Arch Virol (2016) 161:1079–1082
DOI 10.1007/s00705-015-2745-z

ANNOTATED SEQUENCE RECORD

Complete genome sequence of a novel zantedeschia mild mosaic virus isolate: the first report from Australia and from *Alocasia* sp.

Dawit B. Kidanemariam[1,2] · Adane D. Abraham[2] · Amit C. Sukal[1] · Timothy A. Holton[3] · James L. Dale[1] · Anthony P. James[1] · Robert M. Harding[1]

Received: 16 September 2015 / Accepted: 22 December 2015 / Published online: 7 January 2016

Abstract The complete genome of an Australian isolate of zantedeschia mild mosaic virus (ZaMMV) causing mosaic symptoms on *Alocasia* sp. (designated ZaMMV-AU) was cloned and sequenced. The genome comprises 9942 nucleotides (excluding the poly-A tail) and encodes a polyprotein of 3167 amino acids. The sequence is most closely related to a previously reported ZaMMV isolate from Taiwan (ZaMMV-TW), with 82 and 86 % identity at the nucleotide and amino acid level, respectively. Unlike the amino acid sequence of ZaMMV-TW, however, ZaMMV-AU does not contain a polyglutamine stretch at the N-terminus of the coat-protein-coding region upstream of the DAG motif. This is the first report of ZaMMV from Australia and from *Alocasia* sp.

Zantedeschia mild mosaic virus (ZaMMV) is a positive-sense, single-stranded RNA virus belonging to the genus *Potyvirus*, family *Potyviridae* [1]. The virus was first reported infecting calla lily (*Zantedeschia* spp.) in Taiwan in 2005 [1, 2] and has subsequently only been reported from Italy [3] and New Zealand (GenBank accession no. DQ407934). Currently, there is only a single published full-length genome sequence of ZaMMV available from Taiwan, designated ZaMMV-TW (GenBank accession no. AY626825).

In 2014, an aroid (*Alocasia* sp.) showing feathery mosaic symptoms typical of those caused by the potyvirus dasheen mosaic virus (DsMV) was observed at Bellthorpe, Queensland, Australia. To determine if the plant was infected with DsMV, symptomatic leaves were collected and initially tested for the presence of potyviruses by RT-PCR. Total RNA was extracted using a lithium-chloride-based protocol [4], and cDNA was synthesised using M-MLV reverse transcriptase (Promega) and oligo(dT)18 primers. PCR was carried out using GoTaq® Green Master Mix (Promega) and degenerate primers designed to amplify a fragment of the CI-coding region of potyviruses [5, 6]. As a positive control, total RNA extracted from DsMV-infected taro leaves was used. An amplicon of the expected size (~700 bp) was generated from extracts derived from both the DsMV-infected taro and *Alocasia* sp. samples. The amplicon from the *Alocasia* sp. sample was subsequently cloned and sequenced, and a BLAST search analysis of the 621-nt sequence revealed 84 % and 93 % identity to ZaMMV-TW at the nucleotide and amino acid level, respectively. As ZaMMV has not previously been reported in Australia, or in *Alocasia* sp., the complete genome sequence of this novel isolate (herein referred to as ZaMMV-AU) was determined.

To obtain the remainder of the virus genome, RT-PCR was carried out using degenerate primers targeting the potyviral HC-Pro-, NIb- and CP-coding regions [6, 7]. The amplicons were cloned and sequenced, and specific primers were subsequently designed in order to amplify the intervening sequences. The 5′-terminal sequence of the genome was obtained by rapid amplification of cDNA ends (RACE) using a 5′/3′ RACE Kit, 2nd Generation (Roche). In all

✉ Robert M. Harding
r.harding@qut.edu.au

[1] Centre for Tropical Crops and Biocommodities, Queensland University of Technology (QUT), Brisbane 4001, Australia

[2] Holetta National Agricultural Biotechnology Research Center, Ethiopian Institute of Agricultural Research, P.O. Box 2003, Addis Ababa, Ethiopia

[3] Biosciences for Eastern and Central Africa, International Livestock Research Institute (BecA-ILRI Hub), P.O. Box 30709, Nairobi, Kenya

cases, amplicons were separated by electrophoresis through 1.5 % agarose gels, purified using the Freeze 'N Squeeze™ DNA Gel Extraction Spin Columns (Bio-Rad) and cloned into pGEM®-T Easy Vector (Promega) following the manufacturer's protocols. For each amplicon, at least three clones were sequenced in both directions using a Big Dye® Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific) following the manufacturer's protocol. Sequencing data were processed and analysed using CLC Main Workbench v6.9.2 (QIAGEN) and Vector NTI Advance® Suite v11 (Invitrogen). Virus sequences were further aligned and analyzed using the ClustalW multiple alignment algorithm in BioEdit version 7.1.9 (http://www.mbio.ncsu.edu/BioEdit/bioedit.html), and phylogenetic trees were constructed from ClustalW-aligned sequences using MEGA version 6.0.6 [8], using the neighbour-joining method and the Kimura 2-parameter model with 1000 bootstrap replications.

