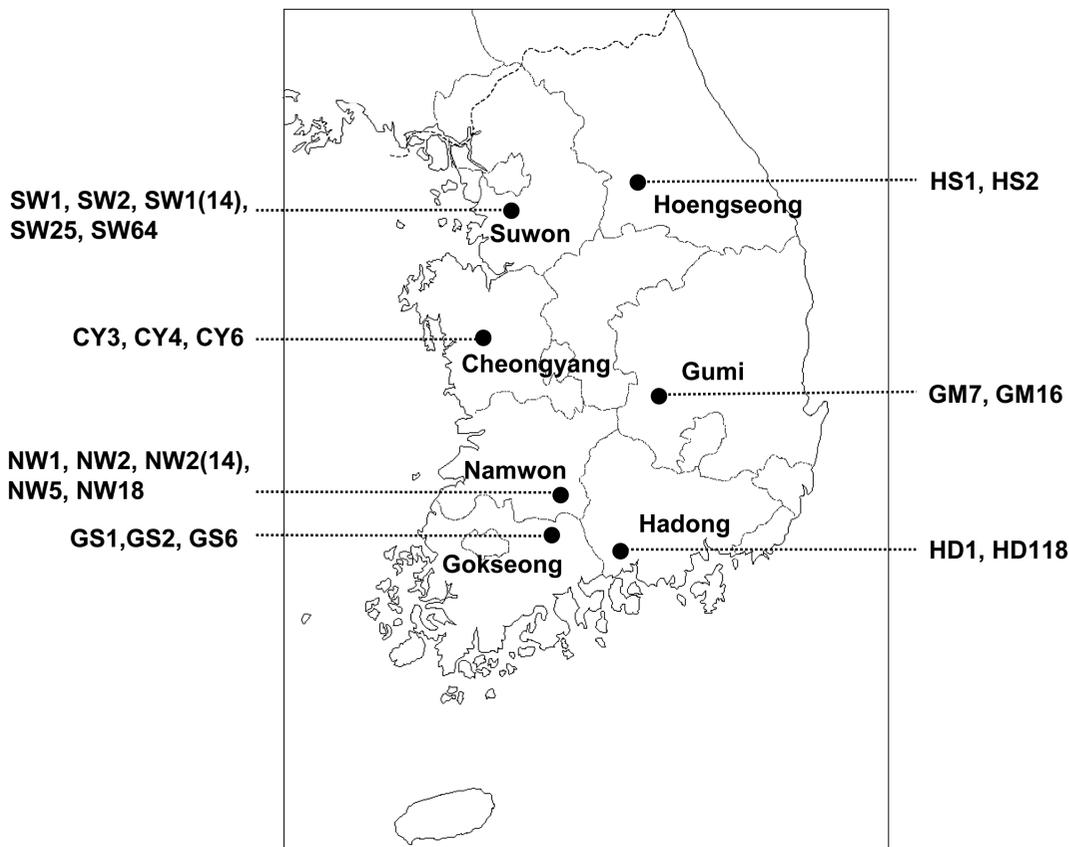


Fig. 1. Geographical locations in Korea at which the CABYV isolates were collected.

Table 1. Survey and detection of CABYV from melon in seven areas in Korea

Area	Date	Sample No.	CABYV No.	Full sequencing	Isolate
Suwon	14.06.09	2	2	2	SW1, SW2
	14.07.11	69	67	3	SW1(14), SW25, SW64
Cheongyang	13.09.06	8	8	2	CY3, CY6
	14.07.25	6	4	1	CY4
Namwon	13.07.26	8	8	2	NW2, NW5
	14.05.09	35	1	1	NW18
	14.08.14	2	2	2	NW1, NW2(14)
Gokseong	14.07.11	1	1	1	GS1
	14.08.14	8	7	2	GS2, GS6
Gumi	13.07.26	21	20	2	GM7, GM16
Hadong	14.09.18	146	123	2	HD1, HD118
Hoengseong	14.08.20	2	2	2	HS1, HS2
All areas		308	245(80%)	22	

Of CABYV-positive samples, the full-length genome sequences of the following 22 CABYV isolates were determined: 5 isolates (SW1, SW2, SW1(14), SW25, and SW64) selected from Suwon in 2014, 3 isolates (CY3, CY6, and CY4) from Cheongyang in 2013 and 2014, 5 isolates (NW2, NW5, NW18, NW1, and NW2(14)) from Namwon in 2013 and 2014, 3 isolates (GS1, GS2, and GS6) from Gokseong in 2013 and 2014, 2 isolates (GM7

and GM16) from Gumi in 2013, 2 isolates (HD1 and HD2) from Hadong in 2014, and 2 isolates (HS1 and HS2) from Hoengseong in 2014 (Table 1 and Fig. 1).

RT-PCR, cloning and sequencing. Total RNA was extracted from infected leaf samples using an Easy-spin™ Total RNA Extraction Kit (Intron, Korea) according to the manufacturer's instructions. RT-PCR was carried out as

Table 2. Primer pairs used for detection and full-length sequencing of the CABYV genome

Primer	Sequence (5'→3')	Loci*	Size (nt)
Primers for detection			
CABYV-u4	ACACGAGTTGCAAGCATTGGAAGT	3341-3364	466
CABYV-d3806	AGTATTCCAGAGCTGAATGCTGGG	3806-3782	
Primers for full-length sequencing			
CABYV-1F-1	ACTATGTTTATAACCCCTGGAGCCAG	214-238	736
CABYV-1R-1	AGTGGGATCTTGTTCCATTCCTGG	950-926	
CABYV-2F	ATATGGTGAAGATGGCGGCTTGG	620-642	1051
CABYV-2R	GAAGCAYTGGTGGTGGGGGAT	1670-1650	
CABYV-3F	ACCACGGCACCCCAAGGACG	1330-1349	1025
CABYV-3R	CCGGTTGAAGGTGAGRCGAGC	2354-2334	
CABYV-4F	GCCCAGTCAGTAAAATCCCCTC	2076-2098	1052
CABYV-4R	ACCGGAATGGCGAGGTCCTC	3127-3108	
CABYV-5F	GTCCCAGGCGTGCAGAAGAG	2892-2911	1031
CABYV-5R	AGCTAAGCTTGCAGTGGGGGTC	3922-3901	
CABYV-6F-1	GGAAGGAGCCCAGGCGAAAC	3679-3698	984
CABYV-6R-1	ATTCGAAGGAAGCGTACCAATCGAC	4663-4639	
CABYV-7F	ACGATGTTTCCCARAGAGGTTGGAA	4496-4520	1022
CABYV-7R	TTAYGAGGTTTTTRTCAGCTAGCACC	5517-5493	
CABYV-5'RACE-R	GCGAGGAAAAATCGCGCAAC	352-333	
CABYV-3'RACE-F	ATGGATARYAGGAAGAAATGGGGA	5314-5337	

