BIOLOGICAL AND MOLECULAR CHARACTERIZATION OF

DAHLIA MOSAIC CAULIMOVIRUS

By

VIHANGA PAHALAWATTA

**A dissertation submitted in partial fulfillment of
the requirements for the degree of**

DOCTOR OF PHILOSOPHY

**WASHINGTON STATE UNIVERSITY
Department of Plant Pathology**

AUGUST 2007

To the Faculty of Washington State University:

 The members of the Committee appointed to examine the dissertation of VIHANGA

PAHALAWATTA find it satisfactory and recommend that it be accepted.

<div align="right">

_____

Chair

_____

_____

_____

</div>

ACKNOWLEDGEMENT

BIOLOGICAL AND MOLECULAR CHARACTERIZATION OF DAHLIA MOSAIC CAULIMOVIRUS

ABSTRACT

By Vihanga Pahalawatta, Ph.D.
Washington State University
August 2007

Chair: Hanu R. Pappu

Dahlia is an economically important ornamental crop in the USA and several countries around the world. Among the viral diseases that affect dahlia, *Dahlia mosaic virus* (DMV), is considered to be the most widespread and to have the greatest impact on flower production. Molecular characterization of DMV showed the association of two distinct caulimoviruses (DMV-D10, DMV-Portland) and a D10-like sequence variant (DMV-Holland) with the disease. The genome structure and organization of DMV-D10 was determined. The double stranded DNA genome shared many features with the members of the genus *Caulimovirus*. However, sequence identity of various open reading frames of DMV-D10 with known caulimoviruses ranged from 32-70% at amino acid level suggesting that D10 is a distinct caulimovirus. Moreover, the aphid transmission factor was absent and the coat protein was smaller than that found in other *Caulimovirus* species. Phylogenetic analysis using amino acid sequences from the different protein coding regions indicated that DMV-D10 is closely related to DMV-Portland and *Figwort mosaic virus.*

A PCR-based assay was used to determine the relative incidence of DMV-D10, DMV-Portland and DMV-Holland in dahlia samples from the USA and the Netherlands. Results

indicated that DMV-D10 was present in 87.4% of samples from the USA and 97.6% of the samples from the Netherlands. Furthermore, 97.6% of samples from the Netherlands were positive for DMV-Holland while only 31.9% of the samples from the USA were positive. The presence of DMV-Portland in samples from both the USA and the Netherlands was comparatively less. The presence of all three DMV sequences in 36.6% of the dahlia samples from the Netherlands and 6.7% of the dahlia samples from the USA highlighted the importance of testing for all three sequences for DMV diagnostics. The distribution of DMV in the dahlia seed was determined by PCR. The viral genome was detected in cotyledons. Seedlings from grow-out tests were also positive for DMV-D10 indicating that the virus is seed transmitted. The virus was also detected in the pollen collected from infected plants.

Viruses in certain genera in family *Caulimoviridae* were shown to integrate their genomic sequences into their host genomes and exist as endogenous pararetroviral sequences. However, members of the genus *Caulimovirus* remained to be an exception and are believed to exist only as episomal elements during virus infection cycle. Using DMV-D10 specific probes, Southern blot hybridization of total plant DNA from dahlia seedlings showed the presence of viral DNA in the host DNA whereas, viral DNA was not detected in other known hosts of DMV. This is the first report of natural integration of a caulimovirus genome into its host and existence of a caulimovirus species as an endogenous pararetroviral sequence in a host plant. Epidemiological and molecular properties of DMV-D10 would facilitate devising management options for reducing the impact of the virus.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEDICATION

This thesis is dedicated to my parents, who have supported me

unconditionally and inspired me to reach my potential.

CHAPTER ONE

BIOLOGICAL AND MOLECULAR CHARACTERIZATION OF DAHLIA MOSAIC CAULIMOVIRUS

GENERAL INTRODUCTION

Plants in genus *Dahlia* Cav. belong to family *Asteraceae*. At present, 35 species belonging to the genus *Dahlia* have been identified with over 20,000 cultivars. Most *Dahlia* species are indigenous to the highlands of Mexico. The exceptions are *D. australis* found in southwestern Guatemala, *D. coccinea* and *D. imperialis* which occur in parts of Central America and Northern South America (Saar *et al.,* 2003). Dahlia plants were first discovered in the 16th century by botanists who accompanied the Spanish Conquistadores to Mexico. The Aztec Indians in Mexico are reported to have used the plant for food and medicinal purposes. The Aztecs referred to dahlia by the name 'Acocotli', however very little is known about the plant prior to the Aztecs.   The first dahlia samples were brought to the Royal Botanical Gardens in Madrid, Spain in 1789 from Mexico. Abbé Cavanilles, the curator at the time flowered them successfully and coined the name 'Dahlia' in honor of the Swedish botanist Andreas Dahl (1751-1789) (Damp, 1981). During the early 1800's dahlias spread throughout Europe. According to Sorenson (1969), the exact origin of the present day garden dahlia, *D. variabilis* hort. non (Willd.) Desf., is unclear although it is widely accepted that it is a hybrid between *D. pinnata* (Cav.) and *D. coccinea*. Observation of chromosomes of *D. variabilis* showed *2n*=64 (Gatt *et al.,* 1998). Plant breeders have been actively breeding dahlias for many years resulting in thousands

of cultivars. At present, dahlia is an important bulb crop used as bedding plants, garden specimen plants and cut flowers, especially in the US, Australia, Canada, Europe, Japan and New Zealand (Albouy, 1995; Armitage,1993).   Propagation of dahlias is mainly through tuberous rooted divisions, stem cuttings, or seed (Dole and Wilkins, 1999).

In recent years, floral and other ornamental crops have achieved the fastest sales growth among the various crops grown in the US. In 2003, ornamental crops ranked fifth among the top eight agricultural sectors that gross at least $13 billion in annual cash receipts, second only to corn and vegetables among crops (Jerardo, 2004). According to USDA statistics in 2005, sales of ornamental crops reached a value of $16.2 billion.   In 2006, ornamental crop sales were projected to reach $16.5 billion with $100 million from floriculture crops (Jerardo, 2006). Imports of cut flowers and nursery stock to the US were valued at $1.4 billion in 2005. Export of these floriculture products from the US was estimated to be approximately one fourth of the value of imports although there has been a steady rise especially to Canada, Europe and Asia (Geisler, 2006).   At the outset, trade in ornamental crops did not pose a serious threat of distributing plant diseases, especially viruses, mainly because most of the germplasm was traded as seed. However with recent advances in storing, handling and transport facilities, the trade of vegetatively propagated clonal material is increasing, along with the risk of dissemination of viral and other graft-transmissible agents. Infected cut flower crops have been implicated in transporting some of the first viruses to be spread across national and international borders (Lawson and Hsu, 2006). Virus diseases in general impede plant growth and damage some or all of the plant organs which results in a reduction of the aesthetic quality and subsequently the marketability.   Quarantine regulations and certification programs are in place to prevent the

movement of plant pathogens. However in cases of asymptomatic infections and instances where adequate detection methods are not available, application of these regulations and certification programs is unfeasible. Among the diseases affecting dahlias, viral pathogens play a significant role primarily due to vegetative propagation of infected plant material.

