SHORT COMMUNICATION

COMPLETE GENOMIC SEQUENCE ANALYSES OF *WATERMELON MOSAIC VIRUS* ISOLATES FROM CHINA

L. Zhao[1], C. Feng[2], X. Hao[1], M. Sou [1], J. Zhang[1, 3], Q. Wang[2] and Y. Wu[1]

[1]College of Plant Protection and [2]College of Horticulture, State Key Laboratory of Crop Stress Biology for Arid Areas, Key Laboratory of Crop Pest Integrated Pest Management on Crop in Northwestern Loess Plateau, Ministry of Agriculture, Key Laboratory of Plant Protection Resources and Pest Management, Ministry of Education, Northwest A&F University, Yangling, 712100, P. R. China
[3]College of Life Science, Henan Normal University, HeNan, XinXiang, 453007, P. R. China

SUMMARY

The complete genomic sequences of *Watermelon mosaic virus* (WMV) of a Chinese isolate (WMV-CHN) and a Shaanxi isolate (WMV-Shaanxi) were each determined from 10 overlapping fragments. Excluding the poly(A) tail, the genomes of the two isolates were 10,037 and 10,046 nucleotides in length and contained one single open reading frame of 9,651 and 9,657 nucleotides, respectively. The identity frequencies between the two complete genomes were 96.5% at the nt level and 98.5% at the amino acid level. The two complete genomic sequences shared 89.7-94.8% identity with other characterized WMV isolates at the nucleotide level and 94.6 to 97.1% at the amino acid level.

Keywords: genome sequence, virus evolution, *Watermelon mosaic virus*.

Watermelon (*Citrullus lanatus*) is an economically horticultural crop and widely grown in China from south (Hainan) to north (Inner Mongolia) and from east (Zhejiang) to west (Tibet). Viral diseases are a threat for sustainable production of watermelon (Xu *et al*., 2004). The major viruses of this crop include *Watermelon mosaic virus* (WMV), *Zucchini yellow mosaic virus* (ZYMV), *Cucumber mosaic virus* (CMV), *Squash mosaic virus* (SqMV), and *Cucumber green mottle mosaic virus* (CGMMV). WMV is a member of the genus *Potyvirus*, one of the largest group of plant viruses comprising more than 200 definite and tentative species (Adams *et al*., 2011). Although WMV is the most common virus found in China (Wang *et al*., 2010), reported from several watermelon-growing regions, little information is available on its genetic characterization based on full-length genomes.

In this study, the complete genomes of isolates WMV-CHN and WMV-Shaanxi were cloned, sequenced and compared with the genomes of other reported isolates, to determine their evolutionary relationship and genetic diversity. To the best of our knowledge, this is the first report on the complete genome sequence of WMV from China.

The WMV-CHN and WMV-Shaanxi isolates were obtained in 2005 and 2011, respectively, from field-grown watermelon plants showing typical virus symptoms in Yangling (Shaanxi, China). Both isolates were transmitted by mechanical inoculation (Sharifi *et al*., 2008) to healthy watermelon plants at the fourth leaf stage, which were kept in an insect-proof greenhouse at 15-25°C. Ten days post inoculation, virus infection was detected by DAS-ELISA (Clark and Adams, 1977) using a commercial kit according to the manufacturer's instructions (Neogen, Beijing, China), and further confirmed by RT-PCR (Wang *et al*., 2010).

Total RNA was extracted from leaf tissue of infected plants using the Trizol reagent (Invitrogen, USA) following the manufacturer's recommendations. First-strand cDNA was then synthesized using MMLV reverse transcriptase (Promega, USA), according to the manufacturer's instructions. For cDNA synthesis, a series of primers extending from 5' untranslated region (UTR) to the 3' UTR of the WMV genome, excluding the Oligo (dT) 18 Primer (Promega, USA), were designed from WMV-Fr sequence (AY437609) (Desbiez and Lecoq, 2004) and other potyvirus sequences (retrieved from GenBank). The complete genomes of WMV-CHN and WMV-Shaanxi were amplified in 10 fragments with the primers listed in Table 1. PCR fragments were generated using *Taq* DNA polymerase (Fermentas, Lithuania) in a programmable thermal cycler (PTC-200, Cepheid, USA) according to the manufacturer's protocols.

