

RESEARCH

Open Access



# Identification of two distinct begomoviruses infecting *Malvastrum coromandelianum*

Yaqin Wang<sup>1</sup>, Xinyue Zhang<sup>1</sup>, Tao Hu<sup>1\*</sup> and Xueping Zhou<sup>1,2\*</sup> 

## Abstract

*Malvastrum coromandelianum* is a common weed plant frequently found around agricultural fields. Three virus isolates (Y249, Y278 and Y281) were obtained from *M. coromandelianum* with yellow vein symptoms in Honghe and Baoshan, Yunnan Province, China. Specific 500 bp products were amplified from total DNA extracts using universal primers for members of the genus *Begomovirus*. The complete viral genome sequences of both Y278 and Y281 were determined to be 2743 nucleotides, and that of Y249 was determined to be 2740 nucleotides. Sequence alignments and phylogenetic analyses support the proposal of creating new species in the genus *Begomovirus*, for which the name malvastrum yellow vein Baoshan virus (MaYVBsV) is proposed for Y278 and Y281, and malvastrum yellow vein Honghe virus (MaYVHhV) is proposed for Y249.

**Keywords:** Geminivirus, *Begomovirus*, *Malvastrum coromandelianum*, Malvastrum yellow vein Baoshan virus, Malvastrum yellow vein Honghe virus

## Background

The genus *Begomovirus* constitutes the largest genus in the family *Geminiviridae*, with over 400 species recognized by the International Committee on Taxonomy of Viruses (ICTV) (Zerbini et al. 2017). Begomoviruses infect economically important crops and cause serious damages in agriculture throughout the tropical and subtropical regions (Brown et al. 2015). Based on genome organizations, begomoviruses can be further classified into monopartite and bipartite subgroups (Hanley-Bowdoin et al. 2013). The genome of bipartite begomoviruses contains two similarly sized single-stranded DNA (ssDNA) components, designated as DNA A and DNA B. The DNA A component encodes six proteins involved in viral replication, encapsidation, transmission and pathogenesis (Fondong 2013). The DNA B component encodes two proteins, which participate in cell-to-cell

and systemic spread throughout the host (Lazarowitz and Beachy 1999). The genome of monopartite begomoviruses comprises a single molecule that is similar to the DNA A component of bipartite begomoviruses. For a few begomoviruses, DNA A/DNA component also encodes V3 (Gong et al. 2021) or C5 (Li et al. 2015) protein involved in suppression of host RNA interference. In addition, two types of ssDNA molecules, referred to as alphasatellite and betasatellite, are frequently found to be associated with the infection of monopartite begomoviruses (Briddon et al. 2018; Yang et al. 2019). Both alphasatellite and betasatellite have approximately half the size of the begomovirus components, and depend on their helper viruses for encapsidation and movement in plants (Zhou 2013). Alphasatellite has a single ORF coding for a replication initiator protein (alpha-Rep), and is capable of self-replication in the host plants (Briddon et al. 2018). Betasatellite encodes a multifunctional  $\beta$ C1 protein involved in pathogenicity, suppression of gene silencing, and repression of other plant defense responses (Li et al. 2018). Recently, a novel  $\beta$ V1 gene was identified

\* Correspondence: [taohu@zju.edu.cn](mailto:taohu@zju.edu.cn); [zzhou@zju.edu.cn](mailto:zzhou@zju.edu.cn)

<sup>1</sup>State Key Laboratory of Rice Biology, Institute of Biotechnology, Zhejiang University, Hangzhou 310058, China  
Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

in betasatellite, which is related with symptom development (Hu et al. 2020).

Weeds serve as the important intermediate hosts of plant viruses, and may participate in disease epidemic. Several types of weeds have been identified to be the alternative hosts of geminiviruses (Yang et al. 2008; Papayiannis et al. 2011; Fiallo-Olive et al. 2012). *Malvastrum coromandelianum* is a weed plant commonly seen in Yunnan Province, China, and is also a reservoir host for geminiviruses (Zhou et al. 2003). In this study, we describe the molecular characterization of two distinct begomovirus species that infect *M. coromandelianum* in Yunnan, China.