Major viruses affecting dahlia crops include *Cucumber mosaic virus* (CMV), *Dahlia mosaic virus* (DMV), *Tobacco streak virus* (TSV), and *Tomato spotted wilt virus* (TWSV). When considering the prevalence and the economic impact of the viruses, DMV is considered to be the most important viral pathogen of dahlias (Brunt, 1971; Albouy, 1995). *Dahlia mosaic virus* was first reported in *Dahlia pinnata* from Germany in 1928 (Brunt, 1971). The disease is considered to be geographically widespread and ubiquitous in dahlia growing areas although the natural host range is limited to *Dahlia spp*. The most characteristic symptoms of the disease are mosaic and vein-banding accompanied by stunting and leaf distortion. Symptoms vary depending on the cultivar infected whereas cultivars with few or no symptoms are considered to be important reservoirs of infection. The economic impact of the disease is due to its negative effect on the quality and quantity of flowers produced by diseased plants.

Dahlia mosaic virus is a member of family *Caulimoviridae*. Family *Caulimoviridae* encompasses all plant viruses that consist of a double stranded DNA (dsDNA) genome. The viruses are monopartite with bacilliform or isometric particles. The defining feature of family *Caulimoviridae* is the presence of a reverse transcription stage in their replication cycle. The following steps have been identified in the caulimovirus replication cycle following the entry of the virus into the host cell using *Cauliflower mosaic virus* as a model. (i) Nuclear localization signal mediated transport of virus particles to the nuclear pore. (ii) Uncoating of viral DNA and

3

import into the nucleus. (iii) Association of DNA with histones to form minichromosomes. (iv) Transcription of the viral DNA by cellular DNA dependant RNA polymerase II to produce two RNA intermediates (19S RNA and 35S RNA). (v) Reverse transcription of RNA intermediates to produce dsDNA (Haas *et al.,* 2002). Thus viruses in family *Caulimoviridae* are not true dsDNA viruses due to the requirement of RNA intermediates for their replication and are referred to as pararetroviruses. Pararetroviruses differ from true retroviruses in two main features. First, retroviruses have a genome composed of RNA while pararetroviruses have a DNA genome. Second, in retroviruses the DNA resulting from the reverse transcription of the RNA genome is integrated into the host DNA with the use of an integrase enzyme, whereas in pararetroviruses integration is not an essential part of the replication cycle. In pararetroviruses, the reverse transcribed DNA behave as free chromosomes in the host cell (Hull and Covey, 1996). Viruses in family *Caulimoviridae* are classified in six genera; genus *Badnavirus* (type member *Commelina yellow mottle virus*), genus *Caulimovirus* (type member *Cauliflower mosaic virus*), genus *Cavemovirus* (type member *Cassava vein mosaic virus*), genus *Petuvirus* (type member *Petunia vein clearing virus*), genus *Soymovirus* (type member *Soybean chlorotic mottle virus*), genus *Tungrovirus* (type member (*Rice tungro bacilliform virus*). Figure 1 represents the genome organization of the type members of the six virus genera listed above (Glasheen *et al*., 2002). *Dahlia mosaic virus* has been classified under genus *Caulimovirus* based on its genome structure and organization. It is composed of isometric virion particles 48-50 nm in diameter. There is a single report on the physical map of the DMV genome (Richins and Shepherd, 1983). Partial nucleotide sequences, representing various regions of the DMV genome that have been submitted to the National Center for Biotechnology Information (NCBI) database is listed in

Table 1. Previous studies on DMV transmission have resulted in identifying 16 aphid species, including *Aphis fabae, Myzus persicae* and *Macrosiphum euphorbiae* (Brierley and Smith, 1950), as vectors transmitting the virus in a non-persistent manner. A virus encoded aphid transmission protein is reported to be necessary for viral transmission (Albouy, 1995). DMV was also shown to be transmitted by mechanical inoculation (Brierley and Smith, 1950; Brierley, 1951; Brunt, 1971). *Verbesina encelioides, Ageratum conyzoides, Zinnia elegan, Amarantus caudatus* and *Chenopodium capitatum* are plant host species susceptible to DMV although the natural host range of the virus is limited to *Dahlia spp.* Furthermore, *V. encelioides* and *Z. elegans* have been reported to be suitable maintenance and propagation hosts for DMV (Brierley and Smith, 1950; Brierley, 1951; Brunt, 1971). *Dahlia mosaic virus* infected plants are also reported to have viral inclusion bodies in the cells (Robb, 1963). Overall there is very limited information on the molecular and biological properties of the virus. This lack of information is a major impediment in the development of control options. The main goal of our study was to understand the important biological and molecular biological properties of DMV. Our specific objectives were to characterize the virus(es) associated with dahlia mosaic disease, determine the genome structure and organization of the virus(es) and to develop suitable diagnostic assays that would assist in rapid detection of the virus as well as help in developing suitable control options.

Genus *Badnavirus*
Commelina yellow mottle virus (ComYMV)
  I   II   III

Genus *Caulimovirus*
Cauliflower mosaic virus (CaMV)
 VII   I   II   III   IV   V   VI

Genus *Cavemovirus*
Cassava vein mosaic virus (CsVMV)
                  I   II   III   IV   V

Genus *Petuvirus*
Petunia vein clearing virus (PVCV)
                  I   II

Genus *Soymovirus*
Soybean chlorotic mottle virus (SbCMV)
 VII   I   A   B   C   IV   V   VI

Genus *Tungrovirus*
Rice tungro bacilliform virus (RTBV)
  I   II   III   IV

Figure 1. Genome organization of viruses belonging to genera in family Caulimoviridae. Coding regions with the same pattern are considered homologs. Roman numerals and capital letters represent the open reading frames (ORFs). ▨ Protein of unknown function; ▨ ORF III ComYMV and RTBV coat protein, aspartic protease, reverse transcriptase and ribonuclease H; ▨ Cell to cell movement protein; ▨ Aphid transmission factor; ▨ Nucleic acid binding protein; ▨ Coat protein; ▨ Aspartic protease, reverse transcriptase, ribonuclease H; ▤ Inclusion body/translational transactivator protein; ▨ ORF I CsVMV RNA binding protein, coat protein, cell to cell movement protein; ▨ ORF II PVCV RNA binding protein, protease domain, reverse transcriptase, ribonuclease H.