* Reference sequence: CABYV-JAN (GQ221224)

either one-step RT-PCR (Genetbio, Korea) for CABYV detection, and two-step RT-PCR including RT using AMV reverse transcriptase (Promega, USA) and PCR using high-fidelity LA Taq polymerase (Takara, Japan) for full-length genome sequencing. Pairs of specific primers for the detection and full-length genome sequencing of CABYV were designed based on previously reported CABYV nucleotide sequences and contig sequences determined by NGS techniques (Lee et al., 2015) (Table 2). cDNA clones containing the 5' end of the genomes were produced using a sense

primer (5'-ACAAAAGATACGAGCGGGTGA TGC-3') complementary to the conserved 24 nt at the 5' terminus and an antisense primer (5'-GCGAGGAAAAATCGCG-CAAC-3') complementary to nt 352-333 in the CABYV genome. In addition, cDNA clones containing the 3' end of the genomes were produced using a sense primer (5'-ATG-GATARYAGGAAGAAATGGGGA-3') complementary to nt 5,314-5,337 and an antisense primer (5'-ACACC-GAAACGCCAGGGGG-3') complementary to the conserved 19 nt at the 3' terminus of the CABYV genome. All

Table 3. Database of the complete nucleotide sequences of CABYV genomes

Virus	Isolate	Host plant	Origin	Genome (nt)	Accession No.	Year collected
<i>Previous studies</i>						
	JAN	cucumber	Japan: Okayama	5682	GQ221224	-
	FJ	squash	China: Fuziang	5682	GQ221223	-
	CHN	cushaw	China: Beijing	5682	EU000535	2006
	Xinjiang	cantaloupe	China: Xinjiang	5682	EU636992	-
	CZ	zucchini	China: Beijing	5691	HQ439023	-
	R-TW82	spong gourd luffa	Taiwan: Tainan	5679	JQ700306	2009
	C-TW20	bitter melon	Taiwan: Kaohsiung	5670	JQ700305	2008
	N	melon	France: Nerac	5669	X76931	1989
	Sq/2003/7.2	squash	Spain: Murcia	5672	JF939812	2003
	Sq/2004/1.9	squash	Spain: Murcia	5672	JF939814	2004
	Sq/2005/9.2	squash	Spain: Murcia	5675	JF939813	2005
<i>This study</i>						
CABYV	SW1	melon	Korea: Suwon	5683	KR231959	2014
	SW2	melon	Korea: Suwon	5683	KR231961	2014
	SW1(14)	melon	Korea: Suwon	5682	KR231960	2014
	SW25	melon	Korea: Suwon	5682	KR231962	2014
	SW64	melon	Korea: Suwon	5681	KR231963	2014
	CY3	melon	Korea: Cheongyang	5683	KR231942	2013
	CY6	melon	Korea: Cheongyang	5682	KR231944	2013
	CY4	melon	Korea: Cheongyang	5684	KR231943	2014
	NW2	melon	Korea: Namwon	5682	KR231955	2013
	NW5	melon	Korea: Namwon	5683	KR231957	2013
	NW18	melon	Korea: Namwon	5683	KR231958	2014
	NW1	melon	Korea: Namwon	5683	KR231954	2014
	NW2(14)	melon	Korea: Namwon	5683	KR231956	2014
	GS1	melon	Korea: Gokseong	5682	KR231947	2014
	GS2	melon	Korea: Gokseong	5683	KR231948	2014
	GS6	melon	Korea: Gokseong	5682	KR231949	2014
	GM7	melon	Korea: Gumi	5683	KR231945	2013
	GM16	melon	Korea: Gumi	5681	KR231946	2013
	HD1	melon	Korea: Hadong	5680	KR231950	2014
	HD118	melon	Korea: Hadong	5683	KR231951	2014
HS1	melon	Korea: Hoengseong	5682	KR231952	2014	
HS2	melon	Korea: Hoengseong	5682	KR231953	2014	
MABYV	CHN	winter melon	China: Beijing	5674	EU000534	2006
	TW1	watermelon	Taiwan: Yunlin	5676	JQ700307	2000

the PCR products were overlapped at least 200 bp to ensure that they were amplified from the same genome. Each PCR fragment was purified using a MEGA Quick-spin™ Kit (Intron, Korea) and cloned into the pGEM-T easy vector (Promega, USA) according to the manufacturer's instructions, followed by transformation into *Escherichia coli* DH5α. The clones of each fragment were completely sequenced by a commercial company (Genotech, Korea). The resultant sequences were assembled using DNA Star v. 5.02 (Lasergene, USA) and have been submitted to GenBank database under the accession numbers listed in Table 3.

Sequence and phylogenetic analyses. The complete nt sequences and the deduced amino acid sequences were aligned using the ClustalX2 program and Geneious methods in Geneious Pro 8 and compared with those of previously reported isolates; i.e., the JAN (Japan), CHN, FJ, Xinjiang, and CZ (China, Xiang et al., 2008b), R-TW82 and C-TW20 (Taiwan, Knierim et al., 2013), N (France, Lecoq et al., 1992), Sq/2003/7.2, Sq/2004/1.9, and Sq/2005/9.2 (Spain, Kassem et al., 2013). The MABYV isolates, CHN and TW1, were used as outgroups (Table 2). Nucleotide and deduced amino acid sequence similarities were analyzed using AlignX implemented in the Vector NTI Suite (Invitrogen, Carlsbad, CA). Pairwise genetic distances and pairwise synonymous (*dS*) and nonsynonymous (*dN*) substitutions were analyzed by Kimura's two-parameter method (Kimura, 1980) and the Pamilo-Bianchi-Li method (Li, 1993; Pamilo and Bianchi, 1993), respectively, using the MEGA6 program (Tamura et al., 2013). The phylogenetic relationships of the CABYV sequences were analyzed by the maximum likelihood (ML) method in MEGA 6. In ML analyses, the phylogenetic trees were constructed using best fit nucleotide substitution models (GTR+G+I for full-length genome and K2+G for IR region and 3' UTR) and best fit amino acid substitution models (JTT+G for all

protein regions). Bootstrap values were calculated using 1,000 random replication. All positions containing gaps and missing data were eliminated. Geneious Pro 8 software was used to calculate the percentage nucleotide and amino acid identities.

