

Complete Genome sequence of a new variant of jujube mosaic-associated virus isolated from jujube (*Ziziphus jujuba* Mill.) grown at Aksu in Xinjiang of China

Jianyu Bai (✉ bjyhao2010@sina.com)

Chinese Academy of Forestry Xinjiang Branch <https://orcid.org/0000-0002-9172-0702>

Baojun Liu

Xinjiang Agricultural University

Guoxin Zhang

Xinjiang Agricultural University

Aixing Gu

Xinjiang Agricultural University

Danbo Song

Xinjiang Agricultural University

Quan Wang

Xinjiang Agricultural University

Hong Li

Chinese Academy of Forestry Xinjiang Branch

Research Article

Keywords: jujube, genome sequence, badnavirus, mosaic

Posted Date: April 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-442593/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

This work reports the discovery and the complete genome sequence of a novel member of the genus *Badnavirus* in the family *Caulimoviridae* from a Chinese jujube tree of known variety which grown in Aksu, Xinjiang, China. The symptoms of jujube leaves infected by the virus showed mosaic and malformation, and round chlorotic spots on infected fruits. The genome of this virus is a circular double-stranded DNA, which has the length of 7086 nt and has a genome structure similar to the one reported for jujube mosaic-associated virus (JuMaV), with five open reading frames (ORFs). High nucleotide and amino acid sequence similarity were seen between JuMaV and the new virus, and the nucleotide (NT) difference of the genome was greater than 20% in the RT/ Rnase H coding region of ORF3. Consequently, the virus was identified a new variant of JuMaV, and propose the name as jujube mosaic-associated virus A (JuMaVA).

Main Text

The genus *Badnavirus* is the largest of the eight known Genera of the *Caulimoviridae*, with nearly twice as many species as all other genera[1]. The genome consists of one circular double-stranded DNA of 6.9 to 9.2 kb[2], which typically encodes three open reading frames (ORFs), all of which are located on the on the plus strand[3,4]. Badnavirus ORF1 encodes a small protein of unknown function. ORF2 encodes a protein of possesses a conserved coiled-coil motif as virion-associated protein (VAP)[5]. ORF 3 encodes a large polyprotein which is processed into the movement protein (MP), coat protein (CP), aspartic protease (AP), reverse transcriptase (RT) and ribonuclease H(RNase H)[6]. Moreover, some badnaviruses have more ORFs of unknown function have been reported from other species[7,8]. The RT/ RnaseH coding region of ORF3 is the most conserved region in the genome, and the nucleotide (NT) difference in this part of the genome is greater than 20%, which is used to distinguish the species of this genus[9]. Badnaviruses are spread by vegetative propagation, mealybug vectors, and in some cases by seed[10].

Chinese jujube (*Ziziphus jujuba* Mill.) is one of the oldest cultivated fruit trees in the world. Its use and cultivation history can be traced back to the Neolithic Age 7000 years ago[11,12]. By 2012, the area of jujube in Xinjiang of china was 345,000 hm², of which Aksu had the largest planting area at 136,000 hm²[13]. In recent years, Jujube mosaic disease (JuMD) is caused by virus infection had been reported many times in China. However, pathogenic mechanism of JuMD needs further research and exploration. Two complete sequences of JuMD can be obtained from the database, jujube mosaic-associated virus (JuMaV, KX852476.1) and jujube associated Badnavirus (JuAB, MN274946.1), respectively. Although sequence analysis of JuMaV has been reported, diseased leaves were collected from Chinese jujube trees grown in Beijing, China[14], there is no information on JuMD in other regions and jujube variety.

In July 2020, leaf and fruit samples were collected from other single plant of *Z. jujuba* cv. 'Huizao' (sample IDs: HZ, AKS-6) growing areas in Aksu, Xinjiang Uygur Autonomous Region, China. These plants showed symptoms of mosaic and malformation on young leaves (Fig. 1A, left panel), and round chlorotic spots on infected fruits (Fig. 1A, right panel). All leaf and fruit samples were mixed into one pool,

respectively, from which total RNA was extracted using a TransZol Up Plus RNA Kit (Transgen) for Small RNA and RNA Sequencing analysis.

A cDNA library of sRNAs library was constructed and sequenced on an Illumina HiSeq XTen sequencing machine (Illumina) with a paired-end 150 bp setup[14]. A total of 179 contigs length sequence between 34 to 448 nt were found in leaf and fruit samples, which had high sequence identity ranging from 46.3% to 94.44% with several badnaviruses. Therefore, we concluded that the virus was a badna-like virus found in mosaic-diseased jujube trees.