Table 1. A description of *Dahlia mosaic virus* sequences available in GenBank

Sequence definition	GenBank accession number	Sequence length (bp)	Authors	Date submitted
Full length transcript promoter; and inclusion body matrix protein gene, partial cds	EF513491	442	Kulvev, B. Chemeris, A.	04/10/2007
Inclusion body matrix protein gene partial cds; and intergenic spacer, partial sequence	EF463101	678	Kulvev, B. Chemeris, A.	04/03/2007
Isolate D2 putative capsid protein gene partial cds	EF203678	1300	Pham, K. T. K. Van Doorn, J. Derks, A. F. L. M.	02/03/2007
Isolate D1 putative capsid protein gene, partial cds	EF203677	1068	Pham, K. T. K. Van Doorn, J. Derks, A. F. L. M.	02/03/2007
Isolate 2 truncated reverse transcriptase gene, partial cds	EF203676	1082	Pham, K. T. K. Van Doorn, J. Derks, A. F. L. M.	02/03/2007
Isolate 1 reverse transcriptase gene, partial cds	EF203675	966	Pham, K. T. K. Van Doorn, J. Derks, A. F. L. M.	02/03/2007
Putative capsid protein gene, partial cds	AY971810	323	Buganov, Y. V.	03/16/2006
Partial ORF4 for putative capsid protein and partial ORF5 for putative reverse transcriptase	AJ515906	1171	Nicolaisen, M.	04/15/2005
Inclusion body matrix protein (VI) gene, complete cds	AY309480	1515	Maiti, I. B. Pattanaik, S.	06/30/2003
Polyprotein (V) gene, complete cds	AY309479	1830	Maiti, I. B. Pattanaik, S.	06/30/2003
Putative capsid protein gene, complete cds	AY291588	1476	Maiti, I. B. Pattanaik, S.	06/30/2003
Putative DNA-binding protein/minor viral capsid protein gene, complete cds	AY291587	363	Maiti, I. B. Pattanaik, S.	06/30/2003
Putative aphid transmission protein gene, complete cds	AY291586	504	Maiti, I. B. Pattanaik, S.	06/01/2003
Putative movement protein gene, complete cds	AY291585	966	Maiti, I. B. Pattanaik, S.	06/01/2003

REFERENCES

Albouy, J. 1995. Dahlia. Pages 265-273 in: Virus and virus-like diseases of bulb and flower crops. G. Lobenstein, R. H. Lawson, A. A. Brunt, A. A. eds. John Wiley and Sons. Chichester, New York.

Armitage, A. M. 1993. Dahlia hybrids. Pages 268-271 in: Specialty Cut Flowers. The production of annuals, perennials, bulbs, and woody plants for fresh and dried cut flowers. Varsity Press/Timber Press, Portland, Oregon.

Brierley, P. 1951. Value of index plants for detecting dahlia viruses. Plant Dis. Rep. 35: 405-407.

Brierley, P., Smith, F. F. 1950. Some vectors, hosts and properties of dahlia mosaic virus. Plant Dis. Rep. 34: 363-370.

Brunt, A. A. 1971. Some hosts and properties of dahlia mosaic virus. Anal. App. Biol. 67: 357-368.

Damp, P. 1981. Growing dahlias. Croom Helm Ltd, London.

Dole, J. M., Wilkins, H. F. 1999. Dahlia. Pages 287-292 in: Floriculture: principles and species. Prentice-Hall Inc, New Jersey.

Gatt, M., Ding, H. Hammett, K. Murray, B. 1998. Polyploidy and evolution in wild and cultivated dahlia species. Anal. Botany 81: 647-656.

Geisler, M. 2006. Agricultural marketing resource center. Floriculture profile. http://www.agmrc.org/agmrc/commodity/specialitycrops/floriculture/floricultureprofile.htm. Accessed online April, 2007.

Glasheen, B. M., Polashock, J. J., Lawrence, D. M., Gillett, J. M., Ramsdell, D. C., Vorsa, N., Hillman, B. I. 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.

Haas, M., Bureau, M., Geldreich, A., Yot, P., Keller, M. 2002. *Cauliflower mosaic virus*: still in the news. Mol. Plant Pathol. 3(6): 419-429.

Hull, R., Covey, S. N. 1996. Retroelements: Propagation and adaptation. Virus genes 11: 105-118.

Jerardo, A. 2006. Floriculture and nursery crops outlook. Electronic outlook report from the economic research service.

http://www.ers.usda.gov/publications/flo/2006/09Sep/FLO05.pdf. Accessed online April, 2007.

Jerardo, A. 2004. Amber waves. The economics of food, farming, natural resources, and rural America.

http://www.ers.usda.gov/amberwaves/February04/Findings/VolumeProduction.htm. Accessed online April, 2007.

Lawson, R., Hsu, H. T. 2006. Quarantine viruses, viroids and phytoplasmas that affect the movement of ornamental plants. Acta Hort. 722: 17-30.

Richins, R. D., Shepherd, R. J. 1983. Physical maps of the genomes of *Dahlia mosaic virus* and *Mirabilis mosaic virus* - two members of the Caulimovirus group. Virology 124: 208-214.

Robb, S. M. 1963. Inclusion bodies in dahlia infected with *Dahlia mosaic virus*. Nature 200: 1339-1340.

Saar, D. E., Polans, N. O., Sørensen, P. D. 2003. A phylogenetic analysis of the genus *Dahlia*

(*Asteraceae)* based on internal and external transcribed spacer regions of nuclear

ribosomal DNA. Systematic botany 28(3): 627-639.

Sørensen, P. D. 1969. Revision of the genus *Dahlia* (Compositae, Heliantheae-Coreopsidinae).

Rhodora 71: 309-365.

CHAPTER TWO

GENOME STRUCTURE, ORGANIZATION AND INTERVIRAL HOMOLOGIES OF A NOVEL AND DISTINCT SPECIES OF GENUS *CAULIMOVIRUS* ASSOCIATED WITH DAHLIA MOSAIC

ABSTRACT

The genome structure and organization of a new and distinct caulimovirus that is widespread in dahlia (*Dahlia pinnata*) was determined. The double stranded DNA genome was ca. 7.0 kb in size and shared many of the features of the members of the genus *Caulimovirus* such as the presence of genes potentially coding for the movement protein, the inclusion body protein, the reverse transcriptase, and an intergenic region consisting of a potential 35S promoter. However, the virus differed from known caulimoviruses in that the aphid transmission factor was absent and the coat protein was smaller than its counterpart in known *Caulimovirus* species. Sequence homology at the amino acid level with known caulimoviruses was low and ranged from 32-70%. The absence of an aphid transmission factor and the highly divergent nature of the genomic sequence are characteristics of this newly characterized caulimovirus that is widely associated with dahlia mosaic disease.

