Arch Virol (2008) 153:2145–2148
DOI 10.1007/s00705-008-0235-2

BRIEF REPORT

Nucleotide sequence and genome organization of a member of a new and distinct *Caulimovirus* species from dahlia

H. R. Pappu · K. L. Druffel · R. Miglino · A. R. van Schadewijk

Received: 5 January 2008 / Accepted: 2 October 2008 / Published online: 31 October 2008

Abstract A distinct caulimovirus, associated with dahlia mosaic, was cloned and sequenced. The caulimovirus, tentatively designated as dahlia common mosaic virus (DCMV), had a double-stranded DNA genome of ca. 8 kb. The genome organization of DCMV was found to be typical of members of the genus *Caulimovirus* and consisted of six major open reading frames (ORFs), ORFs I–VI, and one minor ORF, ORF VII. Sequence comparisons with the DNA genomes of two known caulimoviruses isolated from dahlia, *Dahlia mosaic virus* (DMV) and an endogenous caulimovirus, DMV-D10, showed that DCMV is a member of a distinct caulimovirus species, with sequence identities among various ORFs ranging from 25 to 80%.

Dahlia mosaic disease [1, 5, 6] is becoming an increasingly important disease of dahlia (*Dahlia variabilis*). Symptoms of dahlia mosaic include mosaic, chlorotic vein banding, systemic chlorosis, and leaf malformation. *Dahlia mosaic virus* (DMV), a distinct species in the genus *Caulimovirus* (family *Caulimoviridae*) has been described [6] and the physical map of the viral genome has been reported [21]. The genomic sequence of an isolate of this species was subsequently determined and is available in GenBank [3]. In addition to DMV, we recently characterized another caulimovirus that was found to be widely prevalent in dahlia in the US and elsewhere [13–17, 19, 20]. Tentatively designated as DMV-D10, this caulimovirus was found to be significantly divergent from DMV at nucleotide and amino acid sequence levels [14, 17]. Moreover, DMV-D10 was found to exist as an endogenous plant pararetorviral sequence (EPRV) in *D. variabilis* [18]. In addition to these two caulimoviruses, a partial genomic sequence representing the viral coat protein and reverse transcriptase genes of another caulimovirus has been recently reported [12] in dahlia samples originally collected in The Netherlands. As part of our ongoing studies on assessing the molecular variability of caulimoviruses associated with dahlia mosaic disease, we obtained the complete nucleotide sequence of a caulimovirus that was found in samples collected in the US. This virus had the highest sequence identity with the partial sequence of the virus originally collected in The Netherlands [12]. We report the molecular characterization and inter-viral relationships of this third and a distinct caulimovirus, tentatively designated as dahlia common mosaic virus (DCMV), that is associated with dahlia mosaic disease.

A minipurification method was used to obtain a partially purified virus preparation. The source of the virus was dahlia plants collected from local dahlia gardens between 2005 and 2007. Dahlia leaf tissue (20 g) was ground in liquid nitrogen, followed by the addition of 80 ml 0.5 M phosphate buffer, pH 7.2, containing 0.75% sodium sulfite. The slurry was squeezed through cheesecloth. To the expressed sap, 4.8 g of urea (1 M) and 2 ml (2.5%) of Triton X-100 were added, and the solution was stirred at room temperature for 1 h. The solution was centrifuged at

Sequences reported here are available in GenBank under the following accessions: EU090952, EU090953, EU090954, EU090955, EU090956, EU090957.

H. R. Pappu (✉) · K. L. Druffel
Department of Plant Pathology, Washington State University, PO Box 646430, Pullman, WA 99164-6430, USA
e-mail: hrp@wsu.edu

R. Miglino · A. R. van Schadewijk
Dutch Flower Bulb Inspection Service, PO Box 300, 2160 AH Lisse, The Netherlands

42,000 rpm in a Ti70 rotor in a Beckman ultracentrifuge for 90 min. The supernatant was discarded and the pellet was resuspended in 2.5 ml of 0.5 M phosphate buffer, pH 7.2. The liquefied pellet was distributed as 1-ml fractions into 1.7-ml microfuge tubes, and to each fraction, 200 µl chloroform and 50 µl of *n*-butanol were added and mixed by inversion for 1 min. The tubes were centrifuged at 7,000*g* for 8 min. The aqueous layer was collected and was subjected to another cycle of centrifugation at 32,000 rpm in a Ti70.1 rotor for 1 h. The supernatant was discarded and the pellet was resuspended in distilled water and stored at −80°C.