The 5'-proximal end of the genomic sequence was determined using the 5'-RACE cDNA amplification kit (TaKaRa, Japan) with the reverse transcription (RT) primer for the 5'-UTR (Table 1), according to the manufacturer's instructions. The 3'-proximal end of the genomic sequence was determined by reverse transcription with the M4-T primer, followed by amplification with the M4 primer and sense primer 3'P1 for 3'-UTR (Chen *et al*., 2002). In all

Corresponding author: Y. Wu
Fax: +86.29.87092716
E-mail: wuyf@nwsuaf.edu.cn

Table. 1. Nucleotide sequences of the primers used in this study

Primer Name	Sequence 5'-3'	Primer position	Length (bp)
5'RT	ATCGGTAGCACTGGTT	536-551	16
5'P1	GATTGTAGCCATTCGTGT	127-144	18
5'P2	GCACCCTACAGTGATGCTCT	269-288	20
F1	ACACGAATGGCTACAAT	127-143	17
R1	TGTACTCCTCCCAACTAA	1661-1678	18
F2	TTTGTGGTGCGAGGTAGA	1383-1401	18
R2	AAGCCTTCCCATCGTCAT	2351-2368	18
F3	CTGAGGGATACAGTAAATACG	2150-2170	21
R3	AGCCATGAATGCAACAA	3911-3927	17
F4	TCCCGTAGCAAAGAGC	3828-3843	16
R4	GCATATCGGCGTTTGA	4750-4765	16
F5	AGGCTCAAACGCCGATAT	4746-4763	18
R5	CTTCACGCCCTACTTTCC	6143-6140	18
F6	GGATGCTGTGGGACTACTT	6041-6059	19
R6	TCAATGGCTCAAAGAACT	7541-7558	18
F7	ATTTGGTGGCTTGTGGGC	7433-7450	18
R7	GCCACTGGTGTCCTTCTTCG	8993-9012	20
F8	ATTTGGATGCAGGAAAGG	8972-8989	18
R8	AGGACAACAAACATTACCGT	10027-10046	20
M4-T	GTTTTCCCAGTCACGAC$(T)_{15}$		32
3'P1	AAGGACACCAGTGGCAAAGG	8998-9017	20
M4	GTTTTCCCAGTCACGAC		17

tests, negative controls were used by replacing DNA template with distilled water. The amplified cDNA fragments were identified by electrophoresis in a 1% agarose gel and visualized by ethidium bromide staining. The specific bands were removed from the agarose gel and purified using TaKaRa agarose gel DNA purification kit (TaKaRa, Japan). The purified PCR fragment was directly cloned into pMD18-T Simple vector (TaKaRa, Japan) according to manufacturer's instructions, and transformed into *Escherichia coli* strain DH5α. DNA sequence analysis was performed on both strands of PCR recombinant clones with an ABI Prism DNA sequencer (model 3730). For each fragment, at least three clones from independent PCR reactions were sequenced in both directions. If there was any difference at any position of the three sequences, at least five more clones were sequenced to identify the base at the position concerned and to obtain the consensus sequence. Nucleotide sequence data were assembled and analyzed using the BLAST program at NCBI website and the Vector NTI program (version 10.0, Invitrogen, USA). The NCBI BLAST program was used to scan the sequence data against other closely related viral gene sequences. Multiple sequence alignment was done using optimal alignment by DNAMAN software package (Lynnon, Biosoft, Canada).

Phylogenetic trees were constructed either by DNAMAN software package using similarity matrix and the neighbor-joining method and max parsimony (MEGA version 5.0) (Tamura *et al*., 2011). Bootstrap analysis with 1000 replicates was performed to evaluate the significance of the internal branches. The validity of the trees was evaluated by the cophenetic correlation coefficient based on the Mantel test (Mohammadi and Prasanna, 2003).