Total DNA was extracted from obviously symptomatic leaves and fruits using DNasecure Plant Kit. In order to identify the virus, two degenerate primers pair BADNA-FP/RP were used to screen sequence of badnavirus from the total DNA[15], designed to amplify 529 bp sequences in the reverse transcriptase/RNase H region has been used for taxonomic purposes within this genus[1]. The PCR of leaves and fruits samples both yielded an amplicon of approximately 500 bp were analysed by gel electrophoresis in 1.5% (w/v) agarose gels. The amplicons were cloned into the pEASY[®]-T1 Cloning vector, and sequenced. BLAST analysis of the sequencing results revealed that the 528-bp PCR fragment shared maximum sequence identity of 81.8% with the corresponding regions of Jujube mosaic-associated virus (JuMaV, KX852476.1) belong to the genus badnaviruses. Based on the JuMaV alignment, 5 primers were designed spanning the whole genome and used to obtain the full-length genome sequence of this badnavirus (Table 1). PCR amplifications were performed in 25- μ l reaction mixtures containing 1 μ l of DNA template, 13 μ l 2 \times Es Taq MasterMix by the manufacturer (CW BIO), 1 mM each primer, 9 μ l of ddH₂O. PCR cycling was as follows: initial denaturation at 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min, with a final extension at 72°C for 10 min. Amplicons were separated by electrophoresis in 1.4% agarose gels purified by EasyPure[®] Quick Gel Extraction Kit, cloned into the pEASY[®]-T1 Cloning vector and transformation of E. coli DH5a, positive clones were obtained and sequenced by Sangon Biotech (Shanghai) Co., Ltd. In all cases, at least three independent clones in both directions were sequenced to determine the exact sequence of the presumed full-length banavirus genome.

Table 1. Primers used for the amplification of full-length DNAs of JuMaVA from sample HZ and PCR detection for JuMaVA in jujube samples.

Primer	Sequence (5'–3')	Position (nt)	Product Size (bp)	Sequence Identities ^a
F1 ₅₅₂₋₅₇₁	agccgtaaagctcacggaag	552-571	2653	99.29%
R1 ₂₉₈₅₋₃₀₀₄	ggcagttgccattgttcacc	2985-3004		
F2 ₂₉₅₉₋₂₉₇₈	gagccccgtttcaaggagga	2959-2978	2017	100%
R2 ₄₉₁₄₋₄₉₃₃	catggctctgctgacatggt	4914-4933		
F3 ₃₅₄₂₋₃₅₆₁	cccacgaatacacttcgcct	3542-3561	1585	99.66%
R3 ₅₀₂₈₋₅₀₄₇	agcctgtccaaagttcctcg	5028-5047		
F4 ₅₀₂₇₋₅₀₄₆	tcgaggaactttggacaggc	5027-5046	2099	83.64%
R4 ₇₀₁₃₋₇₀₃₂	gcttacacgcaaagcaaacg	7013-7032		
F5 ₆₄₆₇₋₆₄₈₆	tacagaatggcagacgctgg	6467-6486	1663	99.59%
R5 ₇₁₅₋₇₃₄	aggtcttgccaagccaactg	715-734		

a: The sequenced results of primer pairs share highest nucleotide sequence identity with JuMaV (accession KX852476.1)

The edited sequences were used for similarity BLAST searches in the NCBI GenBank databases (<http://www.ncbi.nlm.nih.gov/genbank/>). Multiple Nucleotide sequences obtained by sequencing were assembled and analyzed using Mega Version7.0. Genome maps were generated using SnapGene® Viewer version4.3.6. Full-length genome sequences were assembled in MEGA version7 and ORFs were predicted using NCBI ORFfinder with minimal ORF length of 75 nt (<https://www.ncbi.nlm.nih.gov/orffinder/>). Use the NCBI conserved domain tool to search for conserved domains of putative gene products (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.c>

gi).

The five nucleotide sequences obtained through sequenced were analyzed using similarity BLAST searches in the NCBI GenBank database for sequence identities with other badnaviruses (Table1). After assembly, a circular DNA of 7086 bp was obtained, shared nucleotide sequence identity 99.54% and amino acid sequences identity 94.61% with JuMaV (accession KX852476.1). This virus was possible a new variant of JuMaV, referred to as jujube mosaic-associated virus A (JuMaVA). The full-length sequence of the circular DNA virus has been registered in GenBank with the accession number MW892537.