Initial clones of DCMV were obtained by amplifying a portion of the viral genome by polymerase chain reaction (PCR). Primers were designed by using Primer Designer version 2.0 (Scientific and Educational Software, Cary, NC). The forward primer, 5′-AAA TTT ATG TTT ACT GTW T-3′, was based on the conserved sequences from the movement gene in both DMV-D10 [14, 17] and DMV (available in GenBank, submitted by I.B. Maiti). The reverse primer (5′-TTC TAT TGG AAT GTA GTC TT-3′) was based on a partial sequence reported for a DMV isolate from The Netherlands [12]. A complementary primer was designed based on the clone (6396-196C: 5′-TTA GAC TGA TTT TTG TAA GC-3′) and used with a primer (5′-TTT TGG TTG AAA AGG AAG CG-3′) designed to correspond to the end of the sequence reported by Nicolaisen [12]. A 4,200-bp fragment was cloned into TOPO 2.1 (Invitrogen, Carlsbad, CA) and sequenced using the ABI system at the Molecular Biology Core Laboratory of the Washington State University, Pullman, WA. The remainder of the genome was cloned by using specific primers derived from the sequence of the 4,200-bp clone. Standard molecular biological techniques were used for the identification of recombinant clones and preparation of plasmid DNA for sequencing [22]. Sequences were compiled using PC-based AlignX (Vector NTI Suite 9, Informax Inc, Bethesda, MD). Sequences of DMV available in GenBank were used for comparisons. Nucleotide sequences and their putative translation products were compared to other caulimovirus sequences available in GenBank [3] using BLASTN and BLASTX [2]. Pairwise alignment of amino acid sequences to determine sequence identity and similarity was carried out using the Needleman–Wunsch Global Alignment in EMBOSS [11]. DNA fragments were assembled into contigs, and the contigs were assembled to give complete genome sequence using ContigExpress (Vector NTI). Amino acid sequence alignments and phylogenetic analysis of each DMV open reading frame (ORF) with ORFs of other caulimoviruses was done using CLUSTALW version 1.83 [24] and MEGA3 [10].

The complete genome of DCMV, assembled from overlapping clones, was 8,012 nucleotides in length and coded for six major open reading frames (ORF), ORFs I to VI, and one minor ORF (ORF VII) of unknown function. The ca. 1.1-kb sequence described from a Dutch isolate [12] shared 90% sequence identity with the corresponding region of DCMV. The size of each ORF and their organization were typical of the members of the genus *Caulimovirus* [9]. ORF1 through VI could potentially code for movement protein (MP), aphid transmission factor (ATF), DNA-binding protein (DNAb), coat protein (CP), reverse transcriptase (RT), and inclusion body protein (IB), respectively. Comparisons of amino acid sequences of various ORFs of DCMV with those of previously characterized isolates of the distinct caulimovirus species DMV and DMV-D10 showed that sequence similarities ranged from 25 to 80% (Table 1).

The putative MP coded by DCMV contained the transport domain (GNLAYGKFMFTVY) with one amino acid variation where A was replaced by C. This domain has been shown for other caulimoviruses to be important for virus cell-to-cell movement. The amino acid sequence of DCMV had 80 and 49% sequence identity with DMV and DMV-D10, respectively (Table 1). Phylogenetic analysis showed that the MP of DCMV was closest to that of DMV and Mirabilis mosaic virus (MMV) (Fig. 1a). The putative ATF coded by the ORF II of DCMV clustered with that of DMV and MMV but showed greater divergence from DMV (Fig. 1b). The amino acid sequence motif IXG, X being any amino acid, which has been found to be necessary for the interaction between the ATF and virus particles for aphid transmission [23], was also found in the ORF II of DCMV. Phylogenetic analysis of the putative DNAb coded by the ORF III of DCMV showed that it forms a cluster within the clade formed by DMV, MMV, DMV-D10, and figwort mosaic virus (FMV) (Fig 1c). The ORF IV codes for a putative CP, and the CP of DCMV had 52% sequence identity with that of DMV and was distinct from DMV and DMV-10 in the cluster dendrogram (Fig. 1d). The deduced amino acid sequence contained an RNA-binding domain that is consistent with a cysteine motif or "zinc finger" ($CX_2CX_4HX_4C$) [7, 8].