INTRODUCTION

Dahlia mosaic caulimovirus (DMV) is an important viral pathogen of dahlia (*Dahlia*

pinnata) in the US and several parts of the world. First reported from Germany in 1928, the virus is considered to be one of the most important disease constraints affecting dahlias. DMV is a member of the family *Caulimoviridae,* genus *Caulimovirus* with a circular double stranded-DNA genome of approximately 7 kb. The symptomatology, propagative hosts and the role of various aphid species in virus transmission were reported (Albouy, 1992; Brierley & Smith, 1950, Brunt, 1971). The most characteristic symptoms of the disease include mosaic and vein-banding accompanied by stunting and leaf distortion. Symptoms vary depending on the cultivar, and cultivars with few or no symptoms are common. The economic impact of the disease is due to its negative affect on the quality and quantity of flowers produced by diseased plants.

The physical map of a caulimovirus (DMV-Portland) genome associated with dahlia mosaic was reported (Richins & Shepherd, 1983), and its genomic sequence representing six open reading frames is available in GenBank (GenBank accessions AY309480, AY309479, AY291588, AY291587, AY291586, and AY291585). However, our surveys of dahlias to determine the incidence of DMV revealed the overwhelming presence of another caulimovirus, designated DMV-D10 that was found to be distinct from DMV-Portland (Pahalawatta *et al*., 2005; Pahalawatta *et al*., 2007a; Pahalawatta *et al*., 2007b; Pappu *et al*., 2005). A partial sequence representing a D10-like sequence variant, DMV-Holland (Nicolaisen, 2003) has also been reported. To better understand the identity and taxonomic position of the caulimoviruses associated with dahlia mosaic, the genome structure and organization of DMV-D10 genome was deciphered and was compared to DMV-Portland and other members of the genus *Caulimovirus*.

MATERIALS AND METHODS

DMV isolates and DNA extraction. D10-infected dahlia plants were collected from Dahlia Trial Garden, Tacoma, WA. Total genomic DNA was extracted from infected dahlia leaves following a modified Dellaporta method (Presting *et al*., 1995). Approximately 0.1g (or a quarter-sized leaf piece) was ground in liquid N_2 followed by the addition of 1ml extraction buffer (100mM Tris, pH 8.0, 50mM EDTA, 500mM NaCl, and 10mM ß-mercaptoethanol) and 140ul 10% sodium dodecyl sulfate. Samples were mixed by inversion and incubated at 65°C for 10 minutes followed by addition of 250ul 8M potassium acetate. Samples were incubated on ice for 5 minutes, spun at maximum speed in the Eppendorf microfuge for 5 minutes and 1ml of the aqueous phase pulled off and added to tubes containing 0.6ml isopropanol. Tubes were inverted to mix and placed on ice for 3 minutes. Samples were spun for 5 minutes at maximum speed in the microfuge. Pellets were rinsed with 1ml 70% ethanol, spun briefly and the ethanol decanted. Pellets were air-dried followed by resuspension in 100ul sterile double distilled water. Ten microliters DNA was diluted 1:5 by addition of 40ul distilled water.

Cloning and sequencing. Initial cloning of the D10 genome was accomplished by using primers V4739 (5'-TGTCCACAAGGTCACTA-3') and C5400 (5'-TKTCRTTRCTRTGGTAATT-3') derived from the conserved regions of the reverse transcriptase gene of caulimoviruses. A 670 bp fragment was amplified, cloned and sequenced. The sequence was utilized to design primers and obtain additional clones of the remaining genomic DNA.

The PCR reaction included a final volume of 2µl 10X PCR (Promega, Madison, WI)

13

buffer and a final concentration of 0.15mM dNTP's, 2mM $MgCl_2$ and 0.6µM of each primer in a total reaction volume of 20µl. PCR amplification was performed in a DNA thermocycler (BioRad, Hercules, CA) programmed for 3 min at 94^oC for initial denaturation and 50 cycles each consisting of 30 sec at 94^oC, 20 sec at T_m-5^oC and 1 min extension per 1000 bp product at 72^oC, followed by a final extension for 7 min at 72^oC. Each PCR reaction included a DNA sample from a DMV-infected plant and a water control. The PCR products were analyzed by 1.5% agarose gel electrophoresis in 1X TAE buffer. PCR products of expected size were cloned using pGEM-T (Promega, Madison, WI) or TOPO TA (Invitrogen, Carlsbad, CA) cloning kits according to manufacturers' instructions. Some PCR products were sequenced directly after purification using QIAquick PCR purification kit (Qiagen, Valencia, CA). The ca. 7 kb DMV-D10 genomic DNA was cloned in 6 cloning steps (Figure 1). In each step two specific primers were utilized for PCR amplification (Table 1). Recombinant plasmids were isolated (Sambrook *et al*., 1989) and were sequenced using M13 forward and reverse primers, whereas PCR products were sequenced using the specific primer pairs that were initially used to obtain the PCR product. A minimum of two clones were sequenced from each cloned region to verify sequence data. Nucleotide sequences were determined using the ABI Prism Sequencing System at the Molecular Biology Core Laboratory of the Washington State University, Pullman, WA.

Sequence analysis. Nucleotide sequences and their putative translation products were compared to other caulimovirus sequences available in GenBank (Benson *et al*., 2005) using BLASTN and BLASTX (Altschul *et al*., 1997). Pairwise alignment of amino acid sequences to determine sequence identity and similarity were carried out using the Needleman-Wunsch Global

Alignment in EMBOSS (Needleman & Wunsch, 1970). DNA fragments were assembled into contigs and the contigs were assembled to give complete genome sequence using ContigExpress (Vector NTI Suite 9.0.0, Informax Inc, Bethesda, MD). 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 (Thompson *et al.*, 1994) and MEGA3 (Kumar *et al.*, 2004). The RNA secondary structure with the DMV-D10 leader sequence was predicted using MFold (Zuker *et al.*, 1999).

Table 1. List of names and GenBank accession numbers of the viruses used for multiple sequence alignment and phylogenetic comparison of DMV-D10 amino acid sequences with amino acid sequences of other caulimoviruses. CaMV- *Cauliflower mosaic virus*; CERV- *Carnation etched ring virus*; FMV- *Figwort mosaic virus*; HLV- *Horseradish latent virus*; MMV- *Mirabilis mosaic virus*; SVBV- *Strawberry vein banding virus*; BRRV- *Blueberry red ringspot virus*; CmYLCV- *Commelina yellow leaf curling virus*; PCSV- *Peanut chlorotic streak virus*; SbCMV-*Soybean chlorotic mottle virus*; NA- Not available.

