Complete Genome Sequence of *Amazon Lily Mosaic Virus* Isolated From Amaryllis (*Hippeastrum Hybridum* Hort.)

Chian-Chi Lin
Asia University College of Medical and Health Science
Tsung-Chi Chen (kikichenwolf@hotmail.com)
Asia University College of Medical and Health Science https://orcid.org/0000-0002-4422-5469

Research Article

Keywords: Potyvirus, Amazon lily mosaic virus, amaryllis, Hippeastrum hybridum Hort

Posted Date: February 8th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1316498/v1

Abstract

In April 2011, a virus was isolated by single lesion isolation on *Chenopodium quinoa* leaves from an amaryllis plant with chlorotic ringspots at a private garden in Changhua County, Taiwan. The Illumina MiSeq sequencing system was used to determine the genomic nucleotide (nt) sequence of the virus. A *de novo* assembled contig with 9377 nt, containing a putative open reading frame of the potyviral polyprotein, was annotated as the potyvirus Amazon lily mosaic virus (ALiMV) by sharing 95.5% nt identity with a partial genomic sequence of ALiMV available in GenBank. Therefore, the amaryllis virus was designated as ALiMV-TW. Through 5´ and 3´ rapid amplification of cDNA ends (RACEs), the complete genome sequence of ALiMV-TW was clarified as 9618 nt. Sequence analysis indicated that the genome and polyprotein of ALiMV-TW share 52.3%-65.1% nt identity and 30.1%-64.2% aa identity with those of other potyviruses, respectively. This is the first report to elucidate the complete genome sequence of *Amazon lily mosaic virus*.

Full Text

Amaryllis (*Hippeastrum hybridum* Hort.) is a perennial bulbous plant of Amaryllidaceae and has attracted worldwide attention due to its charming and colorful flowers. The Netherlands, South Africa, Japan, Brazil and the United States are the world's major exporters [1]. Amaryllis is a very popular garden flower in Taiwan. Virus-like symptoms can often be observed on amaryllis plants. The orthotospovirus capsicum chlorosis virus, the cucumovirus cucumber mosaic virus, the potyvirus hippeastrum mosaic virus (HiMV), and the carlavirus nerine latent virus (NeLV) infecting amaryllis have been reported in Taiwan, and the mixed infection of HiMV and NeLV is common in the field [2–4].

In April 2011, amaryllis plants showing chlorotic ringspots were found at a private garden in Changhua County, Taiwan (Fig. 1a). A virus isolate, denoted AV-N1, was isolated from a diseased amaryllis plant by three successive single lesion transfers on *Chenopodium quinoa* leaves. AV-N1 induced local chlorotic spots on *C. quinoa* leaves 7 days post-inoculation (dpi) (Fig. 1b) and caused systemic mosaic and curling symptoms on *Nicotiana benthamiana* plants 10 dpi (Fig. 1c). The virus was maintained in *C. quinoa* as a local-lesion host and *N. benthamiana* as a systemic host by mechanical transmission under temperature-controlled conditions (25-28°C) in the greenhouse for study.

To determine the genomic sequence of AV-N1, 100 mg of AV-N1-infected *C. quinoa* leaf tissue 7 dpi was used to purify total RNA using the Plant Total RNA Miniprep Purification kit (GMbiolab, Taichung, Taiwan) following the manufacturer's instructions. The total RNA was used as a template to prepare the random primer-primed cDNA library, following the Illumina TruSeq® RNA Sample Preparation v2 Guide as previously described [5]. Sequencing was conducted at Genetech Biotech Co., Ltd. (Taipei, Taiwan) using an Illumina MiSeq sequencing platform to synthesize paired-end reads. After the trimming of adaptors and low-quality sequences, the raw reads were randomly clipped into 23-mers using CLC Genomics Workbench 6.0.2 (CLC bio, Aarhus, Denmark) to facilitate *de novo* sequence assembly. The contigs were then subjected to the Basic Local Alignment Search Tool (BLAST) searches against the non-redundant

protein database of the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov). A contig with 9377 nt, containing a putative open reading frame (ORF) of the potyviral polyprotein, was annotated as Amazon lily mosaic virus (ALiMV), a *Potyvirus* species, by sharing 95.5% nt identity with a partial genomic sequence of ALiMV (acc. no. AB158523) available in the GenBank of NCBI. Based on sequence homology, AV-N1 was identified as an amaryllis isolate of ALiMV and renamed ALiMV-TW.