## Results

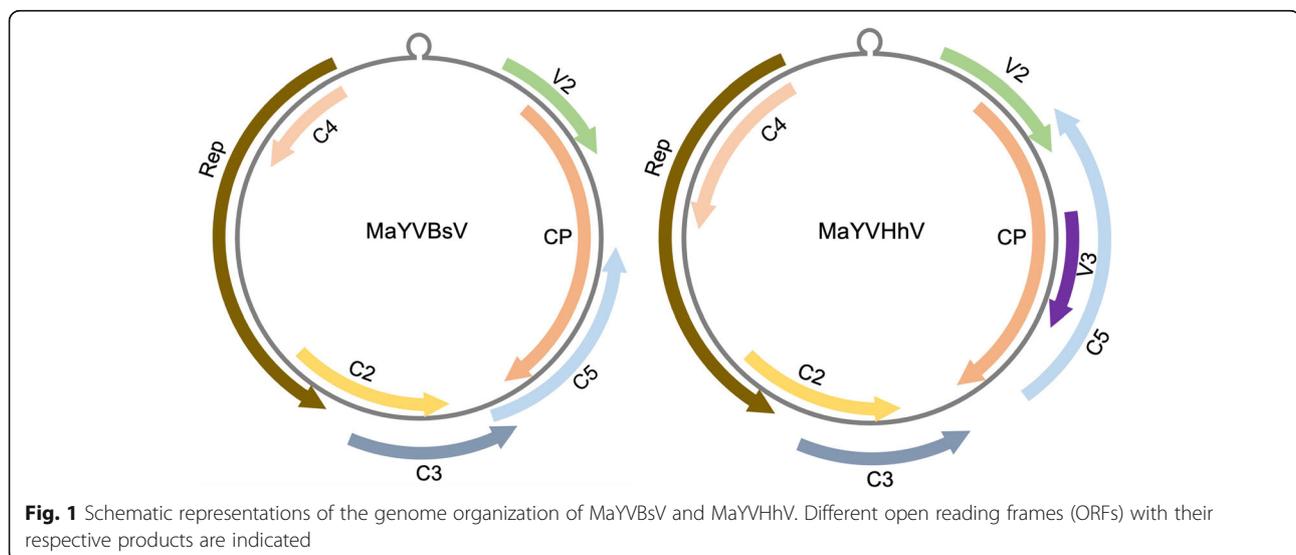
### Identification of begomoviruses in *M. coromandelianum*

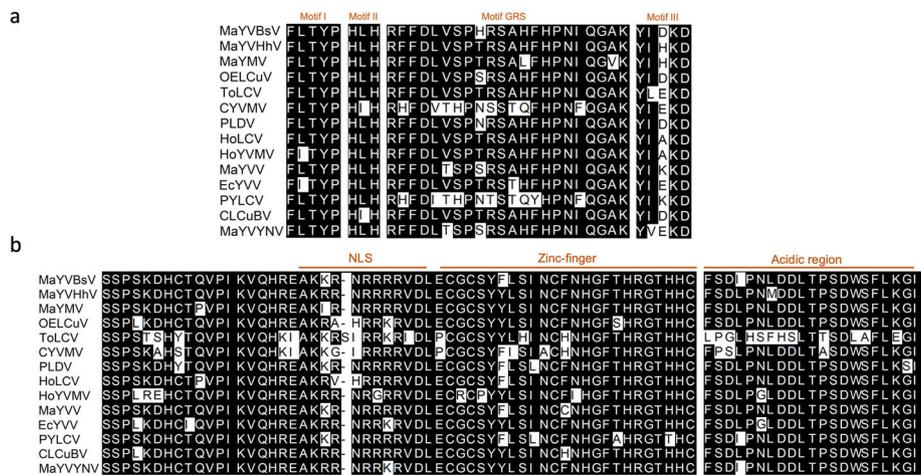
Three *M. coromandelianum* leaf samples, Y249, 278 and Y281, with yellow vein symptoms were collected in Yunnan Province. To test whether the symptoms were caused by geminivirus infection, degenerate primers PA and PB, universal for members of the genus *Begomovirus*, were designed to detect a conserved region within DNA-A/DNA component of begomoviruses, and an amplicon of ~500 base pairs (bp) in size was obtained from each sample, suggesting that these samples were infected by begomoviruses. Based on the obtained sequences, primer pairs were designed to amplify the rest part of the viral genome from each sample. The complete nucleotide sequences of both Y278 and Y281 are 2743 bp in length (Accession numbers FN386459 and FN386460, respectively). These two sequences are nearly identical to each other (99.5% similarity), demonstrating that they are isolates of a single begomovirus species. Comparative sequence analysis of Y278 and Y281 with other begomoviruses showed that they share