Recombination analyses. Recombination events on the full-length sequences of 33 CABYV isolates and 2 MABYV isolates were analyzed using RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, and 3Seq methods implemented in the RDP4 software (Recombination Detection Program, ver. 4) with default settings and a Bonferroni corrected P-value cut-off of 0.01. To reduce the possibility of false detection of recombination, only recombination events supported by at least three methods were selected. To further investigate the putative recombination signals, phylogenetic network analysis was performed using SplitsTree v. 4.1 program (Huson and Bryant, 2006).

Results

Genome characterization of Korean CABYV isolates.

We collected melon leaf specimens that showed yellowing and mosaic symptoms from seven melon-producing areas of Korea during the years 2013-2014. These leaves were analyzed for CABYV using RT-PCR. Of the 308 leaf samples collected, 245 (80%) were positive for CABYV (Table 1). Of the CABYV-positive samples, we selected 22 CABYV isolates based on geographic location, and determined their full-length genome sequences (Table 3). The representative symptoms included yellowing, local chlorosis, mosaic patterns on infected leaves, and informal net formation on fruits (Fig. 2). The complete genomes of Korean CABYV isolates ranged from 5,680 to 5,684 nt, and encoded six open reading frames (ORFs) that were separated into two regions by a non-coding internal region (IR) of 199 nt. Their genomic organization is typical of



Fig. 2. Symptoms induced by CABYV on naturally infected melon plants in Korea. Yellowing (A), leaf mosaic (B), and informal net on melon fruits (C)

members of the genus *Polerovirus*. The 5' and 3' non-coding regions (NCR) are 20 and 164-167 nt in length, respectively. The 5'-proximal ORFs (ORF 0, 1, and 2) encode P0, P1, and the ribosomal frameshift protein P1-P2 with sizes of 239 aa, 631 aa, and 1,056 aa, respectively. The 3'-proximal ORFs (ORF 3, 4, and 5) encode P3 (CP), P4 (MP), and readthrough protein P3-P5 with sizes of 199 aa, 191 aa, and 667-668 aa, respectively (Fig. 3A). The genome organization of these Korean isolates is similar to those of other Asian group CABYVs, including CABYV-JAN and CABYV-CHN.

Genetic diversity in genome region of CABYV population. The molecular variability of 33 isolates of CABYV population, including 22 Korean CABYV isolates and 11 previously reported CABYV isolates, was compared using both complete nucleotide and deduced amino acid sequences. Widespread nucleotide variations were detected throughout the genomes of CABYV. Especially, more significant variations were observed in 3'-UTR and P0, P1-P2 and P3-P5 regions while P3 (CP) and P4 (MP) region were relatively conserved (Figs. 3B and 3C).

Nucleotide diversity for different genomic regions of

the CABYV population was estimated by Kimura's two-parameter method. Low nucleotide diversity values were observed in CP and MP regions and other regions showed the relatively high diversity values (Table 4). Also, pairwise genetic differences at nonsynonymous (dN) and synonymous (dS) nucleotide position were estimated using the Pamilo-Bianchi-Li method. The ratio between nucleotide diversity values in nonsynonymous and synonymous positions (dN/dS) provides an estimation of the degree and direction of the selective constraints acting on the coding regions of CABYV. On the whole, the values of the dN/dS ratio for all the coding regions except P0 were under 1, indicating that these genes are under negative or purifying selection. While, the ratios of dN/dS for P0 gene was greater than 1, considered as evidence for positive selection.

Analysis of phylogenetic relationships. The complete nucleotide and deduced amino acid sequences of 22 Korean CABYV isolates were compared to those of 11 previously reported CABYV isolates. Two MABYV isolates were included as outgroup isolates in the phylogenetic analyses (Table 3). Full-length genome sequence-based phylogenetic analyses revealed that the Korean CABYV isolates

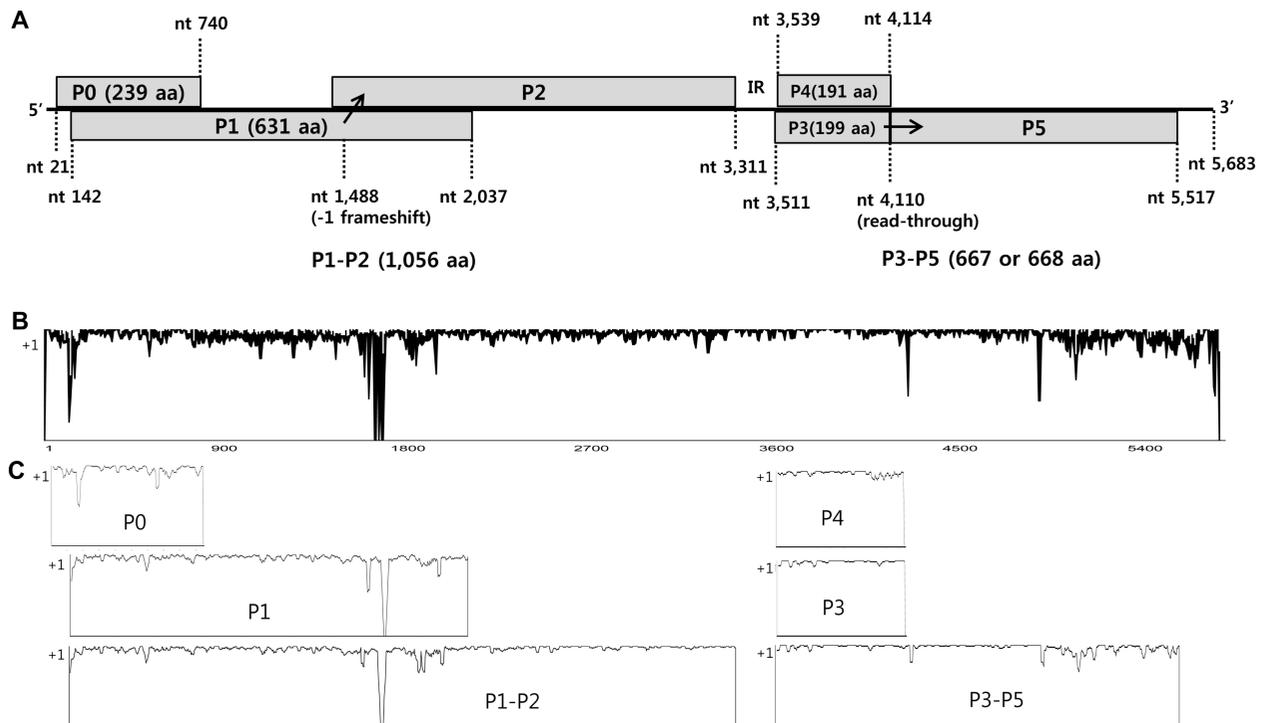


Fig. 3. Genome organization of Korean CABYV isolates (A). The six proteins are separated by IR into two regions: 5' proximal and 3' proximal proteins. A ribosomal frame shift (-1) in the P1-P2 protein is indicated at nt 1,488 and readthrough of the P3-P5 protein occurs at nt 4,110. Nucleotide (B) and deduced amino acid (C) sequence similarities in the CABYV population. Full-length sequences of 33 CABYV isolates were aligned by ClustalX2 and analyzed using AlignX by setting window size to estimate similarities. '+1' on y-axis means that sequences are perfectly conserved.