Virus	Viral Protein (Open reading frame)					
	Movement (I)	Aphid transmission (II)	DNA-binding (III)	Capsid (IV)	Polyprotein (V)	Inclusion body (VI)
DMV-Portland	AAP44107	AAP44108	AAP44109	AAP44110	AAP75615	AAP75616
CaMV	P03546	P03549	P03552	P03543	P03555	P03558
CERV	P05396	P05397	P05398	P05399	P05400	P05401
FMV	P09520	P09521	P09522	P09519	P09523	P09524
HLV	AAW56085	AAW56086	AAW56087	AAW56088	AAW56089	AAW56090
MMV	NP_659393	NP_695394	NP_695395	NP_695396	NP_695397	NP_695398
SVBV	CAA65972	CAA65973	CAA65974	CAA65969	CAA65970	CAA65971
BRRV	NP_395472	----------	---------	NP_395468	NP_395469	NP_395470
CmYLCV	NP_861411	NP_861407	---------	NP_861409	NP_861410	NP_861547
PCSV	NP_042515	----------	---------	NP_042512	NP_042513	NP_042514
SbCMV	P15631	----------	---------	P15627	P15629	P15628

Figure 1. The locations of the clones used in the sequencing of the entire genome of *Dahlia mosaic virus*-D10. Letters A to F represent clones. The length of each clone is indicated in base pairs (bp). The hatched box represents a single clone.
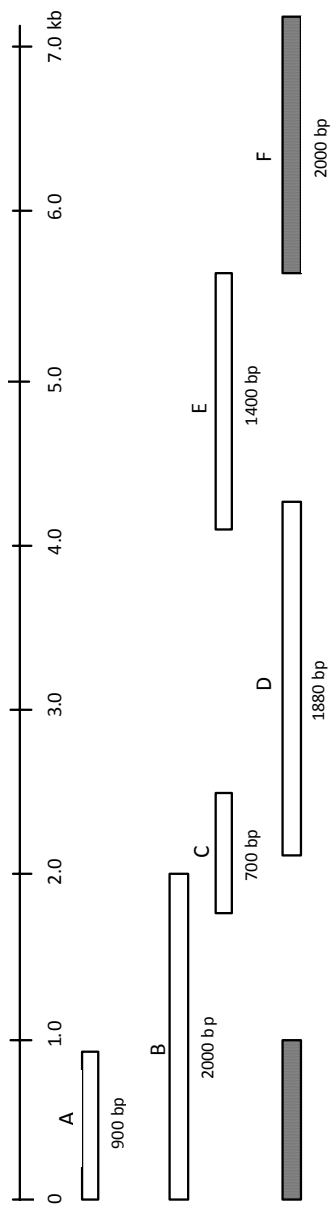
Table 1. Primers used to clone and sequence *Dahlia mosaic virus*- D10. The location of each clone is indicated in Fig. 1.

Clone	Primer	Primer sequence (5'-3')	T_m (^{o}C)
A	ORF1- START	ATG GAT CGT AAA GAT T	55
	ORF1- END1	CTG TTT TTC TGT GTT TCT ACT GG	
B	ORF1-START	ATG GAT CGT AAA GAT T	55
	D5 1396-360F	CCG GTT TTA TAA CTT CTT GG	
C	D5 1337-602F	GAG GAT CTA TGG GAA AAC AC	64
	D10 1322-532F	ATT ATG GCT GTC TCC TTT GG	
D	D10 1321-457F	TTT GAA AAC CCT ATT GAC CC	65
	D10 CORF6	CTC TGT GTT CTC CAT CGA AG	
E	V6100	TCC GTT GAT GGC TGA CAG TT	70
	DAPROR	AGG TAG GGG TTT ACT CAC ACA TT	
F	DAPROF	AAT GTG TGA GTA AAC CCC TAC CT	67
	ORF1-END1	CTG TTT TTC TGT GTT TCT ACT GG	

RESULTS

Genome structure. The D10 genomic sequence consisted of 7046 base pairs (bps) with five complete ORFs and one partial ORF that corresponded to those present in known caulimoviruses. The size of the DMV-D10 genome was comparable to other known caulimoviruses which range from 7.2 kb to 8.0 kb. The A+T composition of the genome was 66.4%, slightly higher than that of other known caulimoviruses for which the A+T composition ranges from 60.0% to 64.6%. Percentage similarity of DMV-D10 proteins to homologous proteins in other caulimoviruses ranged from 25%-74% (Table 2). Figure 2 is a graphical representation of the DMV D10 genome that shows the approximate location of each ORF discussed below. There are several discernable differences between the genome organization of DMV-D10, DMV-Portland and *Cauliflower mosaic virus* (CaMV) (Fig. 3). The complete nucleotide sequence of DMV-D10 and the amino acid translation is given in Figure 4. The DMV-D10 sequence is numbered beginning with the 5' end of the probable tRNAmet primer binding site (5' TGGTATCAGAGCCTTG3') that is predicted to be the primer for negative strand DNA synthesis (Richins *et al*., 1987; Glasheen *et al*., 2002).

Coding regions. ORF I of DMV-D10 consisted of 990 bps encoding the movement protein (MP). A putative transport domain (GNLAYGKFMFTVY; amino acids 173-185), identified in other caulimoviruses to be important for virus cell-to-cell movement, was present within this region (Glasheen *et al*., 2002; Hasegawa *et al*., 1989). Phylogenetic analysis showed that the MP of DMV-D10 was closest to that of *Figwort mosaic virus* (FMV), whereas the MP of DMV-Portland clustered with that of *Mirabilis mosaic virus* (MMV) (Fig. 5a).

19

Table 2. Percentage identity and percentage similarity of *Dahlia mosaic virus* –D10 open reading frames (ORFs) to other caulimoviruses. Percentage similarities are listed within brackets.

