Phytopathologia Mediterranea

The international journal of the Mediterranean Phytopathological Union



Citation: Z. Li, X. Lu, P. Sun, J. Sheng, S. Song (2019) Complete genome sequence of *Alfalfa mosaic virus*, infecting *Mentha haplocalyx* in China. *Phytopathologia Mediterranea* 58(3): 663-669. doi: 10.14601/Phyto-10887

Accepted: July 23, 2019

Published: December 30, 2019

Copyright: © 2019 Z. Li, X. Lu, P. Sun, J. Sheng, S. Song. This is an open access, peer-reviewed article published by Firenze University Press (http://www.fupress.com/pm) and distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Editor: Nihal Buzkan, Kahramanmaraş Sütçü Imam University, Turkey. Short Notes

Complete genome sequence of *Alfalfa mosaic virus*, infecting *Mentha haplocalyx* in China

Zhengnan LI^{1,*}, Xingkai LU^{3,*}, Pingping SUN¹, Jinhua SHENG^{1,#}, Shuang SONG^{2,#}

¹ College of Horticulture and Plant Protection, Inner Mongolia Agricultural University, Hohhot 010018, China

² College of Agriculture, Northeast Agricultural University, Harbin 150030, China

³ Apple Industry Research Institute of Zhaotong, Zhaotong 657000, China

*Zhengnan Li and Xingkai Lu contributed equally to this study

*Corresponding author:

Shuang Song: songshuang.000@163.com

Jinhua Sheng: sjinhua@163.com

Summary. *Mentha haplocalyx (Lamiaceae)* is a herbaceous perennial economic plant which is widely cultivated in China. Plants of *M. haplocalyx* with mosaic symptoms were collected from Zhaotong, Yunnan Province. *Alfalfa mosaic virus* (AMV) was detected from symptomatic leaf samples using small RNA sequencing and RT-PCR. The complete genome sequence of AMV-Mint was determined. AMV-Mint RNA1 was 3,644 nt, RNA2 was 2,594 nt and RNA3 was 2,040 nt, encoding P1 of 1,126 amino acids (aa), P2 of 794 aa, MP of 300 aa and CP of 218 aa. The genome structure of AMV-Mint was similar to those reported previously, except for an insertion of one adenine nucleotide at 2,389 nt of the RNA2, which results in reading frame shift mutation of the P2. Phylogenetic analysis based on RNA3 sequences grouped AMV-Mint into the Group I clade, with closest relationship to AMV-Lst. No recombination event was detected in the genome of AMV-Mint. This is the first report of the complete genome sequence of AMV from *M. haplocalyx*.

Keywords. AMV, mint, small RNA sequencing, insertion mutation, phylogenetic analysis.

INTRODUCTION

Mentha haplocalyx (*Lamiaceae*), commonly known as mint, is a widely cultivated herbaceous perennial plant (Dorman *et al.*, 2003; She *et al.*, 2010). It is an important vegetable crop, and a traditional Chinese medicinal plant with functions including spasmolytic, analgesic, antibacterial and promotion of gas secretion (Dorman *et al.*, 2003; She *et al.*, 2010). In addition, it is one of the most important sources of essentials, which can be used in food, pharmaceutical, flavour and fragrance industries (Zheljazkov *et al.*, 2013). Because cultivation of mint depends mainly on vegetative propagation, infections by viruses are serious threat to mint production. Several viruses have been

reported on cultivated *Mentha* species, including *Toma*to spotted wilt virus (Sether et al., 1991), Tobacco mosaic virus (Samad et al., 2000), Strawberry latent ringspot virus (Postman et al., 2004) and *Tomato leaf curl Paki*stan virus (Samad et al., 2009).

Alfalfa mosaic virus (AMV) belongs to Alfamovirus in the Bromoviridae, having a tripartite positive singlestranded RNA genome. The RNA1 encodes the replicase protein P1 and RNA2 the protein P2 (Bergua *et al.*, 2014). The RNA3 encodes the movement protein (MP) in 5'-proximal half and the coat protein (CP) in 3'-proximal half (Bergua *et al.*, 2014). Based on phylogenetic analyses of the CP gene or RNA3 genome sequences, AMV isolates have been clustered into two groups (Group I and Group II) (Parrella *et al.*, 2011; Song *et al.*, 2019). AMV occurs widely with a broad host range, infecting more than 430 plant species, including many important crops such as potato, tomato, pepper and soybean (Fleysh *et al.*, 2001; Abdalla *et al.*, 2015).