Furthermore, both 5′ and 3′ rapid amplification of cDNA ends (RACEs) were performed as the previously described methods [6] to exlucidate the complete genome of ALiMV-TW. Specific primers designed from the nt sequence of the ALiMV contig were used in RACEs. In 5′ RACE, the primer ALiMV-P1-1269R (5′-CCGCTATCCCCTTTCCGTAAAGTA-3′) was used to synthesize the first cDNA strand that was then tailed with PolyC(3c3t11c) (5′-CCCTTTCCCCCCCCCCC-3′) [6]. Subsequently, the primers ALiMV-P1-1137R (5′-GTATGCGTGCTGCTTCTCGATTG-3′) and PolyG(11g3a3g) (5′-GGGGGGGGGGGAAAGGG-3′) [6] were used to amplify the 5′ end of the ALiMV genome. The primer ALiMV-Nlb-8642F (5′-AACACTGGTACTTCCGGGACGCA-3′) paired with Oligo-dT (5′-TTTTTTTTTTTTTTTTTTTTTTTTB-3′) were used in 3′ RACE. Revere transcription (RT) was conducted at 50°C for 30 min, and then inactivation at 94°C for 2 min. Polymerase chain reaction (PCR) was run with the conditions of hot start at 94°C for 2 min, and then 30 cycles of 94°C for 1 min, 60°C for 2 min and 72°C for 3 min, and a final extension of 72°C for 7 min. All amplicons were cloned by the TOPO TA Cloning kit (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) using standard procedures, and then sequenced by Mission Biotech company (Taipei, Taiwan) using an ABI 3730 DNA sequencer. Three clones for each fragment were sequenced. The obtained nt sequences of the 5′- and 3′-ends were assembled with the contig sequence to obtain the complete genome sequence of ALiMV-TW, deposited in the GenBank database (acc. no. OL355031).

Analysis of the genome structure indicated that the ALiMV-TW genome is 9618 nt long and contains a single ORF, ranging from nt 126 to nt 9359, encoding a polyprotein of 3077 amino acids (aa) (350.8 kDa). According to the consensus sequences for P1, HC-Pro and Nla protease cleavage sites [7], the positions of P1 (nt 126-992), HC-Pro (nt 993-2360), P3 (nt 2361-3467), 6K1 (nt 3468-3623), CIP (nt 3624-5525), 6K2 (nt 5526-5681), VPg (nt 5682-6245), Nla (nt 6246-6983), Nlb (nt 6984-8540) and CP (nt 8541-9356) were deduced (Fig. 2a).

The complete genome sequences of 136 potyviruses available in the GenBank database (Table S1) were used for sequence comparison with ALiMV-TW. Multiple sequence alignments were performed using the ClustalW method in the MEGA X software (http://www.megasoftware.net) [8]. Sequence Demarcation Tool Version 1.2 (SDTv1.2) was used to perform pairwise alignments and generate color-coded pairwise matrices and nt and aa identity calculations [9]. Phylogenetic analyses were conducted in the MEGA X software using the Neighbor-Joining method with 1000 bootstrap replicates, and the maximum composite likelihood method was used to compute the evolutionary distances. The results revealed that the full-length genome and polyprotein of ALiMV share 52.3%-65.1% nt identity and 30.1%-64.2% aa

identity, respectively, with those of other potyviruses (Fig. 2b & c). And, ALiMV is close to the potato virus Y (PYV) clade (Fig. 2d & e).

Using the One-Step RT-PCR kit (GMbiolab), a one-step RT-PCR method with the primer pair ALiMV-P1-967F (5´-CAATAAGACACCTAGTGCCCGAAAGC-3´)/ALiMV-P1-1269R was developed to specifically detect ALiMV in plants. The amplification conditions were set as RT at 50°C for 30 min, inactivation at 94°C for 2 min, and then PCR at 35 cycles of 94°C for 30 s, 58°C for 30 s and 72°C for 1 min, and a final extension of 72°C for 7 min. An expected 303-bp DNA fragment was amplified and its nt sequence was directly sequenced by Mission Biotech company using an ABI 3730 DNA sequencer to verify the correctness.

ALiMV was first discovered in Amazon lily (*Eucharis grandiflora*) in Japan. Before this report, only part of the genome sequence was available in the GenBank database [10–11]. Here, we have identified the amaryllis virus as ALiMV and elucidated the whole genome sequence of the virus. Although an ALiMV isolate (acc. no. AY590143) was previously reported in Taiwan [12], that virus should be re-identified as HiMV due to the over 90% nt identity. Therefore, this is also the first report of ALiMV in Taiwan.

Declarations

Author contributions

Funding acquisition: Tsung-Chi Chen.