the highest sequence identities (88.4 and 88.5%) with pea leaf distortion virus (PLDV). The complete genome of Y249 is 2740 bp in length (Accession number FN552749). Sequence analysis showed that Y249 genome shares the highest sequence similarity (86.5%) with malvastrum yellow mosaic virus (MaYMV). These sequence identities are below the threshold value of 91% for species demarcation within the genus *Begomovirus* (Fauquet and Stanley 2003; Brown et al. 2015). According to the established principles of geminivirus taxonomy and nomenclature (Fauquet et al. 2008), the virus isolates Y278 and Y281 were named as malvastrum yellow vein Baoshan virus (MaYVBsV), and the virus isolate Y249 was named as malvastrum yellow vein Honghe virus (MaYVHhV) (Fig. 1).

### Genome organization of MaYVBsV and MaYVHhV

Both MaYVBsV and MaYVHhV have canonical begomovirus genome arrangement (Fig. 1). In the virion-sense strand of the genome, the ORFs V1 encode 256 amino acids (aa) (MaYVBsV) and 257 aa (MaYVHhV) viral coat proteins (CP) (pfam 00844). Putative proteins encoded by ORFs V2 of MaYVBsV and MaYVHhV are 112 aa (12.9 kDa) and 116 aa (13.3 kDa), respectively, predicted to be the movement protein-like V2 protein (pfam 01524). MaYVHhV has an additional ORF V3 located downstream of the ORF V2, encoding a polypeptide with a molecular weight of 7.4 kDa. The function of the ORF V3 is elusive, while in silico analysis revealed a putative transmembrane helix at position 13–35 aa. In the complementary-sense strand, the C1 ORF of both MaYVBsV and MaYVHhV codes for putative 363 aa-long polypeptides with predicted molecular mass of 41.2 kDa. These proteins were identified as the replication-associated protein (Rep) (pfam 00799), containing four





**Fig. 2** Multiple alignments of protein sequences from MaYVBsV, MaYVHhV and other begomoviruses. **a** An alignment of Rep sequences of MaYVBsV, MaYVHhV and other begomoviruses. The four conserved motifs I (DNA binding), II (metal binding), GRS (Geminivirus Rep Sequence) and III (DNA cleavage and ligation) are indicated. **b** An alignment of C2 protein sequences of MaYVBsV, MaYVHhV and other begomoviruses. The positions of functional regions are indicated

conserved motifs of Reps of geminiviruses (Nash et al. 2011; Fondong 2013) (Fig. 2a). P-loop NTPase domains were identified at aa 217–264 in both proteins, which are involved in nucleotide binding. ORFs C2 of the two viruses encode putative proteins of 150 aa, identified as the transcriptional activator protein (TrAP) (pfam 01440). Three known functional regions of C2 proteins were identified, including nuclear location signal (NLS), cysteine-rich zinc finger-like domain which confers DNA-binding activity, and the acidic region at C-terminal required for transactivation activity (Fig. 2b). Putative proteins encoded by ORFs C3 and C4 of MaYVBsV and MaYVHhV are 134 aa and 143 aa, identified as the replication enhancer protein (REn) (pfam 01407) and C4 protein (pfam 01492), respectively. The

C5 ORFs potentially encode proteins of 167 aa for MaYVBsV and 208 aa for MaYVHhV. By analogy to similarly located ORF of other members of the genus *Begomovirus*, the noncoding regions of both MaYVBsV and MaYVHhV are 272 nt in length, with a predicted hairpin structure containing the conserved nonanucleotide motif TAATATT/AC.