Mentha haplocalyx plants exhibiting obvious viruslike symptoms were recently found in Zhaotong, in the Yunnan Province of China. This paper describes identification of AMV associated with this disease. Furthermore, the complete genome sequence of the AMV *M. haplocalyx* isolate (AMV-Mint) was determined and analyzed.

MATERIALS AND METHODS

Sample collection and RNA extraction

In August 2018, *M. haplocalyx* plants with disease symptoms of mosaic (Figure 1) were collected in Zhaotong, Yunnan Province of China. All of the samples were immediately frozen using liquid nitrogen and stored at -80'. Total RNA was extracted from symptomatic leaf tissues each of approx. around 200 mg, using TRIzol reagent (Invitrogen), according to the manufacturer's instructions. The quality of the extracted RNA was estimated using Bio-Analyzer 2100 (Agilent Technologies) and quantity of RNA was determined using NanoDrop ND-100 (NanoDrop Technologies).

Sequencing of small RNA, and data analysis

The qualified RAN was sent to Biomarker Technologies (Beijing, China) for the construction of a library of small RNAs, as described by Mi *et al.* (2008), followed by sequencing using an IIIumina HiSeq200 platform. The raw data obtained from original image data by base-calling were firstly processed by trimming adapt-



Figure 1. *Mentha haplocalyx* leaf with mosaic symptoms (a), and symptomless leaf (b).

er sequences and removing reads shorter than 18 nt or longer than 35 nt. The obtained clean reads were aligned to the databases Silva (Pruesse *et al.*, 2007), GtRNAdb (Chan *et al.*, 2009), Rfam (Griffiths-Jones *et al.*, 2003) and Repbase (Jurka *et al.*, 2005), using Bowtie software (Langmead *et al.*, 2009) to eliminate non-coding RNA and repeated sequences. The remaining clean reads were assembled using the Velvet program with a minimal overlapping length (*k*-mer) of 17 (Wu *et al.*, 2010). To identify potential viral sequences, the obtained contigs were compared against the nucleotide (nt) sequence databases in NCBI using BLASTn, and nonredundant protein (nr) using BLASTx, with e-value of 10⁻⁵ (Wu *et al.*, 2010).

Validation of candidate viruses

In order to confirm the presence of viruses identified by small RNA sequencing, primers were designed based on the obtained virus contigs (Table S1). Single-stranded cDNA was synthesized from total RNA using random hexamer primers. PCR was carried out in 25 μ L reaction mixtures containing 12.5 μ L Premix *LA Taq* DNA polymerase (TaKaRa), 1.0 μ L 10 μ M primers and 1.0 μ L cDNA. The PCR products were examined by 1% agarose gel electrophoresis.

Determination of full-length genome and sequence analysis

To determine the full-length genome sequence of the AMV *M. haplocalyx* isolate (AMV-Mint), four primers (Table S1) were designed based on the sequences of contigs mapped to AMV genome, and were used in RT-PCR assays. The 5' and 3' terminal sequences of genomic RNA were obtained using the SMARTerTM RACE cDNA Amplification Kit (Clontech) with the primers described previously (Song *et al.*, 2019). The amplicons were cloned into the pMD18-T simple vector (TaKaRa) and sequenced by Sangon Biotech (Shanghai) Co., Ltd. Three independent clones were sequenced for each amplicon.

The full-length genome sequences of RNA1-3 were assembled using Vector NTI (Invitrogen) based on overlapping fragments, and were used for BLAST analysis (https://blast.ncbi.nlm.nih.gov/Blast.cgi). All the genome sequences of AMV available in GenBank database (Table 1) were downloaded and analyzed. The SDT software (Muhire *et al.*, 2014) was used to determine pairwise nucleotide and amino acid sequence similarities.