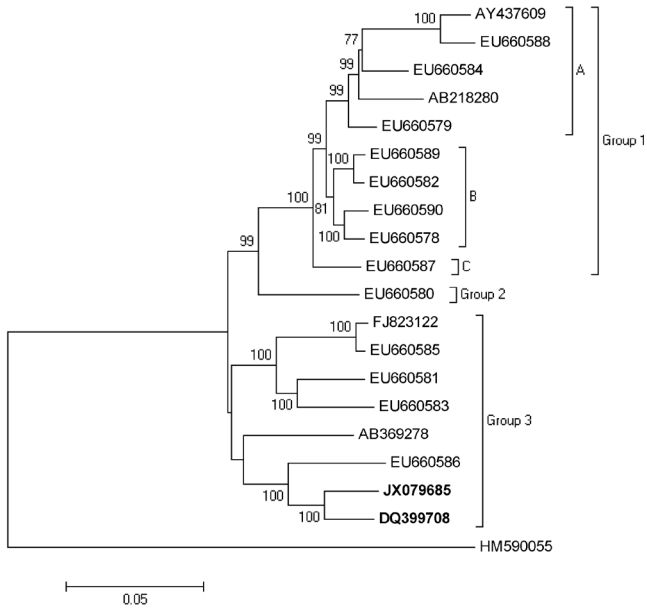

Fig. 1. Unrooted phylogenetic tree constructed with the nucleotide sequences of *Watermelon mosaic virus* from watermelon (JX079685 and DQ399708) and other hosts. The phylogenetic tree was constructed using the neighbor-joining method with 1000 bootstrap replicates. Condensed tree cut-off value of 50%, and bootstrap support were indicated as percentages. The scale bar represents the number of nucleotide replacements per site. GenBank accession numbers of viruses included in the analysis are shown. The sequence of *Soybean mosaic virus* (SMV) (GenBank accession No. HM590055) was used as an outgroup (for more detailed information see Table 2).

The complete genomic sequences of isolates WMV-CHN and WMV-Shaanxi were determined from 10 overlapping cDNA clones and deposited in GenBank under the accession Nos. DQ399708 and JX079685. The genomes of the two isolates were 10,037 and 10,046 nucleotides (nt) in length, respectively, excluding the poly (A) tail. The identity frequencies between two complete genomes were 96.5% at the nt level, and 98.5% at the amino acid (aa) level. Two complete genomic sequences shared 89.7-94.8% identities with other reported WMV isolates at the nt level and 94.6 to 97.1% at the aa level. As with other monopartite viruses of the genus *Potyvirus*, the genomes of isolates WMV-CHN and WMV-Shaanxi encoded a single polyprotein, with an initiation codon at nt position 133 and termination codons at nt positions 9,783 and 9,789, respectively, resulting in putative polyproteins of 3,217 and 3,219 aa. WMV-Shaanxi differed from WMV-CHN by having an insertion of 6 nt after position 8,954, which added two aa residues in the N-terminal domain of the coat protein. The genomes of both isolates contains nine putative cleavage sites, at which the single polyprotein encoded by the coding region is processed into 10 mature proteins by three virus-encoded proteinases P1, HC-Pro and NIa-Pro (Carrington *et al*., 1990). The 5' untranslated regions (5'-UTR) of both isolates were 132 nt in length, whereas the 3'-UTR, whose secondary structure might

Table. 2. Comparison of the nucleotide sequence of WMV-CHN, WMV-Shaanxi and other sequenced WMV isolates from different geographical origins.