ORF V of DCMV encodes a polyprotein and contains all of the motifs that are conserved in caulimovirus replicases,

Table 1 Amino acid sequence identities of each of the open reading frames (ORF) of dahlia common mosaic virus with the corresponding ORFs of dahlia mosaic virus (DMV) and DMV-D10

	ORF1	ORF2	ORF3	ORF4	ORF5	ORF6
DMV (%)	80	65	45	52	71	76
D10 (%)	49	ND[a]	40	25	70	40

[a] Not determined. ORF2 is not present in the D10 genome. The low sequence identity with ORF4 of D10 was because the coat protein potentially coded by the ORF4 in D10 was truncated [17]

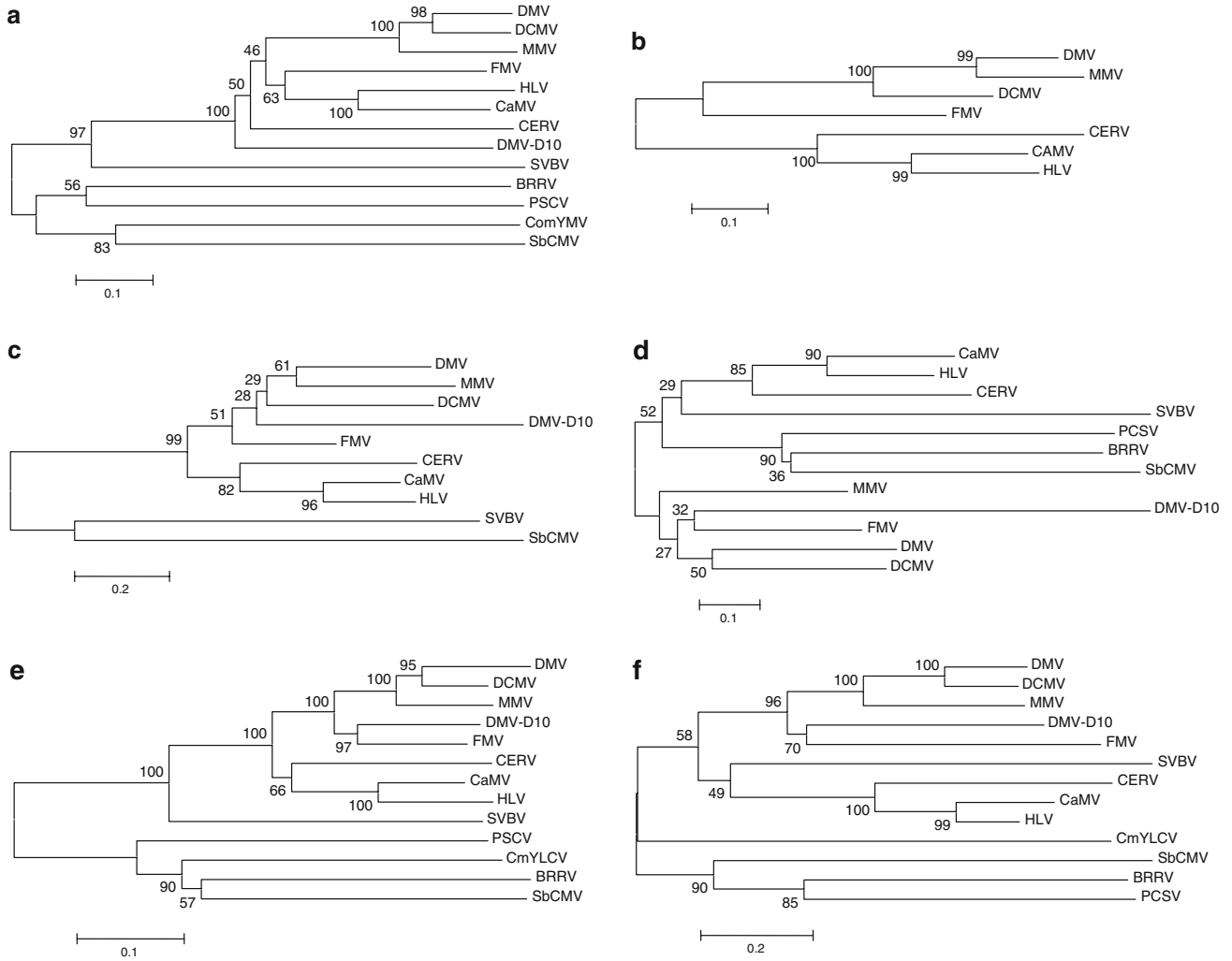

Fig. 1 Phylograms drawn from Clustal W alignments of the different open reading frames (ORF) of selected members of the family *Caulimoviridae*. **a–f** ORFs I through VI, respectively. *BRRV* Blueberry red ringspot virus, *CaMV* cauliflower mosaic virus, *CERV* carnation etched ring virus, *CYLV* Cestrum yellow leaf curling virus, *ComYMV* commelina yellow mottle virus, *DCMV* dahlia common mosaic virus, *DMV* dahlia mosaic virus, DMV-D10, *FMV* figwort mosaic virus, *HLV* horseradish latent virus, *MMV* Mirabilis mosaic virus, *PCSV* peanut chlorotic streak virus, *SbCMV* soybean chlorotic streak virus, *SVBV* strawberry vein banding virus. Bootstrap values are indicated at branching points in the phylogram as a percentage of 1,000 iterations

such as aspartic protease, reverse transcriptase (RT) and ribonuclease H (RNase H) [9]. DCMV shared 70–71% sequence identity in CP with DMV-D10 and DMV (Table 1). The phylogram based on the protein sequences showed a close clustering of DCMV with MMV and DMV (Fig. 1e). The ORF-VI encodes the viral IB. Phylogenetic relationships among the IBs of caulimoviruses showed clustering of DCMV with DMV (Fig. 1f), as they shared 76% sequence identity (Table 1). The ORF-VII of DCMV potentially encodes a minor protein preceded by an intergenic region. A similar ORF is present in all known caulimoviruses, but functional proteins have not been found in infected plants [25]. The large intergenic region between ORFs VI and VII of DCMV contains a putative 35S promoter homolog. The TATA box as well as the conserved residues TC around the TATA box were identified (not shown).