Genome and conserved domains analysis, genome structures of JuMaVA were representative of typical members of the genus *Badnaviruses* (Fig. 1B). Several highly conserved motifs were found in the

JuMaVA genomes, which had been reported in well-characterized badnavirus genomes[16,17]. The JuMaVA genomes contain a tRNA^{Met} binding site (TGGTATCAGAGC₁₋₁₂) which begins at the 5' end of the viral plus-strand, and occurs in the plant host and serves to initiate viral RNA transcription[18]. There was a TATA box (TATAAATA₇₀₀₂₋₇₀₀₉) and a poly(A) (AATAAAAA₇₀₅₇₋₇₀₆₄) signal in the non-coding intergenic region of upstream of the tRNA^{Met} site. The genome organization contains five ORFs typical for badnaviruses. ORF1, ORF2, ORF4 shared the highest nt (99.77%, 99.51%, 100%) and amino acid (100%, 99.25%, 100%) sequence identity with the JuMaV. Same as that of JuMaV, ORF3 was split into ORF3a, ORF3b by a 70-nt intergenic non-coding region, which is highly variable between the viruses[19]. An uncharacterized superfamily domain approximately 200 residues long in ORF1, referred to as 'domain of unknown function (DUF1319)'[20], be restricted to badnaviruses. ORF2 had no conserved domains were identified. The ORF3b of JuMaVA contained four domains: a zinc-finger domain, a pepsin-like aspartate protease domain, an RT-LTR domain, and an RNase H domain. ORF3a contained the 5'-largest domain was the coding region of JuMaVA movement protein (MP).

Nucleotide and amino acid sequences of the the putative JuMaVA RT and RNase H region were analyzed for sequence identities with other badnaviruses. JuMaVA had the highest sequence identity (73.93%) in nucleotide and 84.41% in amino acid sequences with JuMaV in RT- RNase H region. Also, the RT-RNase H and complete genome sequences shared in the range of 69.62-73.93% and 66.16-74.30% nt sequence identity with well-characterized badnaviruses, 73.26-80.11% and 42.30-72.18% in amino acids.

Phylogenetic trees were generated, which based on full nucleotide sequences of JuMaVA and nucleotide sequences of RT-RNase H region with neighbour-joining (NJ) method using Mega Version7.0 to determine the taxonomic position of JuMaVA. RT/ RNase H sequences were selected for analysis because they play a critical role in viral replication, therefore are subject to more stringent variability constraints. The full-genome phylogenetic tree shows that JuMaV is in a group with JuMaV (Fig. 1C). The phylogenetic tree based on the RT/RNase H nucleotide sequences indicated that JuMaVA clusters with JuMaV (Fig. 1D). The presence of these hallmark features further support the identification of JuMaVA as badnaviruses. Moreover, JuMaVA shares nucleotide sequence identity 99.5% with JuMaV in full sequence and 74.25% in RT- RNase H region. Based on the difference of the genome is greater than 20% in RT- RNase H region, we propose that JuMaVA as a new variant of JuMaV is more accurate.

Conclusion

In conclusion, the first complete genome sequence of jujube mosaic virus was identified isolated from Chinese Jujube trees in Aksu, Xinjiang, China. Not only did we isolate the virus from diseased leaves, but we also isolated the virus from the symptomatic jujube fruit for the first time. Based on the analysis of full-length genome sequences, the virus was identified as the closest relative of JuMaV belonging to a member of the badnavirus species in the family *Caulimoviridae*. We propose this new member of the genus *Badnavirus* to be designated "Jujube mosaic-associated virus A" (JuMaVA), which is a new variant of JuMaV infecting jujube (*Ziziphus jujuba* Mill.) grown at Aksu in Xinjiang of China. The work confirmed that the transmission of the virus is widespread and variable. Therefore, we should pay more attention to

the harm of this virus to jujube trees in daily management. The pathogenic principle and pathogenicity of JuMaVA are the main direction of our research in the future.

Declarations

Funding This work was supported by the Natural Science Foundation of Xinjiang Uygur Autonomous Region (grant number 2017D01A28); “Excellent Postdoctoral Professionals and General Funding Staff of Living Subsidies” of the Department of Human Resources and Social Security of Xinjiang Uygur Autonomous Region.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