Supervision: Tsung-Chi Chen.

Investigation: Chian-Chi Lin, Tsung-Chi Chen.

Data analysis: Chian-Chi Lin, Tsung-Chi Chen.

Manuscript writing: Chian-Chi Lin, Tsung-Chi Chen.

Compliance with Ethical Standards

Funding

This study was funded by the Ministry of Science and Technology, Taiwan, R.O.C. (grant number MOST 110-2313-B-468-001).

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

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

References

1. Wang Y, Chen D, He X, Shen J, Xiong M, Wang X, Zhou D, Wei Z (2018) Revealing the complex genetic structure of cultivated amaryllis (*Hippeastrum hybridum*) using transcriptome-derived microsatellite markers. Sci Rep 13;8:10645
2. Chen CC, Huang CH, Cheng YH, Chen TC, Yeh SD, Chang CA (2009) First report of *Capsicum chlorosis virus* infecting amaryllis and blood lily in Taiwan. Plant Dis 93:1346
3. Chen CC, Chiang FL (2015) Identification and detection of *Cucumber mosaic virus* isolated from bloody lily (*Haemanthus multiflorus*). J Taiwan Agric Res 64:74–79
4. Chen CC, Chang CA, Chiang FL (2016) Serological and molecular identification of *Nerine latent virus* infecting blood lily (*Haemanthus multiflorus* Martyn). J Taiwan Agric Res 65:194–206
5. Chou WC, Lin SS, Yeh SD, Li SL, Peng YC, Fan YH, Chen TC (2017) Characterization of the genome of a phylogenetically distinct tospovirus and its interactions with the local lesion-induced host *Chenopodium quinoa* by whole-transcriptome analyses. PLoS ONE 12:e0182425
6. Li JT, Yeh YC, Yeh SD, Raja JAJ, Rajagopalan PA, Liu LY, Chen TC (2011) Complete genomic sequence of watermelon bud necrosis virus. Arch Virol 156:359–362
7. Goh CJ, Hahn Y (2021) Analysis of proteolytic processing sites in potyvirus polyproteins revealed differential amino acid preferences of NIa-Pro protease in each of seven cleavage sites. PLoS ONE 16:e0245853
8. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549
9. Muhire BM, Varsani A, Martin DP (2014) SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. PLoS ONE 9:e108277
10. Terami F, Honda Y, Fukumoto F (1995) Amazon lily mosaic virus, a new potyvirus infecting Amazon lily (*Eucharis grandiflora*). Ann Phytopathol Soc Jpn 61:1–6
11. Fuji S, Terami F, Furuya H, Naito H, Fukumoto F (2004) Nucleotide sequence of the coat protein genes of *Alstroemeria mosaic virus* and Amazon lily mosaic virus, a tentative species of genus *Potyvirus*. Arch Virol 149:1843–1849
12. Wu WC, Chang YC (2003) A new mosaic disease of Amazon lily in Taiwan. New Dis Rep 8:29

Figures

Figure 1

Symptoms associated with Amazon lily mosaic virus (ALiMV) infection. **(a)** Chlorotic ringspots were observed on amaryllis leaves. **(b)** Chlorotic spots were induced on a *Chenopodium quinoa* leaf 7 days

post-inoculation (dpi) with ALiMV. **(c)** Systemic mosaic and curling symptoms on a *Nicotiana benthamiana* plant were caused by ALiMV infection 10 dpi.

Figure 2

Genome sequence analysis of Amazon lily mosaic virus (ALiMV). **(a)** Genome structure of ALiMV. The nucleotide positions corresponding to individual open reading frames are illustrated. **(b and c)** Color-coded pairwise matrices of the genome and polyprotein of potyviruses. Pairwise Sequence Demarcation Tool Version 1.2 (SDTv1.2) was used to perform pairwise alignments and to generate color-coded matrices. Comparisons of ALiMV with other potyviruses are highlighted with red boxes. **(d and e)** Phylogenetic trees of potyviruses are drawn based on the nucleotide sequence of the complete genome and the amino acid sequence of the polyprotein. Phylogenetic analyses were conducted in the MEGA X software using the Neighbor-Joining method with 1000 bootstrap replicates, and the maximum composite likelihood method was used to compute the evolutionary distances. The virus names and the accession numbers of genome sequences are shown in Table S1. ALiMV is indicated by arrows. Artichoke latent virus (ArLV) belonging to the genus *Macluravirus* is used as an outgroup.

Supplementary Files

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

- Submission2515699.txt
- TableS1.Accnosofpotyviruses.docx