The complete genome sequence of ZaMMV-AU was assembled from the consensus sequences of amplicons generated using degenerate and specific primers and 5′ RACE. The genome comprised 9942 nucleotides (GenBank accession no. KT729506) including the 5′ UTR (198 nt) and 3′ UTR (240 nt), but excluding the 3′ polyA-tail. Sequence analysis identified a single putative open reading frame of 9501 nt, encoding a 3167-amino-acid polyprotein with a predicted MW of 359.14 kDa. Sequence comparison of the complete genome of ZaMMV-AU to ZaMMV-TW revealed 82 % identity, while comparison of the polyprotein coding region revealed 79.5 % and 86.3 % identity at the nucleotide and amino acid level, respectively. The nucleotide and amino acid sequences of the putative protein-coding and non-coding region of ZaMMV-AU and ZaMMV-TW were also compared (Table 1). These analyses revealed nucleotide sequence identities ranging from 61.3 % (5' UTR) to 88 % (3' UTR) and amino acid sequence identities ranging from 58.7 % (P1) to 100 % (6K1). Further, when the nucleotide sequence of ZaMMV-AU was compared to the partial sequences of the Italian and New Zealand ZaMMV isolates, there was 86.6 % and 80.3 % identity, respectively. Phylogenetic analysis of the complete genome sequence of ZaMMV-AU and other selected *Potyviridae* members showed that it groups with ZaMMV-TW within the bean common mosaic virus (BCMV) subgroup of the genus *Potyvirus* (Fig. 1a).

Analysis of the amino acid sequence revealed the presence of putative potyviral proteinase cleavage sites, which would result in cleavage of the polyprotein into ten putative mature proteins [9–11] (Fig. 1b). A PIPO-encoding ORF (81 amino acids), embedded within the P3 cistron, was also identified, while the presence of a DAG motif in the CP-coding region indicates that ZaMMV-AU may be aphid-transmissible. The amino acid sequence of ZaMMV-TW contains an unusual stretch of 39 glutamine residues at the N-terminus of the CP-coding region, upstream of the DAG motif, for which the function is unknown [1]. Despite analyzing this region in sequences of 10 individual clones from two different cloning experiments, such a polyglutamine stretch is not present in the amino acid sequence of ZaMMV-AU. In ZaMMV-AU, this region comprises a smaller number of amino acids and is lysine rich (9/36) (Fig. 1c). The differences between ZaMMV-TW and –AU across this region raise questions about their biological significance.