Phylogenetic and recombination analyses

Multiple sequence alignments were performed by Clustal W algorithm of MEGA 5.0 software (Tamura *et al.*, 2011) with default settings. Phylogenetic neighbourjoining (NJ) and maximum likelihood (ML) trees were constructed based on RNA3 sequences with 1,000 bootstrap replicates, and Tamura three-parameter was used as the best-fitting model of substitution determined by MEGA 5.0.

Recombination analysis was carried out using seven different algorithms of the Recombination Detection Program v.3.44 (RDP3) (Heath *et al.*, 2006) including RDP, GENECONV, CHIMERA, BOOTSCAN, MAX-CHI, SISCAN and 3SEQ. Only recombination events predicted by at least five algorithms with P value <0.01 were accepted.

RESULTS AND DISCUSSION

Analysis of small RNA sequencing data

From small RNA sequencing, a total of 22,918,853 good quality clean reads of 18-35 nt were obtained from 28,510,612 raw reads. RNAs of lengths 21-24 nt were the most abundant. After removing the non-coding RNA and repeated sequences, a total of 9,096,901 clean reads were used for further analyses. Using the Velvet program, we assembled 2,314 contigs from these clean reads with contigs N50 of 63. More than half of these contigs were 30 to 60 nt in length. BLAST analysis against the NCBI database identified 112 contigs mapped to viral genomes, including 46 contigs mapped to AMV, 39 mapped to *Watermelon mosaic virus* (WMV), 19 to *Blackcurrant reversion virus* (BRV), four to *Soybean mosaic virus* (SMV), two to *Cherry leaf roll virus* (CLRV), and two mapped to *Grapevine bulgarian latent* virus (GBLV).

Validation of candidate viruses

To validate the presence of the above-mentioned candidate viruses, a total of eight primer pairs were designed based on the candidate viral contigs (Table S1). However, only AMV was detected by RT-PCR using the primer pair AMV-F1/R1. None of WMV, BRV, SMV, CLRV or GBLV was RT-PCR positive, even when three primer pairs were used for WMV. This strongly indicated that AMV infection was associated with the disease of *M. haplocalyx*.

Next-generation sequencing (NGS) of small RNAs in plant tissues has been widely used for identification of plant viruses, because these virus-derived small RNAs generated by plant RNA-silencing machinery can be sequenced by NGS, assembled *in silico* and used for searching in databases (Seguin *et al.*, 2014; *Liang et al.*, 2015). However, the results in the present study and in previous reports showed that false positive results are common using this method, and further confirmation by PCR or other methods is necessary (Song *et al.*, 2019).

Determination and characterization of the genome sequence of AMV-Mint

The nearly full-length and terminal fragments of the genomic RNAs of AMV-Mint were amplified, respectively, using RT-PCR and RACE technology. The complete genomic sequences were assembled and submitted to the GenBank under accession numbers MK883819, MK88320 and MK883821. The RNA1 of AMV-Mint was 3,644 nt, RNA2 was 2,594 nt, and RNA3 was 2,040 nt, having a similar genome organization to those reported previously (Trucco *et al.*, 2014; Song *et al.*, 2019). The RNA1 encodes P1 of 1,126 amino acids (aa) from 100 to 3,480 nt. The RNA2 encodes P2 of 794 aa from 55 to 2,439 nt. The RNA3 encodes MP of 300 aa from 241 to 1,143 nt, and CP of 218 aa from 1,193 to 1849 nt.