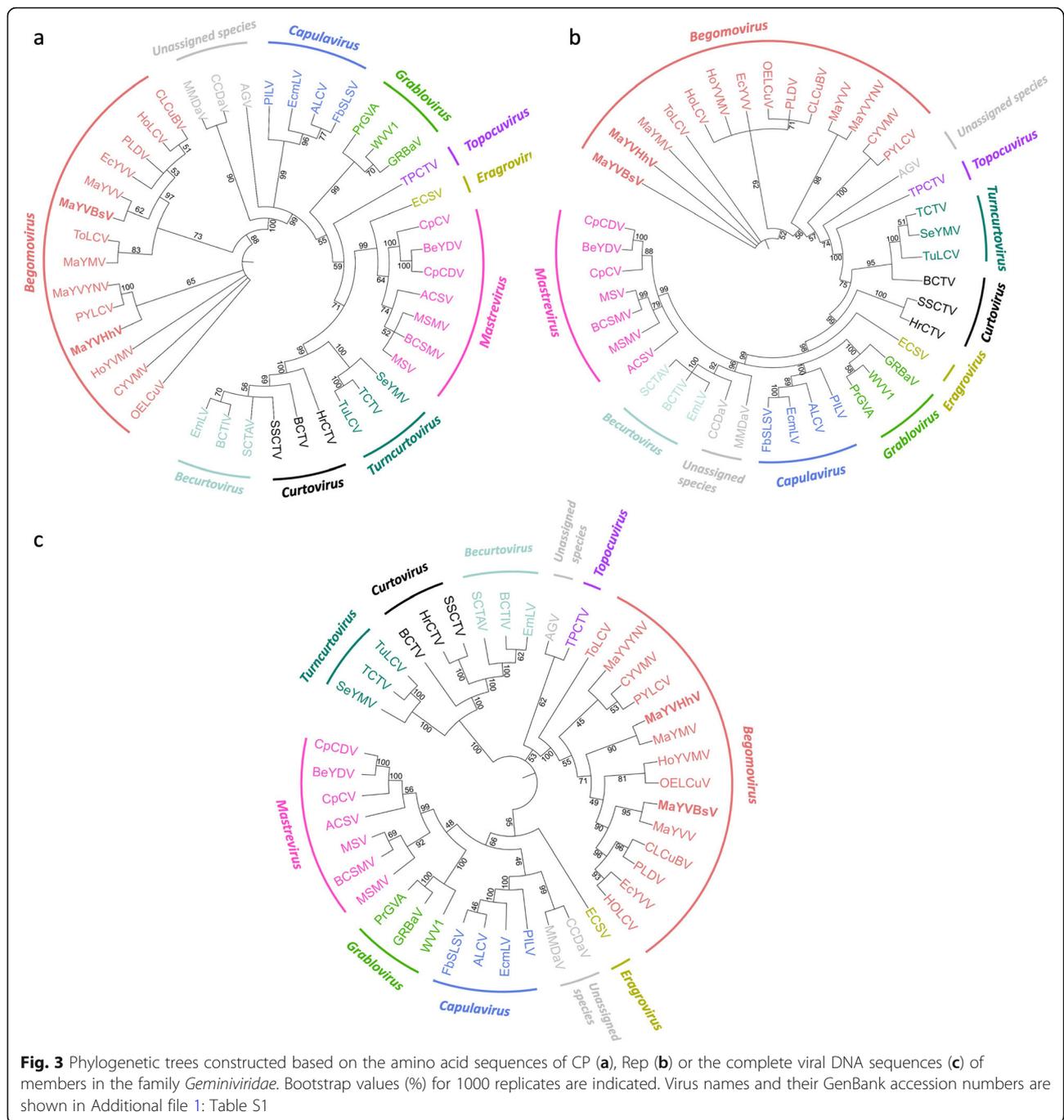
**Phylogenetic relationship of MaYVBsV and MaYVHhV with other geminiviruses**