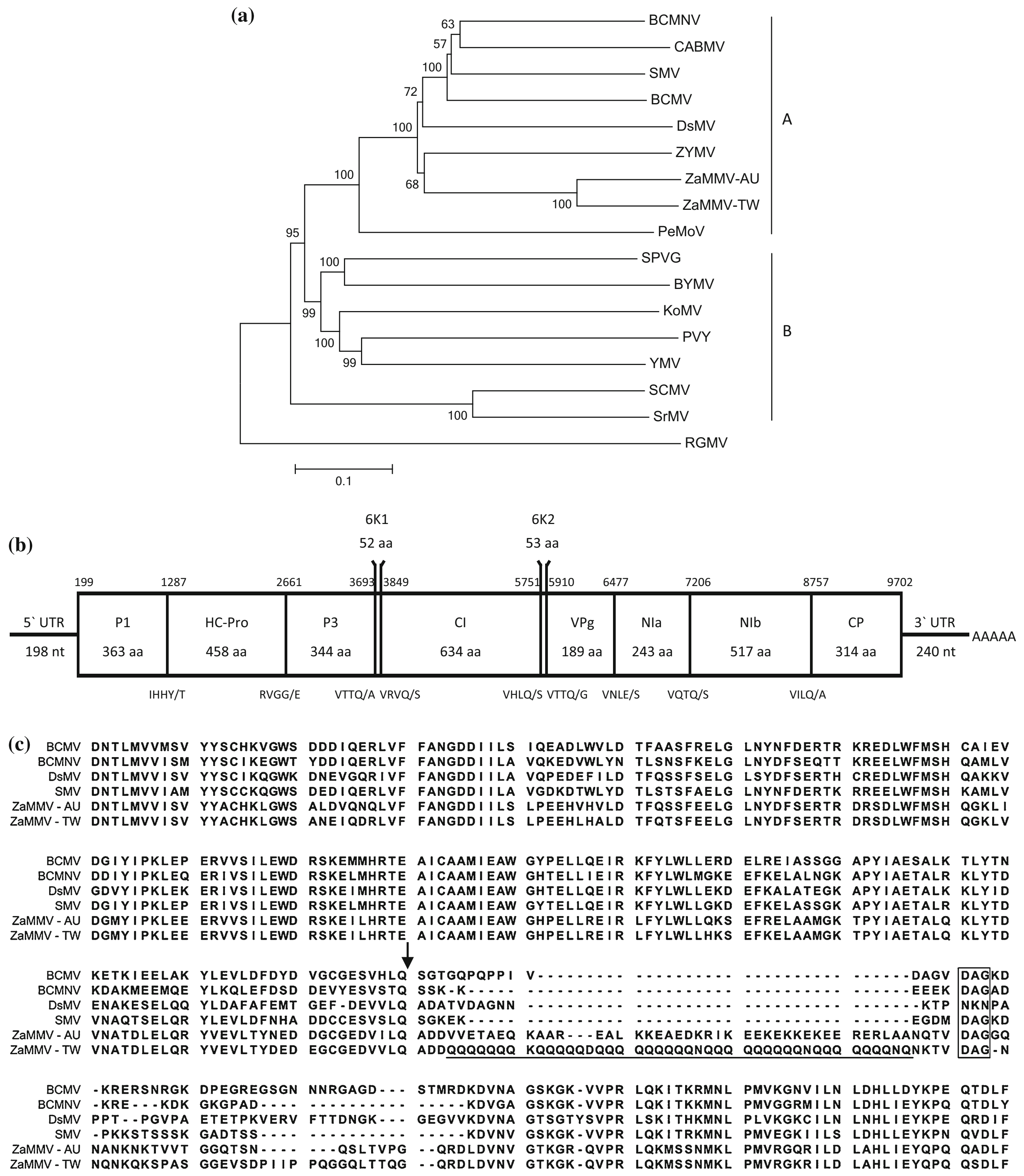
According to the current species demarcation criteria for viruses within the family *Potyviridae* [9], members of different species are distinguished by having less than 80 % CP amino acid sequence identity and less than 76 % nucleotide sequence identity, either in the CP-coding region or over the whole genome. Based on comparisons over the whole genome, the virus sequence isolated from *Alocasia* sp. in this study should be considered a strain of ZaMMV. However, based on comparisons using only the CP-coding region, the reported sequence could be considered a new potyvirus. We have chosen the whole-genome comparison as the criterion for classification due to the presence of the unusual stretch of amino acids in the CP-coding region upstream of the DAG motif. When this region was excluded from comparisons, the amino acid sequences of ZaMMV-TW and ZaMMV-AU shared 89.8 % identity.

To our knowledge, this is the first report of ZaMMV from Australia, and it is also the first report of ZaMMV infecting an *Alocasia* sp. This report provides a useful reference for further work investigating the occurrence of viruses in *Alocasia* sp. and its relatives, particularly the

Table 1 Comparison of the nucleotide and amino acid sequences of the putative coding and non-coding regions of ZaMMV-AU and ZaMMV-TW

	5′ UTR	P1	HC-Pro	P3	PIPO*	6K1	CI	6K2	VPg	NIa	NIb	CP	3′ UTR
% Nucleotide sequence identity	61.3	63.6	80.6	81.4	85.8	85.9	83.2	81.1	84.1	82.7	83.6	75.5	88
% Amino acid sequence identity		58.7	90	83.4	78.8	100	93.4	92.5	93.7	92.2	93.4	78.1	