Table 4. Nucleotide diversity for different genomic regions of the CABYV population

Genomic region	Nucleotide diversity			
	d	d_N	d_S	d_N/d_S
P0	0.0714 ± 0.0049	0.0745 ± 0.0070	0.0536 ± 0.0084	1.3899
P1	0.0898 ± 0.0034	0.0543 ± 0.0030	0.1986 ± 0.0116	0.2734
P1-P2	0.0807 ± 0.0028	0.0638 ± 0.0026	0.1266 ± 0.0069	0.5039
P3(CP)	0.0243 ± 0.0031	0.0132 ± 0.0027	0.0448 ± 0.0077	0.2946
P3-P5	0.0611 ± 0.0030	0.0250 ± 0.0024	0.1579 ± 0.0099	0.1583
P4(MP)	0.0245 ± 0.0033	0.0250 ± 0.0040	0.0262 ± 0.0065	0.9541

clustered with the Asian group, including Japanese and Chinese isolates (Fig. 4).

The reconstructed phylogenetic trees based on amino acid sequences of the six proteins (P0, P1, P1-P2, P3, P4,

and P3-P5) and the nucleotide sequences of two non-coding regions (IR and 3' UTR) showed that the Korean CABYV isolates clustered with the Asian group, similar to the tree based on nucleotide sequences (Supplementary Fig. 1 and

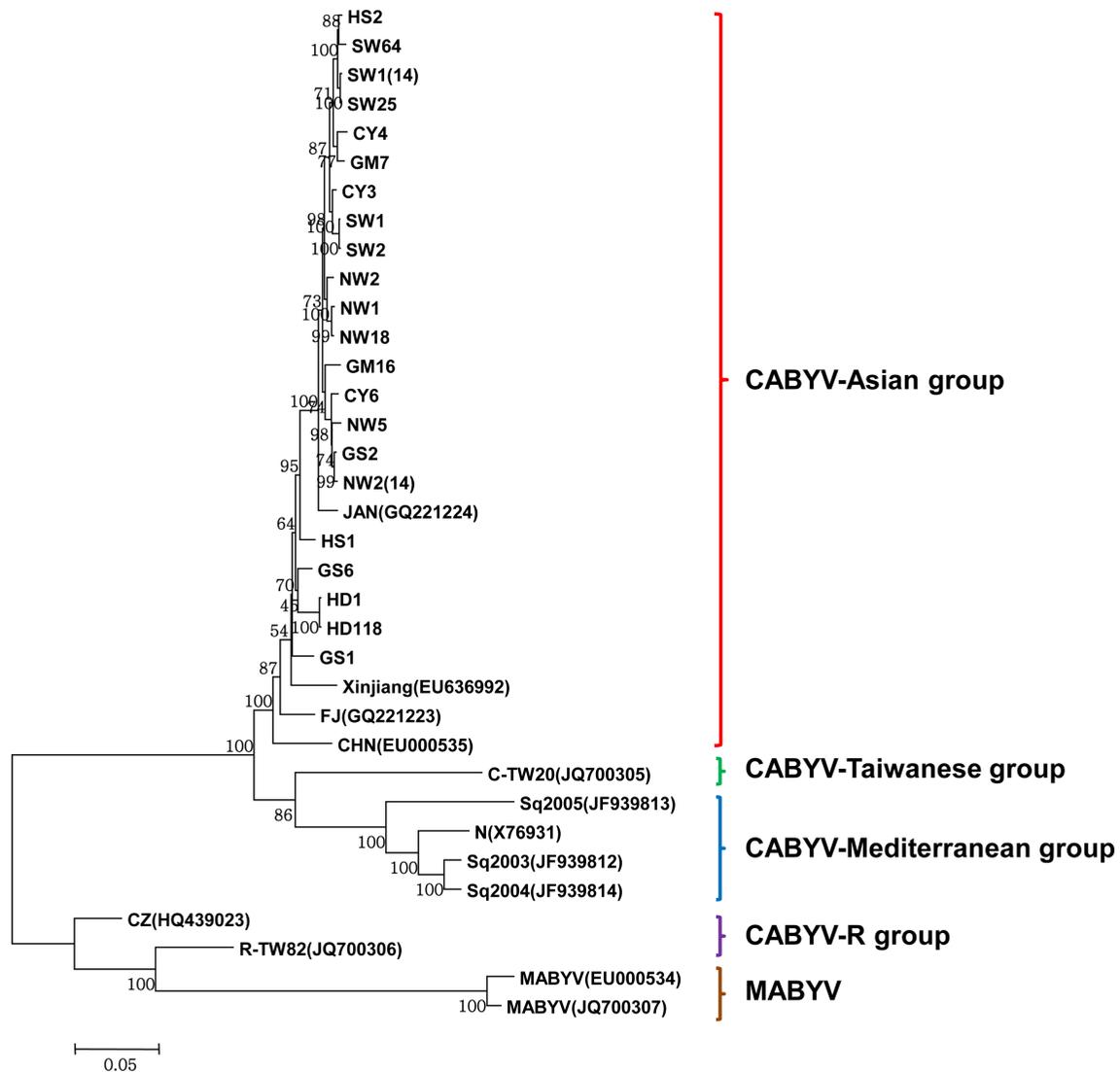


Fig. 4. Phylogenetic trees reconstructed using the complete nucleotide sequences of the CABYV isolates. Phylogenetic trees were reconstructed using maximum likelihood in MEGA 6.

2). However, in the case of the 3' proximal proteins and 3' UTR, the Korean CABYV isolates were differentiated into two subgroups within the Asian group. In addition, the Chinese isolate CZ and Taiwanese isolate R-TW82, which belong to the CABYV-R group, grouped with MABYV based on full-length genome sequences and amino acids of 5' proximal ORFs, but grouped with CABYV based on 3' proximal ORFs. These results suggest that recombination occurred within CABYV isolates and between CABYV and MABYV isolates.