GenBank Accession No.	Location of virus isolation (year)	Name	Host	Genome identity (%)		Amino acid identity (%)	
				WMV-CHN	WMV-Shaanxi	WMV-CHN	WMV-Shaanxi
AB218280	Pakistan (2003)	WMV-PK	Melon	90	90.4	95.2	95.4
AY437609	France (1972)	WMV-Fr	Zucchini	92.5	92.8	96.4	96.6
EU660580	Chile (1987)	CHI87-620	Zucchini	90.7	90	95	94.9
EU660581	France (2000)	FMF00-LL1	Zucchini	91.6	90.7	95.2	94.9
EU660586	France (2004)	FBR04-**37**	Zucchini	94.8	93.1	97.1	97
EU660587	France (2006)	C06-188	Cucumber	89.7	90.1	94.9	95.1
EU660588	France (2006)	C06-666	Zucchini	92.2	92.5	96.6	96.7
EU660589	France (2005)	C05-337	Zucchini	89.9	90.3	94.8	95
EU660590	Italy (2000)	ITA00-G	Zucchini	89.7	90.1	94.9	95.1
DQ399708	China(2005)	WMV-CHN	Watermelon	-	96.5	-	98.5
EU660578	France (2000)	FMF00-LL2	Zucchini	89.9	90.6	95.1	95.4
EU660584	Iran (2002)	IR02-54	Zucchini	90.6	90.9	95.5	95.7
EU660**82**	Chile (2002)	[a]CHI02-481	Unknown	89.7	90.2	94.6	94.8
EU660579	Turkey (1991)	[a]TURK91	Unknown	90.5	90.9	95.3	95.5
EU660583	France (2003)	FMF03-141	Zucchini	91.3	90.2	95.1	95
EU660585	France (2005)	C05-270	Melon	91.8	90.8	96.1	96
AB369278	Korea	KOREA	Melon	92.8	92.1	96.6	96.2
FJ823122	Italy (2009)	Lecce	Watermelon	91.7	90.7	96	95.8
JX079685	China (2011)	WMV-Shaanxi	Watermelon	96.5	-	98.5	-
HM590055	China (2004)	SMV 4469-4	Soybean	79.3	79.3	86.3	86.5

be involved in potyviral genome replication (Haldeman-Cahill *et al*., 1998), were AU-rich and differed slightly in length (254 and 257 nt, respectively).

To look for the possible evolutionary relationships of WMV-CHN and WMV-Shaanxi with other known WMV isolates, the nucleotide and amino acid sequences of the complete genome were analyzed using the software of DNAMAN package. WMV-CHN and WMV-Shaanxi sequences shared 94.8% and 93.1% nt identity and 97.1% and 97% aa identity with the French WMV strain FBR04-37 (Desbiez and Lecoq, 2008).

A phylogenetic tree of complete genomic sequences showed that WMV isolates clustered into three groups (Fig. 1). A single isolate from Chile (CHI87-620) was in Group 2, whereas 10 isolates from France, Iran, Pakistan, Turkey, Chile, and Italy were in Group 1, and eight isolates from France, Italy, Korea and China were in Group 3. Members of Group 1 are divided into three subgroups (A, B, and C). Estimates of WMV complete genomic sequence diversity (available in GenBank) indicated that the French isolates are more variable than any of those from other regions of the world. All subgroups of Group 1 and Group 3 comprise French isolates. WMV-CHN and WMV-Shaanxi isolates fall into Group 3 and are most closely related to FBR04-37 and KOREA. These data suggest that FBR04-37 from France may have originated either in China or France or, alternatively, that FBR04-37, WMV-CHN and WMV-Shaanxi have the same origin and were introduced to China and France. Long distance transmission of WMV may take place through movement of infected materials between countries. Since WMV has a wide host range, it might have been imported with non-cucurbit species, orchids or other ornamentals, legumes, etc. [e.g. Laney *et al.* (2012) have recently reported that black locust (*Robinia pseudoacacia*), an ornamental legume tree, is susceptible to WMV], but the route of introduction remains unknown (Lecoq *et al*., 2003). The high percentage of seed transmission suggests that this may be an important avenue for virus dissemination. Comparison of phylogenetic and diversity analysis of WMV isolates demonstrated that viral determinants of symptomatology, host range or resistance breaking maybe located in different regions along the genome, and in many cases involve more than one protein coding region (Desbiez and Lecoq, 2008; Hjulsager *et al*., 2006; Saenz *et al*., 2000). Therefore, to clarify the biology and characteristics of WMV, a better knowledge of the worldwide structure of WMV populations and further analysis of the complete genome sequence are needed.