* Predicted from ZaMMV-TW sequence annotation



```
     BCMV DNTLMVVMSV YYSCHKVGWS DDDIQERLVF FANGDDIILS IQEADLWVLD TFAASFRELG LNYNFDERTR KREDLWFMSH CAIEV
    BCMNV DNTLMVVISM YYSCIKEGWT YDDIQERLVF FANGDDIILA VQKEDVWLYN TLSNSFKELG LNYDFSEQTT KREELWFMSH QAMLV
     DsMV DNTLMVVISV YYSCIKQGWK DNEVGQRIVF FANGDDIILA VQPEDEFILD TFQSSFSELG LSYDFSERTH CREDLWFMSH QAKKV
      SMV DNTLMVVIAM YYSCCKQGWS DEDIQERLVF FANGDDIILA VGDKDTWLYD TLSTSFAELG LNYNFDERTK RREELWFMSH KAMLV
ZaMMV-AU  DNTLMVVISV YYACHKLGWS ALDVQNQLVF FANGDDIILS LPEEHVHVLD TFQSSFEELG LNYDFSERTR DRSDLWFMSH QGKLI
ZaMMV-TW  DNTLMVVISV YYACHKLGWS ANEIQDRLVF FANGDDIILS LPEEHLHALD TFQTSFEELG LNYDFSERTR DRSDLWFMSH QGKLV

     BCMV DGIYIPKLEP ERVVSILEWD RSKEMMHRTE AICAAMIEAW GYPELLQEIR KFYLWLLERD ELREIASSGG APYIAESALK TLYTN
    BCMNV DDIYIPKLEQ ERIVSILEWD RSKELMHRTE AICAAMIEAW GHTELLIEIR KFYLWLMGKE EFKELALNGK APYIAETALR KLYTD
     DsMV GDVYIPKLEK ERIVSILEWD RSKEIMHRTE AICAAMIEAW GHTELLQEIR KFYLWLLEKD EFKALATEGK APYIAETALK KLYID
      SMV DGIYIPKLEP ERIVSILEWD RSKELMHRTE AICAAMIEAW GYTELLQEIR KFYLWLLGKD EFKELASSGK APYIAETALR KLYTD
ZaMMV-AU  DGMYIPKLEE ERVVSILEWD RSKEILHRTE AICAAMIEAW GHPELLREIR LFYLWLLQKS EFRELAAMGK TPYIAETALQ KLYTD
ZaMMV-TW  DGMYIPKLEE ERVVSILEWD RSKEILHRTE AICAAMIEAW GHPELLREIR LFYLWLLHKS EFKELAAMGK TPYIAETALQ KLYTD

     BCMV KETKIEELAK YLEVLDFDYD VGCGESVHLQ SGTGQPQPPI V--------- ---------- ---------- ------DAGV DAGKD
    BCMNV KDAKMEEMQE YLKQLEFDSD DEVYESVSTQ SSK-K----- ---------- ---------- ---------- ------EEEK DAGAD
     DsMV ENAKESELQQ YLDAFAFEMT GEF-DEVVLQ ADATVDAGNN ---------- ---------- ---------- -------KTP NKNPA
      SMV VNAQTSELQR YLEVLDFNHA DDCCESVSLQ SGKEK----- ---------- ---------- ---------- ------EGDM DAGKD
ZaMMV-AU  VNATDLELQR YVEVLTYNED DGCGEDVILQ ADDVVETAEQ KAAR---EAL KKEAEDKRIK EEKEKKEKEE RERLAANQTV DAGGQ
ZaMMV-TW  VNATDLELQR YVEVLTYDED EGCGEDVVLQ ADDQQQQQQQ KQQQQDQQQ QQQQQQNQQQ QQQQQQNQQQ QQQQNQNKTV DAG-N

     BCMV -KRERSNRGK DPEGREGSGN NNRGAGD--- STMRDKDVNA GSKGK-VVPR LQKITKRMNL PMVKGNVILN LDHLLDYKPE QTDLF
    BCMNV -KRE---KDK GKGPAD---- ---------- -----KDVGA GSKGK-VVPR LQKITKKMNL PMVGGRMILN LDHLIEYKPQ QTDLY
     DsMV PPT--PGVPA ETETPKVERV FTTDNGK--- GEGVVKDVNA GTSGTYSVPR LSKITHKMNL PLVKGKCILN LNHLIEYKPE QRDIF
      SMV -PKKSTSSSK GADTSS---- ---------- -----KDVNV GSKGK-VVPR LQKITRKMNL PMVEGKIILS LDHLLEYKPN QVDLF
ZaMMV-AU  NANKNKTVVT GGQTSN---- ---QSLTVPG --QRDLDVNV GTKGR-QVPR LQKMSSNMKL PMVRGQRILD LAHLIEYQPQ QADLF
ZaMMV-TW  NQNKQKSPAS GGEVSDPIIP PQGGQLTTQG --QRDLDVNV GTKGK-QVPR LQKMSSNMKL PMVRGKRILD LAHLIEYQPQ QSDLF
```