Genus *Caulimovirus*	ORF I	ORF III	ORF IV	ORF V	ORF VI
Dahlia mosaic virus DMV-Portland	37.0	35.2	6.9	54.4	29.3
	(53.0)	(50.8)	(11.6)	(64.8)	(40.1)
Cauliflower mosaic virus	41.7	25.9	7.5	58.3	18.6
	(57.1)	(47.4)	(11.4)	(71.2)	(28.4)
Carnation etched ring virus	36.0	27.7	8.5	60.3	19.1
	(49.5)	(33.0)	(11.3)	(71.5)	(31.8)
Figwort mosaic virus	39.4	32.8	7.2	70.1	26.6
	(53.4)	(50.4)	(11.6)	(79.2)	(36.3)
Horseradish latent virus	39.3	35.9	7.3	56.2	19.1
	(54.6)	(45.0)	(10.5)	(68.0)	(29.0)
Mirabilis mosaic virus	35.7	31.3	6.5	62.7	26.0
	(51.7)	(45.5)	(12.7)	(73.0)	(37.3)
Strawberry vein banding virus	24.8	21.6	6.2	46.3	18.2
	(41.6)	(32.1)	(11.9)	(58.8)	(29.6)
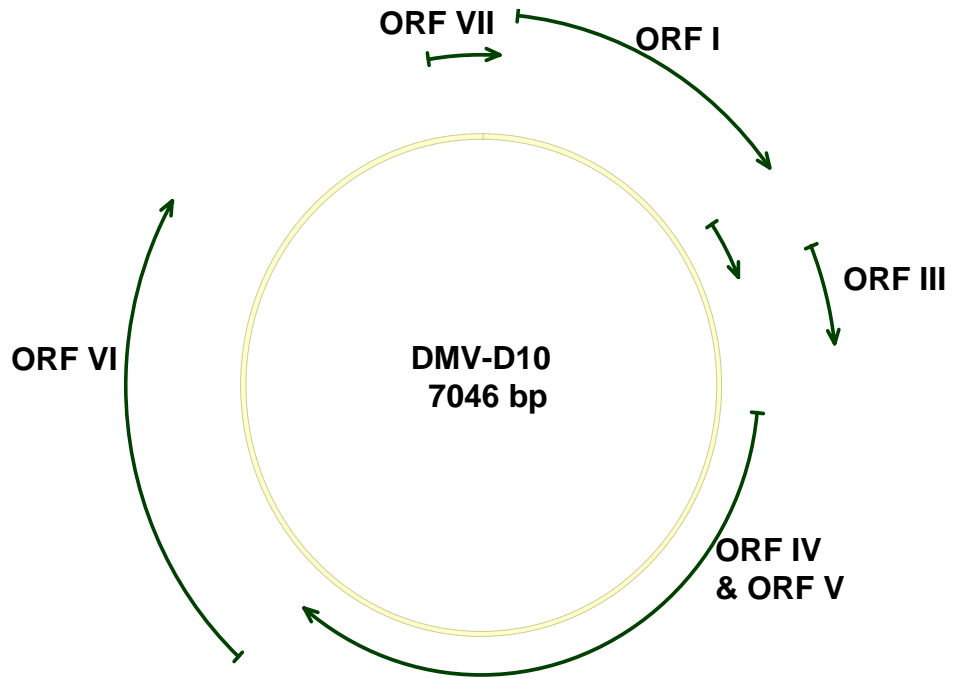
Figure 2. *Dahlia mosaic virus* genome organization. The double-lined closed circle represents

the viral genome. Lines with arrows represent the various open reading frames. Figure not

drawn to scale.

CaMV

DMV-Portland
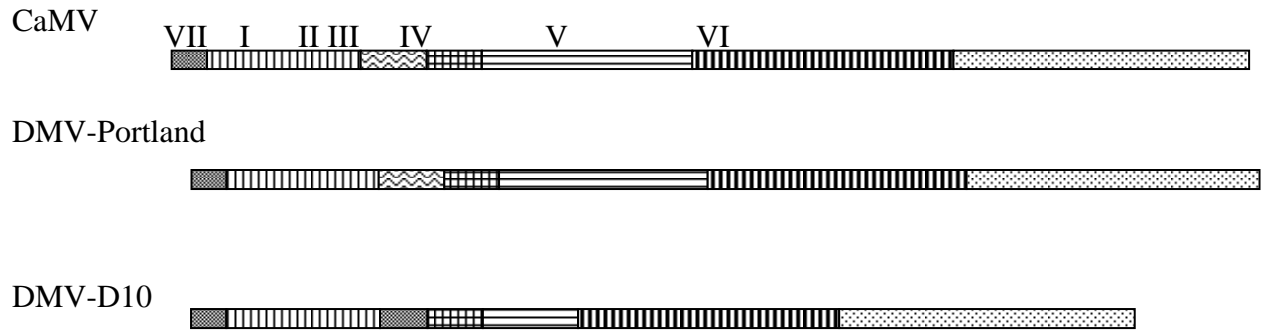
DMV-D10

Figure 3. Genome organization of *Cauliflower mosaic virus* (CaMV) , *Dahlia mosaic virus* DMV-Portland and DMV-D10. Coding regions with same patterns are considered homologs. Roman numerals signify open reading frames. ▨ Protein of unknown function; ⬚ Movement protein; ▧ Aphid transmission factor; ▦ Nucleic acid binding protein; ▤ Coat protein; ⬚ polyprotein; ▨ Inclusion body protein.