Compared to the other AMV genome sequences available in GenBank, AMV-Mint shared nucleotide sequence similarities of 96.0-98.0% for RNA1, 94.9-97.9% for RNA2, and 94.7-98.5% for RNA3 (Table 1). P1, P2, MP and CP of AMV-Mint shared nucleotide sequence similarities of 96.3-97.9%, 94.9-98.1%, 93.4-98.4% and 94.2-98.3%, respectively, compared with other isolates. Amino acid sequence similarities were 98.1-99.6% for P1, 91.9-97.2% for P2, 93.3-97.7% for MP and 94.5-99.1% for CP (Table 1). An insertion of one adenine nucleotide was

GenBank No.	Isolate	Host	Country	Length (nt)	Nucleotide/Amino acid similarities (%)		
RNA1					Genome	P1	
KC881008	Manfredi	Medicago sativa	Argentina	3,643	98.0	97.9 / 99.4	
MH332897	Gyn	<i>Gynostemma pentaphyllum</i>	China	3,643	97.9	97.8 / 99.5	
HQ316635	HZ	Nicotiana glutinosa	China	3,643	97.8	97.7 / 99.0	
MF990284	175	Solanum tuberosum	Canada	3,631	97.6	97.6 / 99.6	
FN667965	Lst	Lavandula stoechas	Italy	3,543	97.2	97.2 / 98.2	
L00163	425 Leiden	-	-	3,644	97.2	97.3 / 98.5	
FR715040	Tec1	Tecomaria capensis	Spain	3,643	96.5	96.3 / 98.2	
KY810767	FERA160224	Nicotiana tabacum	England	3,643	96.0	96.3 / 98.1	
RNA2					Genome	P2	
FN667966	Lst	Lavandula stoechas	Italy	2,593	97.9	98.1 / 97.2	
KY810768	FERA160224	Nicotiana tabacum	England	2,593	97.7	97.9 / 96.5	
KC881009	Manfredi	Medicago sativa	Argentina	2,593	97.7	97.9 / 96.7	
MH332898	Gyn	Gynostemma pentaphyllum	China	2,598	97.0	97.1 / 96.1	
FR715041	Tec1	Tecomaria capensis	Spain	2,594	96.5	96.5 / 95.8	
X01572	A1M4	-	-	2,593	95.4	95.5 / 91.9	
HQ316636	HZ	Nicotiana glutinosa	China	2,595	94.9	94.9 / 92.9	
RNA3					Genome	MP	СР
FN667967	Lst	Lavandula stoechas	Italy	2,038	98.5	98.4 / 97.3	98.3 / 98.2
KC881010	Manfredi	Medicago sativa	Argentina	2,038	98.0	97.6 / 97.3	98.0 / 97.7
M59241	-	alfalfa	America	2,188	98.0	98.1 / 97.3	98.0 / 97.7
HQ316637	HZ	Nicotiana glutinosa	China	2,041	97.8	97.1 / 97.0	98.2 / 98.2
X00819	S	-	-	2,055	97.6	97.6 / 96.3	97.4 / 97.7
K02703	425 Madison	-	-	2,037	97.5	97.3 / 97.7	97.3 / 97.2
K03542	-	clover	America	2,142	97.5	96.9 / 96.0	97.6 / 96.8
MH332899	Gyn	Gynostemma pentaphyllum	China	2,040	97.5	97.1 / 97.3	97.1 / 97.7
MF990286	175	Solanum tuberosum	Canada	2,041	97.5	97.2 / 97.0	97.0 / 97.7
AF332998	N20	-	Australia	2,257	97.3	96.7 / 96.0	97.6 / 96.3
AB126031	AZ	-	-	2,037	97.1	96.3 / 96.0	97.9 / 99.1
AB126032	Kr	-	-	2,037	96.8	96.9 / 96.0	96.2 / 96.3
FR715042	Tec1	Tecomaria capensis	Spain	2,037	95.5	94.2 / 95.3	94.4 / 94.5
KC767662	178	Actinidia fortunatii	New Zealand	1,986	95.3	94.6 / 94.3	95.0 / 95.0
KC767661	176	Actinidia glaucophylla	New Zealand	1,976	95.2	94.7 / 93.3	96.0 / 95.0
KY810769	FERA160224	Nicotiana tabacum	England	2,039	95.1	94.0 / 95.3	94.5 / 94.5
AF015716	VRU	garden lupin	England	2,038	95.1	93.4 / 95.7	95.1 / 95.0
KC767660	175	Actinidia guilinensis	New Zealand	1,977	94.9	93.9 / 94.3	95.0 / 94.5
AF015717	15/64	garden lupin	England	2,038	94.7	93.4 / 94.0	94.2 / 94.5