Amino acid sequence comparisons of the six viral proteins of MaYVBsV or MaYVHhV were performed with the 12 begomoviruses, which have the highest genome sequence identities with these two viruses (Table 1). MaYVBsV has the highest amino acid sequence

**Table 1** Amino acid sequence identities between proteins encoded by MaYVBsV or MaYVHhV and other begomoviruses

Virus	CP	V2	Rep	C2	C3	C4
PLDV	95.3/79.8 <sup>a</sup>	86.1/83.6	91.5/91.1	92.0/91.3	93.3/91.0	80.2/80.4
MaYMV	83.6/83.3	90.2/92.2	92.0/92.6	91.3/94.6	94.0/92.5	77.9/79.0
ToLCV	76.2/73.9	60.7/58.6	88.4/88.6	60.0/60.0	62.7/64.2	67.3/67.6
CYVMV	79.7/91.4	67.0/62.1	74.7/74.7	70.9/67.9	66.4/67.9	44.1/48.8
HoLCV	96.5/81.7	87.5/84.5	90.9/90.4	84.0/85.9	86.6/85.1	70.3/81.4
HoYVMV	74.6/84.4	85.7/80.2	84.9/84.9	82.0/87.3	86.6/83.6	70.7/79.0
MaYVW	96.1/83.7	88.4/90.4	88.2/86.8	92.7/92.6	91.8/92.5	47.9/55.8
EcYVW	81.6/73.9	78.6/75.9	89.8/89.3	89.3/91.3	87.3/88.1	72/79.0
PYLVCV	82.8/94.6	91.7/87.6	79.6/78.2	94.0/89.9	91.0/88.8	44.1/49.4
CLCuBV	93.0/80.9	84.8/81.9	89.4/89.0	84.7/87.9	88.8/86.6	70.3/72.6
OELCuV	78.5/89.5	83.0/80.2	90.0/89.7	86.7/90.6	79.9/79.1	80.8/74.0
MaYVYNV	82.8/94.6	92.0/88.5	85.4/85.1	95.3/93.3	95.5/92.5	51.3/55.0

<sup>a</sup> Amino acid sequence identity compared with MaYVBsV/ MaYVHhV



identities with hollyhock leaf curl virus (HoLCV) for CP (96.5%), malvastrum yellow mosaic virus (MaYMV) for Rep (92.0%), and okra enation leaf curl virus (OELCuV) for C4 (80.8%). The V2, C2 and C3 of MaYVBsV share the highest amino acid identities with those of malvastrum yellow mosaic virus (MaYMV) (92.0, 95.3 and 95.5%, respectively) (Table 1). On the other hand, the CP of MaYVHhV has the highest amino acid sequence identity with those of malvastrum yellow vein Yunnan virus (MaYVYNV) and pepper yellow leaf curl virus

(PYLCV) (94.6%). The V2, Rep, and C2 share the highest amino acid sequence identities with those of MalYMV (92.2, 92.6 and 94.6%, respectively). MaYVHhV C3 has the highest sequence similarity with C3 of malvastrum yellow vein virus (MaYVV), MalYMV and MaYVYNV (92.5%). MaYVHhV C4 protein has the highest amino acid sequence identity with that of HoLCV (81.4%) (Table 1). Phylogenetic analyses were further performed based on the amino acid sequences of two taxonomically relevant gene products, CP and Rep, and the full-length

genome sequences of geminiviruses. The trees showed that MaYVBsV and MaYVHhV always cluster with other begomoviruses (Fig. 3).