Sequence comparison. The nucleotide and amino acid sequence identities between CABYV isolates are summarized in Table 5. For the full-length genome nucleotide sequences, Korean CABYV isolates had 96-99% nt sequence similarity. CABYV-CY3, a Korean CABYV isolate, showed 94-98% nt sequence similarity with the Asian group including Japanese and Chinese isolates, 87-89% with the Mediterranean group, 88% with the Taiwanese group, 81-84% with the CABYV-R group, and 72% with the other polerovirus, MABYV.

Regarding the deduced amino acid sequences of six indi-

Table 5. Nucleotide and amino acid sequence identities (%) between the Korean CABYV isolate CY3 and other CABYV isolates

Virus	Isolate	Full genome	P0	P1	P1-P2	IR	P3	P4	P3-hP5	3'UTR
		(nt)	(aa)	(aa)	(aa)	(nt)	(aa)	(aa)	(aa)	(nt)
CABYV	CY4	98.8	97.5	99.2	99.0	100	99.0	98.4	98.7	98.8
	CY6	98.5	95.8	97.6	98.1	100	99.5	99.5	99.0	98.8
	GM7	98.9	97.5	99.5	99.4	100	99.5	99.5	99.0	99.4
	GM16	98.3	95.4	97.6	98.4	99.5	99.5	99.0	98.8	98.2
	GS1	96.7	97.1	98.4	98.8	98.5	99.5	98.4	95.7	92.2
	GS2	98.7	97.5	98.6	99.1	100	99.0	99.0	98.8	99.4
	GS6	97.5	96.7	98.4	99.1	100	99.5	99.5	96.0	92.3
	HD1	96.8	96.2	98.1	98.2	99.5	97.0	97.4	95.2	92.2
	HD118	96.8	96.2	98.4	98.3	99.5	96.5	97.4	95.2	92.2
	HS1	97.4	98.3	98.7	99.1	99.5	99.0	99.0	95.8	91.0
	HS2	99.0	97.9	99.4	99.2	99.5	99.0	99.0	98.8	98.8
	NW1	98.9	97.1	98.9	99.1	100	98.5	98.4	98.7	99.4
	NW2	98.9	97.1	99.0	99.1	100	98.5	99.5	98.5	98.2
	NW2(14)	98.7	97.1	98.6	99.1	100	99.5	99.0	98.5	99.4
	NW5	98.4	97.1	98.6	98.7	100	99.0	98.4	98.5	98.8
	NW18	98.9	97.9	98.9	99.1	100	99.0	99.0	98.8	99.4
	SW1	99.3	97.5	99.5	99.5	100	99.5	99.5	99.3	99.4
	SW1(14)	99.0	97.5	99.4	99.4	100	98.5	99.0	98.4	98.8
	SW2	99.3	97.5	99.5	99.5	100	99.5	99.5	99.3	99.4
	SW25	99.0	97.5	99.5	99.5	100	99.5	99.5	98.5	98.8
	SW64	98.8	97.1	99.5	99.3	100	99.0	97.9	98.8	97.6
	JAN	97.9	96.7	97.9	98.9	100	99.0	99.5	99.0	92.2
	FJ	95.7	92.1	96.2	96.8	97.5	98.5	96.9	95.5	90.4
	CHN	94.7	91.6	93.7	95.4	97.5	97.5	97.9	95.2	91.0
	Xinjiang	95.6	96.2	95.7	95.7	98.5	98.0	97.4	96.0	92.2
	CZ	83.8	81.0	64.4	75.0	98.0	97.0	98.4	94.9	91.6
	R-TW82	81.4	82.4	66.2	75.0	92.0	95.0	88.5	90.4	81.4
	C-TW20	87.5	76.2	83.7	89.5	92.5	96.5	86.9	90.0	86.1
	N	89.0	80.3	88.0	91.9	94.5	94.5	89.0	90.4	69.0
	Sq/2003/7.2	88.8	81.6	87.3	90.9	94.0	95.0	89.6	91.0	68.5
Sq/2004/1.9	89.1	81.6	87.0	91.2	95.0	92.0	90.6	90.9	68.5	
Sq/2005/9.2	87.1	74.9	83.4	87.3	94.0	95.5	86.9	89.4	68.5	
MABYV	CHN	71.5	73.1	62.4	72.7	69.2	81.5	67.0	62.0	84.4
	TW1	71.8	73.6	62.6	72.5	70.1	81.5	66.5	62.5	83.8

Table 6. Recombination in Korean CABYV populations

Type	Recombinant isolate	Recombination site in genome		Genes affected	Parental isolates ^a	RDP4 ^b	P-value ^c
		Start	End				
1	CZ	nt 36	nt 236	P0	CHN × MABYV	RGMC	5.740 × 10 ⁻²⁰
2	R-TW82	nt 703	nt 3391	P1, P1-P2, IR	Xinjiang × MABYV	RGBMCS3	2.048 × 10 ⁻¹³⁸
3	Sq/2003/7.2, Sq/2004/1.9	nt 1292	nt 1472	P1	Sq2005 × FJ	MCS	7.641 × 10 ⁻⁰⁴
4	N	nt 1292	nt 1629	P1	Sq2005 × GS6	MCS	8.382 × 10 ⁻⁰⁶
5	HS1, NW1, NW2, NW18, JAN	nt 1176-1395	nt 2466-3153	P1, P1-P2	GM7 × CY6	RMCS3	9.765 × 10 ⁻⁰⁶
6	Xinjiang	nt 1419	nt 3139	P1, P1-P2	GS1 × N	BMC	1.181 × 10 ⁻¹⁴
7	GS6	nt 1653	nt 259	P0, P1	GS2 × SW1(14)	RGBMS3	5.173 × 10 ⁻⁰⁸
8	HD1, HD118	nt 3388	nt 11	IR, P3, P4, P3-P5, 3'UTR	HS2 × CZ	RGBMCS3	1.055 × 10 ⁻¹⁹
9	GS1, HS1, FJ, Xinjiang	nt 4601-4893	nt 5653-11	P3-P5, 3'UTR	HS2 × CZ	RGBMCS3	1.055 × 10 ⁻¹⁹
10	Sq/2005/9.2	nt 4820	nt 7	P3-P5, 3'UTR	SW25 × Sq2004	RGBMCS3	6.448 × 10 ⁻³⁸
11	R-TW82	nt 5193	nt 5451	P3-P5	C-TW20 × GS1	GMC	3.153 × 10 ⁻⁰⁵
12	N, Sq/2003/7.2, Sq/2004/1.9, 2005/9.2	nt 5508-5513	nt 5608-5614	3'UTR	SW64 × MABYV	RGB	5.129 × 10 ⁻⁰⁹

^a 'Parental isolates' indicates the most likely isolates among those analyzed; Major parent × minor parent.