In this study, two complete genomic sequences of Chinese WMV isolates from watermelon are reported, which, to our knowledge, are the first complete WMV genome sequences obtained in this country. This study helps our understanding of the genetic diversity of WMV isolates infecting watermelon in Shaanxi and can assist in designing better virus control strategies and diagnostic tools, and in understanding virus epidemiology and evolution.

ACKNOWLEDGEMENTS

The study was supported by the National Department Public Benefit Research Foundation (201303028) and the 111 Project from the Education Ministry of China (No. B07049).

REFERENCES

Adams M.J., Zerbini F.M., French R., Rabenstein F., Stenger D.C., Valkonen J.P.T., 2011. Family *Potyviridae*. In: King A.M.Q., Adams M.J., Carstens E.B., Lefkowits E.J. (eds). Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses, pp. 1069-1089. Elsevier-Academic Press, Amsteram, The Netherlands.

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

Chen J., Chen J.P., Adams M.J., 2002. Characterisation of some carla- and potyviruses from bulb crops in China. *Archives of Virology* **147**: 419-428.

Clark M.F., Adams A.N., 1977. Characteristics of the microplate method of enzyme-linked immunosorbenta for the detection of plant viruses. *Journal of General Virology* **34**: 475-483.

Desbiez C., Lecoq H., 2004. The nucleotide sequence of Watermelon mosaic virus (WMV, *Potyvirus*) reveals interspecific recombination between two related potyviruses in the 5' part of the genome. *Archives of Virology* **149**: 1619-1632.

Desbiez C., Lecoq H., 2008. Evidence for multiple intraspecific recombinants in natural populations of Watermelon mosaic virus (WMV, *Potyvirus*). *Archives of Virology* **153**: 1749-1754.

Haldeman-Cahill R., Daros J.A., Carrington J.C., 1998. Secondary structures in the capsid protein coding sequence and 3' nontranslated region involved in amplification of the tobacco etch virus genome. *Journal of Virology* **72**: 4072-4079.

Hjulsager C.K., Olsen B.S., Jensen D.M.K., Cordea M.I., Krath B.N., Johansen I.E., Lund O.S., 2006. Multiple determinants in the coding region of Pea seed-borne mosaic virus P3 are involved in virulence against *sbm-2* resistance. *Virology* **355:** 52-61.

Laney A.G., Avanzato M.V., Tzanetakis I.E., 2012. High incidence of seed transmission of Papaya ringspot virus and Watermelon mosaic virus, two viruses newly identified in *Robinia pseudoacacia*. *European Journal of Plant Pathology* **134**: 227-230.

Lecoq H., Desbiez C., Wipf-Scheibel C., Girard M., 2003. Potential involvement of melon fruit in the long distance dissemination of cucurbit potyviruses. *Plant Disease* **87**: 955-959.

Mohammadi S.A., Prasanna B.M., 2003. Analysis of Genetic Diversity in Crop Plants – Salient Statistical Tools and Considerations. *Crop Science* **43**: 1235-1248.

Saenz P., Cervera M.T., Dallot S., Quiot L., Quiot J.B., Riechmann J.L., Garcıa J.A., 2000. Identification of a pathogenicity determinant of Plum pox virus in the sequence encoding the C-terminal region of protein P3+6K$_1$. *Journal of General Virology* **81**: 557-566.

Sharifi M., Massumi H., Heydarnejad J., Pour A.H., Shaabanian M., Rahimian H., 2008. Analysis of the biological and molecular variability of Watermelon mosaic virus isolates from Iran. *Virus Genes* **37**: 304-313.

Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**: 2731-2739.

Wang W.L., Zhang H., Yu X.Q., Wu Y.F., Zhang W.B., Zhang C.P., 2010. Establishment and appIication of muItipIex RT-PCR for simuItaneous detection of five watermelon viruses ZYMV, WMV, TMV, SqMV and CMV. *Acta Phytopathologica Sinica* **40**: 27-32.

Xu Y., Kang D., Shi Z., Shen H., Wehner T., 2004. Inheritance of resistance to Zucchini yellow mosaic virus and Watermelon mosaic vrus in watermelon. *Journal of Heredity* **95**: 498-502.

Received September 28, 2013
Accepted November 14, 2013