The genome organization, sequence relationships and phylogenetic analysis of DCMV suggested that it is a member of a distinct species within the genus *Caulimovirus*. While the overall genome organization of DCMV conformed to the features of members of the genus *Caulimovirus*, significant sequence divergence was noticed when compared to DMV and DMV-D10. Several partial sequences (EF203677, EF203676, EF463101) with high (over 90%) sequence identity to the DCMV sequence we reported here were recently submitted to GenBank [3] from different parts of the world. They could be potential sequence variants of DCMV. This study reports the complete genome characterization of an isolate of this new and

distinct species of genus *Caulimovirus* associated with dahlia. With the two previously reported caulimoviruses in dahlia (DMV and DMV-D10), it is now clear that there are at least three distinct caulimoviruses that are associated with dahlia mosaic [4, 12, 17, 21]. In our recent surveys of dahlias in the US, we found symptomatic plants that were positive for only one of these viruses, suggesting that each virus is potentially capable of initiating infection in dahlia. However, the pathogenecity of DCMV remains to be proven. Among the three caulimoviruses, DMV-D10, which exists as an EPRV [18], was found to be the most prevalent caulimoviral sequence in dahlias in the US [16, 19], followed by DCMV and DMV (S. Eid, K. L. Druffel and H. R. Pappu, unpublished). Dahlia is primarily propagated by vegetative means, and our finding of another distinct caulimovirus highlights the need for developing virus-specific detection and management options for these viruses.

Acknowledgments Research was supported in part by the Samuel and Patricia Smith Endowment for Dahlia Virus Research, created by the American Dahlia Society, and a grant from the USDA ARS Northwest Nursery Crop Research Center. PPNS No. 0506, Department of Plant Pathology, College of Agricultural, Human and Natural Resource Sciences, Agricultural Research Center, Project # WNPO 0545, Washington State University, Pullman, WA 99164-6430, USA.