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

References

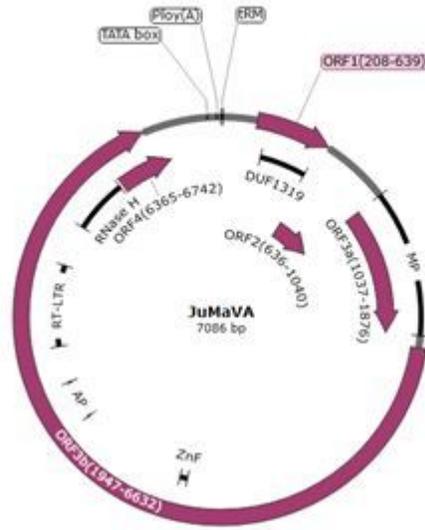
1. King A. M. Q., Adams M. J., Carstens, E.B., Lefkowitz, E.J. (2012) Virus taxonomy: classification and nomenclature of viruses. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, CA, USA
2. Kidanemariam DB, Sukal AC, Abraham AD, et al (2018) Identification and molecular characterization of Taro bacilliform virus and Taro bacilliform CH virus from East Africa. *Plant Pathol* 67:1977–1986. <https://doi.org/10.1111/ppa.12921>
3. Bouhida M, Lockhart BEL, Olszewski NE (1993) An analysis of the complete sequence of a sugarcane bacilliform virus genome infectious to banana and rice. *Journal of General Virology* 74:15–22. <https://doi.org/10.1099/0022-1317-74-1-15>
4. Xu D, Mock R, Kinard G, Li R (2011) Molecular analysis of the complete genomic sequences of four isolates of Gooseberry vein banding associated virus. *Virus Genes* 43:130–137. <https://doi.org/10.1007/s11262-011-0614-8>
5. Stavelone L, Herzog E, Leclerc D, Hohn T (2001) Tetramerization Is a Conserved Feature of the Virion-Associated Protein in Plant Pararetroviruses. *J Virol* 75:7739–7743. <https://doi.org/10.1128/JVI.75.16.7739-7743.2001>
6. Geering A.D. (2014) *Caulimoviridae* (Plant pararetroviruses). John Wiley & Sons, Ltd, Chichester, UK
7. Kazmi SA, Yang Z, Hong N, et al (2015) Characterization by Small RNA Sequencing of Taro Bacilliform CH Virus (TaBCHV), a Novel Badnavirus. *PLoS ONE* 10:e0134147. <https://doi.org/10.1371/journal.pone.0134147>
8. Yang IC, Hafner GJ, Dale JL, Harding RM (2003) Genomic characterisation of taro bacilliform virus. *Archives of Virology* 148:937–949. <https://doi.org/10.1007/s00705-002-0969-1>

9. Geering A, Hull R (2012) Caulimoviridae. In: King AMQ, Adams MJ, Carstens EB, eds. Virus Taxonomy, Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier, New York, NY, USA, pp 429–443
10. Bhat A, Hohn T, Selvarajan R (2016) Badnaviruses: The Current Global Scenario. *Viruses* 8:177. <https://doi.org/10.3390/v8060177>
11. Qu Z. , Wang Y. (1993) Chinese Fruit Trees Record-Chinese Jujube. China Forestry Publishing House, Beijing
12. Liu M. , Wang J. (2016) Fruit scientific research in new China in the past 70 years:Chinese jujube. *J. Fruit. Sci* 36:1369-1381
13. Xinjiang Uygur Autonomous Region Bureau of Statistics (2012) Xinjiang Statistical Yearbook. China Statistics Press, Beijing
14. Guo J, Wang Y, Wang G, et al (2020) Molecular Characteristics of Jujube Yellow Mottle-Associated Virus Infecting Jujube (*Ziziphus jujuba* Mill.) Grown at Aksu in Xinjiang of China. *Viruses* 13:25. <https://doi.org/10.390/v13010025>
15. Yang I C, Hafner G J, Dale J L, Harding R M (2003) Genomic characterisation of taro bacilliform virus. *Archives of Virology* 148:937–949. <https://doi.org/10.1007/s00705-002-0969-1>
16. Bömer M, Rathnayake A I, Visendi P, et al (2018) Complete genome sequence of a new member of the genus Badnavirus, *Dioscorea bacilliform RT virus 3*, reveals the first evidence of recombination in yam badnaviruses. *Arch Virol* 163:533–538. <https://doi.org/10.1007/s00705-017-3605-9>
17. Sukal A, Kidanemariam D, Dale J, et al (2017) Characterization of badnaviruses infecting *Dioscorea* spp. in the Pacific reveals two putative novel species and the first report of *dioscorea bacilliform RT virus 2*. *Virus Research* 238:29–34. <https://doi.org/10.1016/j.virusres.2017.05.027>
18. Medberry S L, Lockhart B E L, Olszewski N E (1990) Properties of *Commelina* yellow mottle virus's complete DNA sequence, genomic discontinuities and transcript suggest that it is a pararetrovirus. *Nucl Acids Res* 18:5505–5513. <https://doi.org/10.1093/nar/18.18.5505>
19. Kreuze J F, Perez A, Gargurevich M G, Cuellar W J (2020) Badnaviruses of Sweet Potato: Symptomless Coinhabitants on a Global Scale. *Front Plant Sci* 11:313. <https://doi.org/10.3389/fpls.2020.00313>
20. Bateman A, Coggill P, Finn R D (2010) DUFs: families in search of function. *Acta Crystallogr F Struct Biol Cryst Commun* 66:1148–1152. <https://doi.org/10.1107/S1744309110001685>