22

```
DMV-D10    TGGTATCAGAGCCTTGTGCTCAAAATTTTAAAAAATAAAAGTTTGGTTTATTCTTGAAAA    60
ORF VII      G   I   R   A   L   C   S   K   F   *

DMV-D10    AGAAAAATTAAAAACTTTATTTTTAAAGTAAAAAATTATTTTAAATGGATCGTAAAGATT   120
ORF I                                                    M   D   R   K   D   L

DMV-D10    TATTTAATATAAATAAAGATGATTCGGAAGAATCTTTCGATCAGGAAACCACTGAAGTAA   180
ORF I         F   N   I   N   K   D   D   S   E   E   S   F   D   Q   E   T   T   E   V   I

DMV-D10    TAAACTTTAGTGAAACAAACAAAGGATTTGTATCAGACTACATGATTAATTCTGATTATT   240
ORF I         N   F   S   E   T   N   K   G   F   V   S   D   Y   M   I   N   S   D   Y   L

DMV-D10    TAGAACAAATCATGAAGCTTAAGCTAAAGCTTGATACAAAACAGGTTTTTAATCAACCTA   300
ORF I         E   Q   I   M   K   L   K   L   K   L   D   T   K   Q   V   F   N   Q   P   S

DMV-D10    GTAATATACAGAGATTAGTTTCAAAAGCTTTCTCTAGAAAAAATAATATCTTTTATTGCT   360
ORF I         N   I   Q   R   L   V   S   K   A   F   S   R   K   N   N   I   F   Y   C   F

DMV-D10    TTAATACTGAAGAATTGTCAGTAGATATAAAAGATACTACAGGTGAATTGTATTTACCAC   420
ORF I         N   T   E   E   L   S   V   D   I   K   D   T   T   G   E   L   Y   L   P   L

DMV-D10    TTCTAACAAAAGGAGAAATAGCCAGAAGACTTCTGACTATTAAACCAGAATTAAGAAAAA   480
ORF I         L   T   K   G   E   I   A   R   R   L   L   T   I   K   P   E   L   R   K   T

DMV-D10    CCATGAATATGGTGCACATCGGAGCAGTAAAAATCCTTCTGAAGGCACAGTTCAGAGATG   540
ORF I         M   N   M   V   H   I   G   A   V   K   I   L   L   K   A   Q   F   R   D   G

DMV-D10    GAATTAACTTCCCGATAAAAATGGCTTTAGTTGATAACAGAATTATCAACAGGCAAGACG   600
ORF I         I   N   F   P   I   K   M   A   L   V   D   N   R   I   I   N   R   Q   D   A

DMV-D10    CTCTACTCGGAGCAGTTCAAGGAAATTTAGCATACGGTAAATTTATGTTTACTGTTTATC   660
ORF I         L   L   G   A   V   Q   G   N   L   A   Y   G   K   F   M   F   T   V   Y   P

DMV-D10    CTAAATTTGCATTACATCGAGATTCAAAAGATTTCGATAAAACCTTAAGTTTCATACATC   720
ORF I         K   F   A   L   H   R   D   S   K   D   F   D   K   T   L   S   F   I   H   Q

DMV-D10    AGTGCGAAAGGACTGACCTCATGGAACCAGGTAACAAAGTATTTACGATTAATTATTTAA   780
ORF I         C   E   R   T   D   L   M   E   P   G   N   K   V   F   T   I   N   Y   L   I

DMV-D10    TTTCGTATGCTTTGACAAATAGTACTCATTCAATTGAGTATAAAGAAAAGGAGAGTATAA   840
ORF I         S   Y   A   L   T   N   S   T   H   S   I   E   Y   K   E   K   E   S   I   T

DMV-D10    CACTTGATGATGTATTCTCAGAAATAGGTACTGTCGAAGGAAGCAAGTTCGCTGAACCTT   900
ORF I         L   D   D   V   F   S   E   I   G   T   V   E   G   S   K   F   A   E   P   S

DMV-D10    CTCAGATACAAGAAAATTGGGCGATTGATATTGCTCGAGAGAAACAAACTCTAGGATTTC   960
ORF I         Q   I   Q   E   N   W   A   I   D   I   A   R   E   K   Q   T   L   G   F   Q

DMV-D10    AACCTAGAAATAGTTTTACAGGAATCCCCTTACAAATAGGCGAGTCCAGTAGAAACACAG  1020
ORF I         P   R   N   S   F   T   G   I   P   L   Q   I   G   E   S   S   R   N   T   E

DMV-D10    AAAAACAACTTATTCGGTCTATGTCCGAAAAAGTAGACATGTATGGAGAAATCCTTAGAG  1080
```

```
ORF I       K  Q  L  I  R  S  M  S  E  K  V  D  M  Y  G  E  I  L  R  E

DMV-D10  AATTGATTGGGAAATGACTTCGTCACAACCAATTGTTATGGCCTTAGCACGCCTAAATAG 1140
ORF I        L  I  G  K  *

DMV-D10  TAAAATCACTTATGAAGAATTAAGTATTGAACTTAACCTCATAAGGATAGAAGAAATAGA 1200

DMV-D10  AGAAAATTTTATGTTAAGTCGTAAAATTATATTCTTAAAAACTAGTTCTCTGAATTCTAA 1260

DMV-D10  ATTAGAAATCAGTAGGCAAAGAATTAAGGACCTCAAAGCTAGGAAAGTCTTAATTTCCAC 1320

DMV-D10  AGCATGAGTCTCTTAAAGCTAGAAGAATTAATTAATAAGTTAATTCTAAAATTAGACAGT 1380
ORF III       M  S  L  L  K  L  E  E  L  I  N  K  L  I  L  K  L  D  S

DMV-D10  TTAGTGTCTAACGAAAACCTAGAGGTTATAGCAGCAAAAATTATCAAGGAAGTCTTAGAC 1440
ORF III  L  V  S  N  E  N  L  E  V  I  A  A  K  I  I  K  E  V  L  D

DMV-D10  GAAATTGATAAGCTTAGCAAGCAACCTTTAGAAACCGAGAAGTCGGAAACCGAGAAATCG 1500
ORF III  E  I  D  K  L  S  K  Q  P  L  E  T  E  K  S  E  T  E  K  S

DMV-D10  GAAGCCAAAAATGAGAAAGAAGTTTTCAAAAATACTCTTTTCCAAACTTCAATGTAGGT 1560
ORF III  E  A  K  N  E  K  R  S  F  Q  K  Y  S  F  P  N  F  N  V  G
DMV-D10  AACCCTTCATTAGGAAGTTCAAAAAGTCCAAATGCTTTAACATGGCCTTTTGGATCAACC 1620
ORF III  N  P  S  L  G  S  S  K  S  P  N  A  L  T  W  P  F  G  S  T

DMV-D10  TCAAATGACTAATAGACGTCAAATGAGAATCCTTCTTAGGAAAATTAATTGGGAAATCCA 1680
ORF III  S  N  D  *

DMV-D10  AATGAATTATTGGTATTTGCAGATCAATAATCAAAGAAGAATCCAACTCCTCAGCGATTG 1740

DMV-D10  GAGCATTCACCCAGGTGATAAACAGTAAGTGAGCTAACTTATTTAGCAGCTCGCAGGTAC 1800

DMV-D10  AAACTCGAAAGAGAAATACAAGATCTTGAAGACCAAAAGGTAATATTATCTTGTGCATAA 1860

DMV-D10  AATGAGTTCTATCTTAGACAGATTAGACAAAATCTTTTTTGATCTCTTGGAAATCCGTCC 1920
ORF IV/V  M  S  S  I  L  D  R  L  D  K  I  F  F  D  L  L  E  I  R  P

DMV-D10  AATTGAACAAATAAAAGAACTCGTTCAACTCATGTCCTTAACAGAACAAGAGTGCGAACA 1980
ORF IV/V  I  E  Q  I  K  E  L  V  Q  L  M  S  L  T  E  Q  E  C  E  Q

DMV-D10  ACAATTTTCTGAAAAAAAAGATTTGAGAAATAAATATCTTTCCGATAAATATTTTTCTGA 2040
ORF IV/V  Q  F  S  E  K  K  D  L  R  N  K  Y  L  S  D  K  Y  F  S  E

DMV-D10  AGAAAAATTTTTAGAAAATAAAGAAAAGGTTTTTAAAATAAACAAACCTATAATTAAAGA 2100
ORF IV/V  E  K  F  L  E  N  K  E  K  V  F  K  I  N  K  P  I  I  K  E

DMV-D10  AGAAAACTCTTTAAAAGAAAAACACTATGCCAACGAATGTCCTAACAGACAACAGTCTCT 2160
ORF IV/V  E  N  S  L  K  E  K  H  Y  A  N  E  C  P  N  R  Q  Q  S  L

DMV-D10  GAAAAAGGTAAAATTAATTTCTGAGGTAGAAACGAAAGGATATTATCCTACAGAAAATAT 2220
ORF IV/V  K  K  V  K  L  I  S  E  V  E  T  K  G  Y  Y  P  T  E  N  I

DMV-D10  TTTTAATGATTATATGCAGGTATATTCATTAAAAGAACAGGAATCTTTGACTAATTCTGA 2280
ORF IV/V  F  N  D  Y  M  Q  V  Y  S  L  K  E  Q  E  S  L  T  N  S  D
```