## Discussion

Due to the broad distribution and rapid propagation characters, weeds may survive in or around crop fields during the non-cropping season, which makes them important reservoir hosts for plant viruses. Besides, mutations and recombinations increase in virus genome when weeds are infected with multiple virus species, which may increase the transmission rate of viruses and further broaden their host range. In our study, two different types of begomovirus, MaYVBsV and MaYVHhV, were identified in *M. coromandelianum*. Sequence alignment analysis of virus genome and gene products showed that MaYVBsV and MaYVHhV were closely related with begomoviruses that collected from different plant hosts, indicating that these two viruses may be transmitted from *M. coromandelianum* to different types of crops and nurseries. In spite of our efforts, we failed to reveal a DNA-B component associated with MaYVBsV or MaYVHhV, indicating that they are monopartite begomoviruses. Further study will focus on detecting the infectivity and host range of MaYVBsV and MaYVHhV.

## Conclusions

Our study reports the detection and characterization of two novel putative begomovirus species infecting *M. coromandelianum* plants. These results will facilitate the development of strategies for managing the spread of geminiviruses.

## Methods

### Plant materials

Virus isolates were collected from *M. coromandelianum* plant displaying yellow vein symptoms in Honghe (Y249) in 2003 and Baoshan (Y278 and Y281) in 2005, in Yunnan Province, China.

### Determination of the full-length genomic sequences of viruses

Total plant DNA was extracted from leaves of naturally infected symptomatic plants as described (Zhou et al. 2001). Degenerate primer pair PA (5'-TAATATTA CCKGWKGVCCSC-3') and PB (5'-TGGACYTTRCAW GBCCTTCACA-3') was designed to amplify the DNA-A/DNA fragment that contains the partial intergenic region and CP gene. Based on the determined sequences, the primer pairs Y249F (5'-GTAAGTGTCC CTACTGTCCGC-3')/ Y249R (5'-TACACGGGTT GAGTAAGGACTG-3') and Y278F (5'-TCAAAGCTTA AATAATTTCCACCG-3')/ Y278R (5'-TTGAGTGC

GTCATCTGATTGGACCAG-3') were designed to amplify the full-length DNA-A/DNA. The PCR products were cloned into pMD18-T vector (TaKaRa, Tokyo, Japan) and the viral genome sequences were determined by Sanger sequencing.

### Sequence analysis

Sequences were assembled and analyzed with SnapGene<sup>®</sup>. Domains were analyzed using the Pfam database (<http://pfam.xfam.org>). Transmembrane helices were predicted using TMHMM 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). Sequence alignments were performed using and MUSCLE by MEGA X. Phylogenetic analyses were performed using the maximum likelihood method by MEGA X. The GenBank accession numbers of sequences analyzed in the study are listed in Additional file1: Table S1.

### Abbreviations

aa: Amino acid; bp: Base pair; CP: Coat protein; MaYVBsV: Malvastrum yellow vein Baoshan virus; MaYVHhV: Malvastrum yellow vein Honghe virus; NLS: Nuclear location signal; Ren: Replication enhancer protein; Rep: Replication-associated protein; TrAP: Transcriptional activator protein

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s42483-021-00085-y>.

**Additional file 1: Table S1.** GenBank accession numbers of the sequences used in the study.

### Acknowledgements

Not applicable.

### Authors' contributions

XZ designed the research; YW, XZ and TH performed the experiments; YW and TH analyzed the data; YW, TH and XZ wrote the paper. All authors read and approved the final manuscript.

### Funding

This work was supported by a grant from the National Natural Science Foundation of China to Yaqin Wang (31972235).

### Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>State Key Laboratory of Rice Biology, Institute of Biotechnology, Zhejiang University, Hangzhou 310058, China. <sup>2</sup>State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China.