Table 1. Nucleotide and amino acid sequence similarities between AMV-Mint and other AMV isolates.

found at 2,389 nt of the AMV-Mint RNA2, resulting in a reading frame shift mutation in the P2 protein from 779 aa to the C-terminal (Figure 2). The AMV-Gyn isolate reported previously (Song *et al.*, 2019) had an insertion of three adenine nucleotides at the same position, resulting in an extra lysine at 779 aa of AMV-Gyn P2 protein. This indicated that the 12 amino acids in the C-terminal region of AMV P2 may be unnecessary for successful AMV infection of host plants.

Phylogenetic and recombination analyses

The RNA3 sequences of AMV-Mint and the other 20 isolates available in GenBank were used to construct the NJ (Figure 3) and ML (Figure S1) phylogenetic trees, and similar topological structures were observed. All of the 20 AMV isolates clustered into two groups. AMV-Mint was located in the Group I, together with 12 of the other available isolates. The closest evolutionary relation-



Figure 2. Alignments of partial RNA3 sequences (2,383-2,439 nt for AMV-Mint) showing an insertion of one adenine nucleotide at 2,389 nt of AMV-Mint RNA2, resulting in a reading frame shift mutation in the C-terminal of the AMV-Mint P2 protein. The inserted nucleotides in RNA2 of AMV-Mint and AMV-Gyn are in red. The codons are underlined, and the corresponding amino acids are indicated in italics below the lines.



Figure 3. Phylogenetic analysis using the neighbour-joining method, based on RNA3 sequences of AMV isolates. Different groups are indicated and separated by dotted lines. AMV-Mint from the present study is indicated in bold font.

ship of AMV-Mint was with AMV-Lst, an isolate from *Lavandula stoechas* from Italy. The other seven isolates in Group II formed two subgroups including IIA (AMV-175, AMV-15/64, AMV-FERA160224 and AMV-VRU) and IIB (AMV-178, AMV-176 and AMV-Tecl),

as reported previously (Parrella *et al.*, 2011; Song *et al.*, 2019). These results showed no phylogenetic correlation to hosts or geographical regions. In addition, no recombination events were detected in the genome of AMV-Mint.

This was the first report of the occurrence, and the complete genome sequence, of AMV from *M. haplocalyx*.

ACKNOWLEDGEMENTS

This research was funded by the project of Research Start-up Funds for High-level Researchers in Inner Mongolia Agricultural University (Grant No. NDYB2018-3), the China Agriculture Research System (CARS-21) and the Natural Science Foundation of Heilongjiang Province, China (LH2019C024).

COMPLIANCE WITH ETHICAL STANDARDS

Funding: This study was financially supported by the project of Research Start-up Funds for High-level Researchers in Inner Mongolia Agricultural University (NDYB2018-3), the China Agriculture Research System (CARS-21) and the Natural Science Foundation of Heilongjiang Province, China (LH2019C024).

Conflict of Interest: The authors declare that they have no conflicts of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent: Informed consent was obtained from all individual participants included in the study.

LITERATURE CITED

- Abdalla O.A., Mohamed S.A., Eraky A.I., Fahmay F.G., 2015. Genetic comparison between coat protein gene of *Alfalfa mosaic virus* isolate infecting potato crop in upper Egypt and worldwide isolates. *International Journal of Virology* 11: 112-122.
- Bergua M.A., Luis-Arteaga M., Escriu F., 2014. Genetic diversity, reassortment, and recombination in *Alfalfa mosaic virus* population in Spain. *Phytopathology* 104: 1241-1250.
- Chan P.P., Lowe T.M., 2009. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Research* 37: 93-97.
- Dorman H.J.D., Kosar M., Kahlos K., Holm Y., Hiltunen R., 2003. Antioxidant properties and composition of aqueous extracts from *Mentha* species, hybrids, varieties and cultivars. *Journal of Agricultural and Food Chemistry* 51: 4563-4569.
- Fleysh N., Deka D., Drath M., Koprowski H., Yusibv V., 2001. Pathogenesis of *Alfalfa mosaic virus* in soybean (*Glycine max*) and expression of chimeric rabies pep-

tide in virus-infected soybean plants. *Phytopathology* 91: 941-947.