^b RDP4-implemented methods that supported the corresponding recombination site: R (RDP), G (GENECONV), B (BootScan), M (MaxChi), C (Chimaera), and S (SiScan), 3 (3Seq).

^c The highest P-value among the RDP4-implemented methods is reported. The corresponding method is shown boldface.

vidual proteins, CABYV-CY3 (as a representative Korean CABYV isolate) showed relatively high sequence identity of 92-100% with the Asian group. In contrast, aa sequence identities between CABYV-CY3 and the Mediterranean group were 75-82% for P0, 83-88% for P1, 87-92% for P1-P2, 92-97% for P3(CP), 87-91% for P4(MP), and 89-91% for P3-P5. In comparison with CABYV-CZ and R-TW82, CABYV-CY3 had lower aa sequence identity of 65-75% for the 5' proximal proteins (P1 and P1-P2), but 89-98% for the 3' proximal proteins (P3, P4 and P3-P5). In addition, CABYV-CY3 shared only 62-82% aa sequence identity with MABYV isolates for each individual protein.

The nt sequence identities of the IR region and 3' UTR were 92-100% and 69-92%, respectively, among the four CABYV groups, and were 70% and 84% with MABYV.

Recombination analysis. Recombination has been shown to significantly contribute to luteovirus diversity. To examine whether recombination events have occurred in the CABYV population, we aligned full-length nt sequences of 33 CABYV and 2 MABYV isolates using the Geneious method in Geneious Pro 8 and analyzed them using the RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan and 3Seq methods implemented in the RDP4 software with a highest acceptable P-value of 0.01. In total, 56 potential recombinant events were detected by at least one method; however, to reduce error, we included only recombination events supported by at least three methods. Using this criterion, 24 recombination events, including 12 recombination types, were detected in 17 CABYV isolates (Table 6). Among the Korean CABYV isolates, nine recombination events were detected in eight isolates of types 5, 7, 8, and 9. In particular, isolates HD1 and HD118 of recombination type 8 were detected as recombinants between the major parent HS2 and minor parent CZ. In this recombination event, the genomic region (nt 3,388-11) was replaced with the homologous region of CZ. This result could explain why these isolates belonged to the same group, which was distantly related to the other Korean CABYV isolates in the phylogenetic trees based on amino acid sequences of the 3' proximal proteins (P3, P4, and P3-P5). The other recombinant isolates, GS1 and HS1, had the same parental isolates as the Chinese isolates, FJ and Xinjiang, with a type 8 recombination event, but their recombination was detected in the P3-P5 region (nt 4,601-4,893). These two regions, IR and P3-P5, were identified as hotspots for recombination of CABYV. On the other hand, as expected from phylogenetic and sequence analyses, the Chinese isolate CZ and Taiwanese isolate R-TW82 were reconfirmed as recombinants of CABYV and MABYV with P-values of 5.740 ×

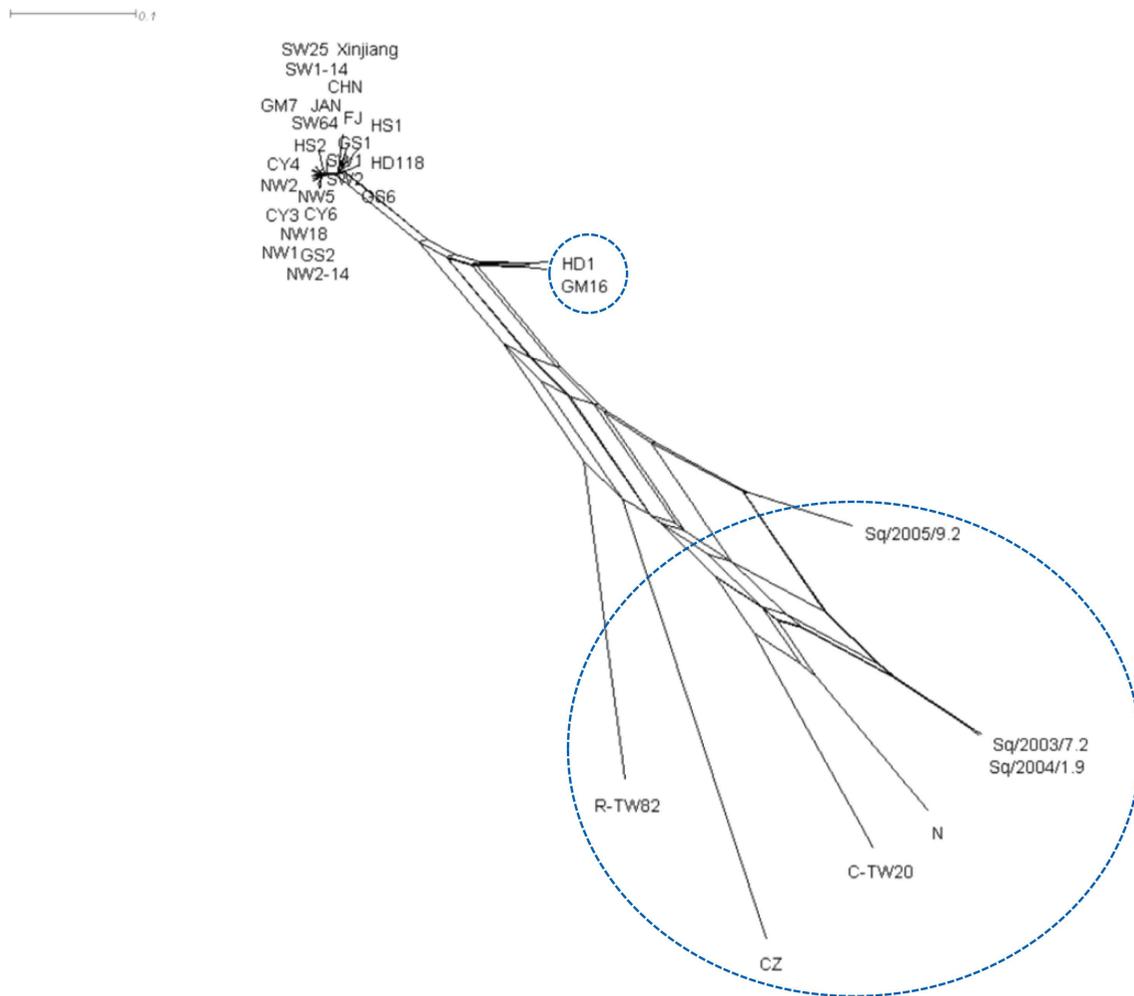


Fig. 5. Split decomposition network of the CABYV population. Phylogenetic network analysis was performed using SplitsTree v. 4.1 program.