Received: 4 February 2021 Accepted: 20 March 2021

Published online: 02 April 2021

## References

- Briddon RW, Martin DP, Roumagnac P, Navas-Castillo J, Fiallo-Olivé E, Moriones E, et al. *Alphasatellitidae*: a new family with two subfamilies for the classification of geminivirus- and nanovirus-associated alphasatellites. *Arch Virol*. 2018;163:2587–600.
- Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JCF, et al. Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. *Arch Virol*. 2015;160:1593–619.
- Fauquet CM, Briddon RW, Brown JK, Moriones E, Stanley J, Zerbini M, et al. Geminivirus strain demarcation and nomenclature. *Arch Virol*. 2008;153:783–821.
- Fauquet CM, Stanley J. Geminivirus classification and nomenclature: progress and problems. *Ann Appl Biol*. 2003;142:165–89.
- Fiallo-Olive E, Navas-Castillo J, Moriones E, Martínez-Zubiaur Y. Begomoviruses infecting weeds in Cuba: increased host range and a novel virus infecting *Sida rhombifolia*. *Arch Virol*. 2012;157:141–6.
- Fondong VN. Geminivirus protein structure and function. *Mol Plant Pathol*. 2013;14:635–49.
- Gong P, Tan H, Zhao SW, Li H, Liu H, Ma Y, et al. Geminiviral genomes encode additional proteins with specific subcellular localizations and virulence function. *bioRxiv*. 2021; doi: <https://doi.org/10.1101/2021.03.01.433473v1>.
- Hanley-Bowdoin L, Bejarano ER, Robertson D, Mansoor S. Geminiviruses: masters at redirecting and reprogramming plant processes. *Nat Rev Microbiol*. 2013;11:777–88.
- Hu T, Song Y, Wang Y, Zhou X. Functional analysis of a novel  $\beta V1$  gene identified in a geminivirus betasatellite. *Sci China Life Sci*. 2020;63:688–96.
- Lazarowitz SG, Beachy RN. Viral movement proteins as probes for intracellular and intercellular trafficking in plants. *Plant Cell*. 1999;11:535–48.
- Li F, Xu X, Huang C, Gu Z, Cao L, Hu T, et al. The AC5 protein encoded by mungbean yellow mosaic India virus is a pathogenicity determinant that suppresses RNA silencing-based antiviral defenses. *New Phytol*. 2015;208:555–69.
- Li F, Yang X, Bisaro DM, Zhou X. The  $\beta C1$  protein of geminivirus-betasatellite complexes: a target and repressor of host defenses. *Mol Plant*. 2018;11:1424–6.
- Nash TE, Dallas MB, Reyes MI, Buhrman GK, Ascencio-Ibañez JT, Hanley-Bowdoin L. Functional analysis of a novel motif conserved across geminivirus replication proteins. *J Virol*. 2011;85:1182–92.
- Papayiannis LC, Katis NI, Idris AM, Brown JK. Identification of weed hosts of tomato yellow leaf curl virus in Cyprus. *Plant Dis*. 2011;95:120–5.
- Yang CX, Cui GJ, Zhang J, Weng XF, Xie LH, Wu ZJ. Molecular characterization of a distinct *Begomovirus* species isolated from *Emilia sonchifolia*. *J Plant Pathol*. 2008;90:475–8.
- Yang X, Guo W, Li F, Sunter G, Zhou X. Geminivirus-associated betasatellites: exploiting chinks in the antiviral arsenal of plants. *Trends Plant Sci*. 2019;24:519–29.
- Zerbini FM, Briddon RW, Idris A, Martin DP, Moriones E, Navas-Castillo J, et al. ICTV virus taxonomy profile: *Geminiviridae*. *J Gen Virol*. 2017;98:131–3.
- Zhou XP. Advances in understanding begomovirus satellites. *Annu Rev Phytopathol*. 2013;51:357–81.
- Zhou XP, Xie Y, Peng Y, Zhang ZK. Malvastrum yellow vein virus, a new *Begomovirus* species associated with satellite DNA molecule. *Chin Sci Bull*. 2003;48:2205–9.
- Zhou XP, Xie Y, Zhang ZK. Molecular characterization of a distinct begomovirus infecting tobacco in Yunnan, China. *Arch Virol*. 2001;146:1599–606.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