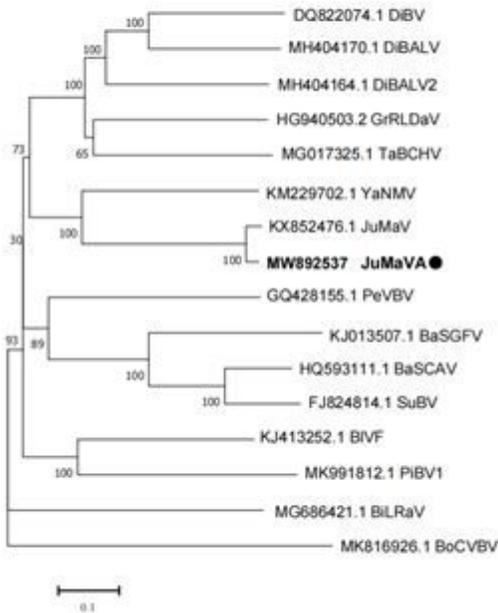
Figures



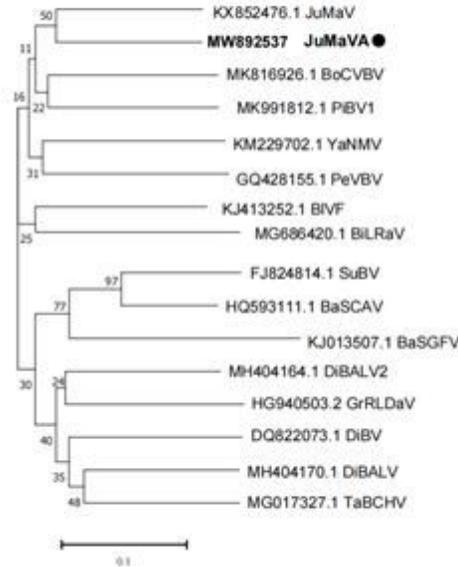
(A)



(B)



(C)



(D)

Figure 1

(A) Diseased leaves (left panel) and symptomatic fruits (right panel) of jujube trees were collected, from which the jujube mosaic-associated virus A was isolated. (B) Schematic representation of the JuMaVA genome. The “tRNA^{Met} binding site” was set as the start site of the JuMaVA genome. ORFs (ORF1, ORF2, ORF3a, ORF3b, ORF4) are the five putative open reading frames (ORFs) of the JuMaVA genome. tRM: tRNA^{Met} binding site, DUF1319: domain of unknown function, ZFD: a zinc-finger domain, AP: a pepsin-like aspartate protease domain, RT-LTR: Reverse transcriptases (RTs) from retrotransposons and retroviruses domain, RNase H: Ribonuclease H-like superfamily domain; Phylogenies were reconstructed using the neighbour-joining (NJ) method with robustness of each internal branch was estimated by 1000 bootstrap replicates, based on nucleotide sequences of the whole genome (C) and the nucleotide

sequences of the RT - RNase H region (D) of Jujube mosaic-associated virus A. The viruses were used for analysis as follow: DiBV, Dioscorea bacilliform virus; DiBALV, Dioscorea bacilliform AL virus; DiBALV2, Dioscorea bacilliform AL virus 2; GrRLDaV, Grapevine roditis leaf discoloration-associated virus; TaBCHV, Taro bacilliform CH virus; YaNMV, Yacon necrotic mottle virus; JuMaV, Jujube mosaic-associated virus isolate Z6; PeVBV, Pelargonium vein banding virus; BaSGFV, Banana streak GF virus; BaSCAV, Banana streak CA virus; SuBV, Sugarcane bacilliform virus; BIVF, Blackberry Virus F; PiBV1, Pitaya badnavirus 1; BiLRaV, Birch leaf roll-associated virus; BoCVBV, Bougainvillea chlorotic vein banding virus

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [jujubemosaicassociatedvirusAcompletegenome.fa](#)