10^{-20} and 2.048×10^{-138} , respectively.

To further confirm the recombination, phylogenetic network analysis was performed using SplitsTree v. 4.1 program. The split decomposition analysis revealed that nine tentative recombinants formed a reticulate network structure (Fig. 5). Seven isolates except GM16 and C-TW20 were detected as recombinants by both RDP4 and SplitsTree v. 4.1 programs.

Discussion

Recently, CABYV was detected in Korea in melons showing yellowing symptoms using NGS and RT-PCR. Of 308 melon samples surveyed in seven areas during 2013-2014, 245 (80%) were positive for CABYV. To investigate the genomic structure, genetic diversity, and the possible origin of Korean CABYV population, we determined the full genome sequences of 22 CABYV isolates from CABYV-

positive melon samples and analyzed their genetic diversity by comparison with the sequences of 11 CABYV isolates and 2 MABYV isolates as outgroups. The complete genomes of the Korean CABYV isolates range between 5,680 to 5,684 nt and encode six open reading frames (ORFs) that are separated into two regions by a non-coding internal region (IR) of 199 nt. The genomic organization of these isolates is typical of the genus *Polerovirus*. The deduced amino acid sizes of five proteins, the exception being the P3-P5 protein, were identical to those of the Asian CABYV group, which includes Japanese and Chinese isolates. Most Korean CABYV isolates had a P3-P5 protein of 668 aa, while some isolates comprised 667 aa due to the lack of one proline in the 5' terminal region of P5.

Sequence comparison revealed that the Korean CABYV isolates shared 95-98% nt sequence identity and 92-100% aa sequence identities for six individual proteins with the Asian group (Table 5). In addition, the Korean CABYV

isolates showed 82-89% nt sequence identity and 75-98% aa sequence identity with three other CABYV groups. Of the individual proteins, the 5' proximal proteins were more variable than the 3' proximal proteins. In particular, P0 was the most variable, while P3 (CP) was the most conserved. These characteristics; i.e., highly variable P0 and conserved P3 (CP), have been reported for other poliovirus species (Hauser et al., 2000; Huang et al., 2005; Xiang et al., 2010).

Using phylogenetic analyses based on full-length genome sequences, the Korean CABYV isolates clustered in the Asian group (Fig. 4). According to previous reports, CABYV isolates are divided into two groups that cluster geographically: the Asian and Mediterranean groups (Shang et al., 2009). However, our phylogenetic results suggest that CABYV isolates are divided into four groups: Asian, Mediterranean, Taiwanese, and R groups. Phylogenetic trees reconstructed using the amino acid sequences of six individual proteins and nt sequences of non-coding regions showed that Korean CABYV isolates consistently grouped into the Asian group. However, using 3' proximal proteins and the 3' UTR, the Korean CABYV isolates were classified into two subgroups within the Asian group. This shows the possibility of recombination in the IR region between the 5' proximal and 3' proximal proteins. Although the CABYV-R group belonged to CABYV, it grouped into MABYV in the 5' proximal protein and 3' UTR-based phylogenetic analyses, due to recombination between CABYV and MABYV isolates (Knierim et al., 2013).

To confirm the results of sequence and phylogenetic analyses, we investigated whether recombination occurred in the CABYV population using the RDP4 software. Twenty-four recombination events of 12 recombination types were detected in the analyzed CABYV population. Among them, nine recombination events occurred in the Korean CABYV isolates. Two recombination types in particular, 8 and 9, were detected in regions IR and the P3-P5 readthrough protein, respectively. These two regions have been reported as hotspots of RNA recombination in the family *Luteoviridae*, including CABYV (Gibbs and Cooper, 1995; Huang et al., 2005; Shang et al., 2009). In addition, recombination types 7 and 8 were detected as recombinants between HS2 as a major parent and CZ as a minor parent. This result could explain why the Korean CABYV isolates differentiated into two groups in phylogenetic trees based on the aa sequences of the 3' proximal proteins (P3, P4, and P3-P5). Especially, some of CABYV isolates detected as tentative recombinants by RDP4 were consistently confirmed as recombinants by split decomposition analysis. Collectively, our results suggest that recombination is a major evolutionary force in the genetic

diversification of the CABYV population in Korea.

In the present study, we analyzed the genetic diversity and structure of the CABYV population collected from melon plants. Our findings revealed that the Korean CABYV isolates belong to the CABYV-Asian group and that their genetic diversity is generated by recombination, as well as accumulation of mutations. Understanding the molecular characterization of viruses is essential for the development of strategies for the virus control.

CABYV, an important pathogen that causes yellowing symptoms in cucurbit crops, has been reported to infect nine cucurbit crops in China (Xiang et al., 2008a). In Korea, many cucurbit species are widely cultivated, and it has been confirmed that CABYV infected cucumber and oriental melon (Choi et al., 2015). Recently, we also could confirm that some watermelons and pumpkins showing mosaic and yellowing symptoms were co-infected with CABYV and other viruses including *Watermelon mosaic virus* or *Zucchini yellow mosaic virus*. CABYV became as one of the major viruses damaging cucurbits in Korea. However, knowledge of host range, pathogenicity, and vector transmission of Korean CABYV isolates is still limited. Further studies are needed for the aim of preventing the spread of CABYV.

Acknowledgements

This research was supported by a grant from the Agenda Program (PJ01130602), funded by the Rural Development Administration of Korea.