References

1. Albouy J (1995) Dahlia. In: Lobenstein G, Lawson RH, Brunt AA (eds) Virus and virus-like diseases of bulb and flower crops. John Wiley and Sons, Chichester, pp 265–273
2. Altschul F, Thomas LM, Alejandro AS, Jinghui Z, Zheng ZMW, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402
3. Benson DA, Karsch-Mizrach I, Lipman DJ, Ostell J, Wheeler DL (2005) GenBank. Nucleic Acids Res 33:D34–D38
4. Bogunov YV (2006) Identification of *Dahlia mosaic virus* with molecular biological methods. Mol Biol 40:184–185
5. Brierley P, Smith FF (1950) Some vectors, hosts and properties of dahlia mosaic virus. Plant Dis Rep 34:363–370
6. Brunt AA (1971) Some hosts and properties of dahlia mosaic virus. Ann App Biol 67:357–368
7. Glasheen BM, Polashock JJ, Lawrence DM, Gillet JM, Ramsdell DC, Vorsa N, Hillman BI (2002) Cloning, sequencing, and promoter identification of Blueberry red ringspot virus, a member of the family Caulimoviridae with similarities to the "Soybean chlorotic mottle-like" genus. Arch Virol 147:2169–2186
8. Hasegawa A, Verver J, Shimada A, Saito M, Goldbach R, Van Kammen A, Miki K, Kameya-Iwaki M, Hibi T (1989) The complete sequence of soybean chlorotic mottle virus DNA and the identification of a novel promoter. Nucleic Acids Res 17: 9993–10013
9. Hull R (2002) Matthew's plant virology, 4th edn. Academic Press, New York, p 1001
10. Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163
11. Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequences of two proteins. J Mol Biol 48:443–453
12. Nicolaisen M (2003) Partial molecular characterization of Dahlia mosaic virus and its detection by PCR. Plant Dis 87:945–948
13. Pahalawatta V, Druffel K, Pappu HR (2005) Molecular characterization of Dahlia mosaic virus genome. Phytopathology 95:S79
14. Pahalawatta V (2007) Biological and molecular properties of Dahlia mosaic virus. PhD thesis. Washington State University, Pullman, WA, USA
15. Pahalawatta V, Druffel K, Pappu HR (2007) Seed transmission of Dahlia mosaic virus in *Dahlia pinnata*. Plant Dis 91:88–91
16. Pahalawatta V, Miglino R, Druffel KL, Jodlowska A, van Schadewijk AR, Pappu HR (2007) Incidence and relative distribution of distinct Caulimoviruses (Genus *Caulimovirus*, Family *Caulimoviridae*) associated with dahlia mosaic in *Dahlia variabilis*. Plant Dis 91:1194–1197
17. Pahalawatta V, Druffel K, Wyatt SD, Pappu HR (2008) Molecular characterization of a novel and distinct species of genus *Caulimovirus* (Family *Caulimoviridae*) associated with dahlia mosaic. Arch Virol 153:733–738
18. Pahalawatta V, Druffel KL, Pappu HR (2008) A new and distinct species in the genus Caulimovirus exists as an endogenous plant pararetroviral sequence in its host, *Dahlia variabilis*. Virology 376:253–257
19. Pappu HR, Wyatt SD (2003) Dahlia mosaic virus. A description of symptoms on dahlia. APSnet Image of the Week. Online/2003/IW000032.asp
20. Pappu HR, Wyatt SD, Druffel KL (2005) Dahlia mosaic virus: molecular detection and distribution in dahlia in the US. HortScience 40:697–699
21. Richins RD, Shepherd RJ (1983) Physical maps of the genomes of Dahlia mosaic virus and Mirabilis mosaic virus—two members of the Caulimovirus group. Virology 124:208–214
22. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, New York
23. Schmidt I, Blanc S, Esperandieu P, Kuhl G, Devauchelle G, Louis C, Cerutti M (1994) Interaction between the aphid transmission factor and virus particles is a part of the molecular mechanisms of cauliflower mosaic virus aphid transmission. Proc Natl Acad Sci USA 91:8885–8889
24. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acid Res 22:4673–4680
25. Wurch T, Kirchherr D, Mesnard J, Lebeurier G (1990) The Cauliflower mosaic virus open reading frame VII product can be expressed in *Saccharomyces cerevisiae* but is not detected in infected plants. J Virol 64:2594–2598