- Griffiths-Jones S., Bateman A., Marshall M., Khanna A., Eddy S.R., 2003. Rfam: an RNA family database. *Nucleic Acids Research* 31: 439-441.
- Heath L., van der Walt E., Varsani A., Martin D.P., 2006. Recombination patters in aphthoviruses mirror those found in other picornaviruses. *Journal of Virology* 80: 11827-11832.
- Jurka J., Kapitonov V.V., Pavlicek A., Klonowski P., Kohany O., Walichiewicz J., 2005. Repbase Updata, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research* 110: 462-467.
- Langmead B., Trapnell C., Pop M., Salzberg S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* 10: R25.
- Liang P., Navarro B., Zhang Z., Wang H., Lu M., Li S., 2015. Identification and characterization of a novel geminivirus with monopartite genome infecting apple trees. *Journal of General Virology* 96: 2411-2420.
- Mi S., Cai T., Hu Y., Chen Y., Hodges E., Qi Y., 2008. Sorting of small RNAs into *Arabidopsis argonaute* complexes is directed by the 5' terminal nucleotide. *Cell* 133: 116-127.
- Muhire B.M., Varsani A., Martin D.P., 2014. SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS ONE* 9: e108277.
- Parrella G., Acanfora N., Orilio A.F., Navas-Castillo J., 2011. Complete nucleotide sequence of a Spanish isolate of alfalfa mosaic virus: evidence for additional genetic variability. *Archives of Virology* 145: 2659-2667.
- Postman J.D., Tzanetakis I.E., Martin R.R., 2004. First report of *Strawberry latent ringspot virus* in a *Mentha* sp. from North America. *Plant Disease* 88: 907.
- Pruesse E., Quast C., Knittel K., B Fuchs.M., Ludwig W., Glockner F.O., 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* 35: 7188-7196.
- Samad A., Gupta M.K., Shasany A.K., Ajayakumar P.V., Alam M., 2009. Begomovirus related to *Tomato leaf curl Pakistan virus* newly reported in Mentha crops in India. *Plant Pathology* 58: 404.
- Samad A., Zaim M., Ajayakumar P.V., Garg I.D., 2000. Isolation and characterization of a TMV isolate infecting scotch spearmint (*Mentha gracilis* Sole) in India. *Journal of Plant Disease and Protection* 107: 649-657.
- Seguin J., Rajeswaran R., Malpica-Lopez N., Martin R.R., Kasschau K., Pooggin M.M., 2014. *De novo* recon-

struction of consensus master genomes of plant RNA and DNA viruses from siRNAs. *PLoS One* 9: e88513.

- Sether D.M., De Angelis J.D., Rossignol P.A., 1991. First report of tomato spotted wilt virus in peppermint (*M. piperita*). *Plant Disease* 75: 644.
- She G., Xu C., Liu B., Shi R., 2010. Two new monoterpenes from *Mentha haplocalyx* Briq. *Helvetica Chimica Acta* 93: 2495-2498.
- Song S., Liu H., Zhang J., Pan C., Li Z., 2019. Identification and characterization of complete genome sequence of Alfalfa mosaic virus infecting Gynostemma pentaphyllum. European Journal of Plant Pathology 154: 491-497.
- 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.
- Trucco V., de Breuil S., Bejerman N., Lenardon S., Giolitti F., 2014. Complete nucleotide sequence of *Alfalfa mosaic virus* isolated from alfalfa (*Medicago sativa* L.) in Argentina. *Virus Genes* 48: 562-565.
- Wu Q., Luo Y., Lu R., Lau N., Lai E.C., Palese P., 2010. Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. Proceeding of the National Academy of Sciences of the United States of America 107: 1606-1611.
- Zheljazkov V.D., Cantrell C.L., Astatkie T., Jeliazkova E., 2013. *Mentha cnadensis* L., a subtropical plant, can withstand first few fall frosts when grown in northern climate. *Industrial Crops and Products* 49: 521-525.