References

- Abou-Jawdah, Y., Sobh, H. and Fayyad, A. 1997. First report of cucurbit aphid-borne yellows luteovirus in Lebanon. *Plant Dis.* 81:1331.
- Al Saleh, M. A., Al-Shahwan, I. M., Amer, M. A., Shakeel, M. T., Kamran, A., Xanthis, C. K., Orfanidou, C. G. and Katis, N. I. 2015. First report of Cucurbit aphid-borne yellows virus in cucurbit crops in Saudi Arabia. *Plant Dis.* 99:894.
- Bananej, K., Desbiez, C., Wipf-Scheibel, C., Vahdat, I., Kheyr-Pour, A., Ahoonmanesh, A. and Lecoq, H. 2006. First report of Cucurbit aphid-borne yellows virus in Iran causing yellows on four cucurbit crops. *Plant Dis.* 90:526.
- D'Arcy, C. J. and Domier, L. L. 2005. Family luteoviridae. In: *Virus taxonomy: Classification and nomenclature of viruses. Eighth report of the international committee on taxonomy of viruses.* eds. by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger and L. A. Ball, pp. 891-900. Elsevier, Academic Press, Netherlands.
- Gibbs, M. J. and Cooper, J. I. 1995. A recombinational event in the history of luteoviruses probably induced by base-

- pairing between the genomes of two distinct viruses. *Virology* 206:1129-1132.
- Guilley, H., Wipf-Scheibel, C., Richards, K., Lecoq, H. and Jonard, G. 1994. Nucleotide sequence of cucurbit aphid-borne yellows luteovirus. *Virology* 202:1012-1017.
- Hausser, S., Stevens, M., Mougél, C., Smith, H. G., Fritsch, C., Herrbach, E. and Lemaire, O. 2000. Biological, serological, and molecular variability suggest three distinct polerovirus species infecting beet or rape. *Phytopathology* 90:460-466.
- Huang, L. F., Naylor, M., Pallett, D. W., Reeves, J., Cooper, J. I. and Wang, H. 2005. The complete genome sequence, organization and affinities of carrot red leaf virus. *Arch. Virol.* 150:1845-1855.
- Huson, D. H. and Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23:254-267.
- Juárez, M., Truniger, V. and Aranda, M. A. 2004. First report of Cucurbit aphid-borne yellows virus in Spain. *Plant Dis.* 88:907.
- Juárez, M., Kassem, M. A., Sempere, R. N., Truniger, V., Moreno, I. M. and Aranda, M. A. 2005. Cucurbit aphid-borne yellows virus (CABYV): a new virus found in cucurbit crops of Southeastern Spain. *Bol. San. Veg. Plagas* 31:587-598 (in Spanish).
- Kassem, M. A., Juárez, M., Gómez, P., Mengual, C. M., Sempere, R. N., Plaza, M., Elena, S. F., Moreno, A., Fereres, A. and Aranda, M. A. 2013. Genetic diversity and potential vectors and reservoirs of Cucurbit aphid-borne yellows virus in southeastern Spain. *Phytopathology* 103:1188-1197.
- 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.
- Knierim, D., Tsai, W. S., Deng, T. C., Green, S. K. and Kenyon, L. 2013. Full-length genome sequences of four polerovirus isolates infecting cucurbits in Taiwan determined from total RNA extracted from field samples. *Plant Pathol.* 62:633-641.
- Lecoq, H., Bourdin, D., Wipf-Scheibel, C., Bon, M., Lot, H., Lemaire, O. and Herrbach, E. 1992. A new yellowing disease of cucurbits caused by a luteovirus, cucurbit aphid-borne yellows virus. *Plant Pathol.* 41:749-761.
- Lee, H. J., Kim, M. K., Lee, S. G., Choi, C. S., Choi, H. S., Kwak, H. R., Choi, G. S. and Chun, C. 2015. Physiological characteristics of melon plants showing leaf yellowing symptoms caused by CABYV infection. *Korean J. Hortic. Sci. Technol.* 33:210-218.
- Lemaire, O. J., Gubler, W. D., Valencia, J., Lecoq, H. and Falk, B. W. 1993. First report of cucurbit aphid-borne yellows luteovirus in the United States. *Plant Dis.* 77:1169.
- Li, W. H. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J. Mol. Evol.* 36:96-99.
- Mayo, M. A. and D'Arcy, C. J. 1999. Family Luteoviridae: a reclassification of luteoviruses. In: *The Luteoviridae*, eds. by H. G. Smith and H. Barker, pp. 15-22. CABI Publishing, Wallingford, UK.
- Mayo, M. A. and Miller, W. A. 1999. The structure and expression of luteovirus genomes. In: *The Luteoviridae*, eds. by H. G. Smith and H. Barker, pp. 23-42. CABI Publishing, Wallingford, UK.
- Mnari Hattab, M., Kummert, J., Roussel, S., Ezzaier, K., Zouba, A. and Jijakli, M. H. 2005. First report of Cucurbit aphid-borne yellows virus in Tunisia causing yellows on five cucurbitaceous species. *Plant Dis.* 89:776.
- Omar, A. F. and Bagdady, N. A. 2012. Cucurbit aphid-borne yellows virus in Egypt. *Phytoparasitica* 40:177-184.
- Orfanidou, C., Maliogka, V. I. and Katis, N. I. 2014. First Report of Cucurbit chlorotic yellows virus in Cucumber, Melon, and Watermelon in Greece. *Plant Dis.* 98:1446.
- Pamilo, P. and Bianchi, N. O. 1993. Evolution of the Zfx and Zfy genes: Rates and interdependence between the genes. *Mol. Biol. Evol.* 10:271-281.
- Pfeffer, S., Dunoyer, P., Heim, F., Richards, K. E., Jonard, G. and Ziegler-Graff, V. 2002. P0 of beet Western yellows virus is a suppressor of posttranscriptional gene silencing. *J. Virol.* 76:6815-6824.
- Shang, Q. X., Xiang, H. Y., Han, C. G., Li, D. W. and Yu, J. L. 2009. Distribution and molecular diversity of three cucurbit-infecting poleroviruses in China. *Virus Res.* 145:341-346.
- Svoboda, J., Leisova-Svobodova, L. and Lecoq, H. 2011. First Report of Cucurbit aphid-borne yellows virus in Squash in the Czech Republic. *Plant Dis.* 95:220.
- 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.
- Tomassoli, L. and Meneghini, M. 2007. First report of Cucurbit aphid-borne yellows virus in Italy. *Plant Pathol.* 56:720.
- Xiang, H. Y., Shang, Q. X., Han, C. G., Li, D. W. and Yu, J. L. 2008a. First report on the occurrence of Cucurbit aphid-borne yellows virus on nine cucurbitaceous species in China. *Plant Pathol.* 57:390.
- Xiang, H. Y., Shang, Q. X., Han, C. G., Li, D. W. and Yu, J. L. 2008b. Complete sequence analysis reveals two distinct poleroviruses infecting cucurbits in China. *Arch. Virol.* 153:1155-1160.
- Xiang, H. Y., Dong, S. W., Zhang, H. Z., Wang, W. L., Li, M. Q., Han, C. G., Li, D. W. and Yu, J. L. 2010. Molecular characterization of two Chinese isolates of Beet western yellows virus infecting sugar beet. *Virus Genes* 41:105-110.
- Yardımcı, N. and Özgönen, H. 2007. First report of Cucurbit aphid-borne yellows virus in Turkey. *Aust. Plant Dis. Notes* 2:59.