◀**Fig. 1** Phylogenetic and sequence analysis of ZaMMV-AU. a) Phylogenetic tree generated by the neighbour-joining method in MEGA 6 [9] using nucleotide sequences of the complete polyprotein ORF of selected potyviruses comprising the bean common mosaic virus (BCMV) subgroup and representative members of other genus *Potyvirus* subgroups. The tree was rooted using ryegrass mosaic virus (RGMV, NC_001814.1), the type member of the genus *Rymovirus*. Bootstrap values greater than 50 % are shown, and the scale bar indicates 0.1 substitutions per site. Subgroup A includes potyviruses from the BCMV subgroup, and subgroup B includes potyviruses from other subgroups. Abbreviations are BCMV (bean common mosaic virus, KC832501), BCMNV (bean common mosaic necrosis virus, AY864314), BYMV (bean yellow mosaic virus, AB439732), CABMV (cowpea aphid-borne mosaic virus, AF348210), DsMV (dasheen mosaic virus, KJ786965), KoMV (konjac mosaic virus, AB219545), PVY (potato virus Y, EF026076), SCMV (sugarcane mosaic virus, AY569692), SMV (soybean mosaic virus, KF135488), SPVG (sweet potato virus G, KF790759), SrMV (sorghum mosaic virus, KJ541740) WMV (watermelon mosaic virus, FJ823122),YMV (yam mosaic virus, NC004752), ZaMMV-AU (zantedeschia mild mosaic virus-Australia, KT729506), ZaMMV-TW (zantedeschia mild mosaic virus-Taiwan, AY626825), ZYMV (zucchini yellow mosaic virus, AY188994-1). b) Genome organisation, predicted mature proteins and their relative position on the genome, and predicted proteinase cleavage sites of ZaMMV-AU (PIPO-encoding ORF not shown). c) Alignment of partial amino acid sequences of the NIb-CP junction of ZaMMV and selected potyviruses from the BCMV subgroup using CLC Main Workbench v6.9.2 (QIAGEN). The polyglutamine amino acid tract present in the ZaMMV-TW isolate is underlined, the characteristic DAG motif is boxed, and the predicted cleavage site between NIb and CP-coding regions is indicated by an arrow

economically important members of the family *Araceae*, such as the cultivated taros (*Colocasia esculenta*).

Acknowledgments DK is the recipient of an Australia Awards Scholarship.

References

1. Huang CH, Chang YC (2005) Identification and molecular characterization of Zantedeschia mild mosaic virus, a new calla lily-infecting potyvirus. Arch Virol 150:1221–1230
2. Huang CH, Hu WC, Yang TC, Chang YC (2007) Zantedeschia mild mosaic virus, a new widespread virus in calla lily, detected by ELISA, dot-blot hybridization and IC-RT-PCR. Plant Pathol 56:183–189
3. Rizzo D, Panattoni A, Stefani L, Paoli M, Nesi B, Lazzereschi S, Vanarelli S, Farina P, Della Bartola M, Materazzi A, Luvisi A (2015) First report of Zantedeschia mild mosaic virus on *Zantedeschia aethiopica* (L) Spreng in Italy. J Plant Pathol 97:1–2
4. Valderrama-Cháirez ML, Cruz-Hernández A, Paredes-López O (2002) Isolation of functional RNA from cactus fruit. Plant Mol Biol Rep 20:279–286
5. Ha C, Revill P, Harding RM, Vu M, Dale JL (2008) Identification and sequence analysis of potyviruses infecting crops in Vietnam. Arch Virol 153:45–60
6. Ha C, Coombs S, Revill P, Harding RM, Vu M, Dale JL (2008) Design and application of two novel degenerate primer pairs for the detection and complete genomic characterization of potyviruses. Arch Virol 153:25–36
7. Yamamoto H, Fuji S (2008) Rapid determination of the nucleotide sequences of potyviral coat protein genes using semi-nested RT-PCR with universal primers. J Gen Plant Pathol 74:97–100
8. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729
9. Adams MJ, Zerbini FM, French R, Rabenstein F, Stenger DC, Valkonen JPT (2012) Family Potyviridae. In: King AMQ, Lefkowitz E, Adams MJ, Carstens EB (eds) Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses, London, pp 1069–1089
10. Adams MJ, Antoniw JF, Fauquet CM (2005) Molecular criteria for genus and species discrimination within the family Potyviridae. Arch Virol 150:459–479
11. Adams MJ, Antoniw JF, Beaudoin F (2005) Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. Mol Plant Pathol 6:471–487