24

```
DMV-D10    TAGTTGTTCAGACAGTTCTTCTTCCCAAAAATTTAATTTGAGGGAAACAATAGATAATGA 2340
ORF IV/V    S  C  S  D  S  F  S  S  Q  K  F  N  L  R  E  T  I  D  N  D

DMV-D10    TGAACAAGTTCATTGTTGTCAAAGCGAATTAGTTTATTCAAATATAACTAATCCAAACTC 2400
ORF IV/V    E  Q  V  H  C  C  Q  S  E  L  V  Y  S  N  I  T  N  P  N  S

DMV-D10    GATTTATATCAAAGGGAAAATATTTCTTAAAGGATATAAAACAATAACTCTACATTGCTA 2460
ORF IV/V    I  Y  I  K  G  K  I  F  L  K  G  Y  K  T  I  T  L  H  C  Y

DMV-D10    TGTAGATACAGGAGCAAGCCTTTGTCTTGCAAGTAAACACGTTATTCCTGAGGATCTATG 2520
ORF IV/V    V  D  T  G  A  S  L  C  L  A  S  K  H  V  I  P  E  D  L  W

DMV-D10    GGAAAACACTCCCAAAGATATAAAAGTAACAATTGCAAATCAGGAAACAATAAAATTAGA 2580
ORF IV/V    E  N  T  P  K  D  I  K  V  T  I  A  N  Q  E  T  I  K  L  D

DMV-D10    CAAGGTTTGTAGAAATCTTATTGTCTATTTTTCAGGGGAACCATTTCCTATTCCCACTAT 2640
ORF IV/V    K  V  C  R  N  L  I  V  Y  F  S  G  E  P  F  P  I  P  T  I

DMV-D10    CTACCAACAAGAAAGTGGTATAGATTTACTCCTTGGAAATAACTTTTGTCAATTATATGG 2700
ORF IV/V    Y  Q  Q  E  S  G  I  D  L  L  L  G  N  N  F  C  Q  L  Y  G

DMV-D10    GCCATTCACCCAATGGATTGACAGAATATCTTTTCATCTTAATCAAGAAATGATTTTGAT 2760
ORF IV/V    P  F  T  Q  W  I  D  R  I  S  F  H  L  N  Q  E  **M**  I  L  I

DMV-D10    CAAAAAGATAACAAAAGCTTTTCAAATTGGAAAAACTGGTTTTCTTGAATCCATGAAAAA 2820
ORF IV/V    K  K  I  T  K  A  F  Q  I  G  K  T  G  F  L  E  S  **M**  K  K

DMV-D10    AGATTCAAAAACTAAACAAATTCCGGGTACCAATATAACCCAAGAAGTTATAAAACCGGA 2880
ORF IV/V    D  S  K  T  K  Q  I  P  G  T  N  I  T  Q  E  V  I  K  P  E

DMV-D10    AAGATTTTTTCTTGAAATACAGAGATATCAAAAAATTGAAGATCTACTTGAAAAAGTATG 2940
ORF IV/V    R  F  F  L  E  I  Q  R  Y  Q  K  I  E  D  L  L  E  K  V  C

DMV-D10    TTTTGAAAACCCTATTGACCCAGAAAAGTCCAAATACTGGATGAATGCCTCAATAGAATT 3000
ORF IV/V    F  E  N  P  I  D  P  E  K  S  K  Y  W  **M**  N  A  S  I  E  L

DMV-D10    AATAGATCCCAAAACCGTGGTTAGGGTAAAACCCATGAAATATAGTCCACAGGATCGCGA 3060
ORF IV/V    I  D  P  K  T  V  V  R  V  K  P  **M**  K  Y  S  P  Q  D  R  E

DMV-D10    AGAATTTGGTAAGCAAATCAAAGAATTACTTGATTTAAAATTAATTATTCCAAGTAAATC 3120
ORF IV/V    E  F  G  K  Q  I  K  E  L  L  D  L  K  L  I  I  P  S  K  S

DMV-D10    TCCTCACATGTCTCCAGCTTTTCTTGTAGAAAACGAAGCAGAAAAAAGGAGAGGAAAAAA 3180
ORF IV/V    P  H  **M**  S  P  A  F  L  V  E  N  E  A  E  K  R  R  G  K  K

DMV-D10    GCGCAGGGTTGTAAATTATAAGGCAATTAATGCAGCTACCAAAGGAGACAGCCATAATCT 3240
ORF IV/V    R  R  V  V  N  Y  K  A  I  N  A  A  T  K  G  D  S  H  N  L

DMV-D10    TCCTTGTATGCAGGAATTATTAACCCTCTTAAGAGGAAAAACCATTTTTTCTACTTTCGA 3300
ORF IV/V    P  C  **M**  Q  E  L  L  T  L  L  R  G  K  T  I  F  S  T  F  D

DMV-D10    CTGCAAGAGTGGTTTTTGGCAGGTATTGCTAAATGAAGAATCTCAGTTATTAACTGCGTT 3360
ORF IV/V    C  K  S  G  F  W  Q  V  L  L  N  E  E  S  Q  L  L  T  A  F
```

25

```
DMV-D10    CACATGTCCAGACGGTCATTACCAATGGAGGGTAGTCCCTTTTGGATTAAAACAGGCACC 3420
ORF IV/V    T  C  P  D  G  H  Y  Q  W  R  V  V  P  F  G  L  K  Q  A  P

DMV-D10    AAGCATATTTCAACGACATATGCAAAATGCCCTAAGAGGATTAGAAAATTATTGTACCGT 3480
ORF IV/V    S  I  F  Q  R  H  M  Q  N  A  L  R  G  L  E  N  Y  C  T  V

DMV-D10    CTACGTTGACGATATCATTGTTTTCTCAGACTCTGAAGAAAAGCATTATTTTCATGTTTT 3540
ORF IV/V    Y  V  D  D  I  I  V  F  S  D  S  E  E  K  H  Y  F  H  V  L

DMV-D10    ATCAGCTTTAAAAACGATTGAAAAATATGGAATTATTCTATCAAAAAAGAAAGCAAATCT 3600
ORF IV/V    S  A  L  K  T  I  E  K  Y  G  I  I  L  S  K  K  K  A  N  L

DMV-D10    TTTTAAAACAAAGATAAACTTTTTAGGTCTTGAAATAGACCAAGGGACTCATTGTCCACA 3660
ORF IV/V    F  K  T  K  I  N  F  L  G  L  E  I  D  Q  G  T  H  C  P  Q

DMV-D10    AAAACATATCTTGGAAAATTTACATAAATTTCCAGATACTTTAGAAGATAAAAAACATCT 3720
ORF IV/V    K  H  I  L  E  N  L  H  K  F  P  D  T  L  E  D  K  K  H  L

DMV-D10    TCAAAGATTTTTAGGTGTTCTAACCTATGCAGAAAGCTACATTCCCAAGCTTGCGGAGCT 3780
ORF IV/V    Q  R  F  L  G  V  L  T  Y  A  E  S  Y  I  P  K  L  A  E  L

DMV-D10    CAGGAAACCCTTACAGGTAAAACTAAAAAAGGATTACGTATGGGAATGGAAACAATCCGA 3840
ORF IV/V    R  K  P  L  Q  V  K  L  K  K  D  Y  V  W  E  W  K  Q  S  D

DMV-D10    TACGAACTATATTAAAAAGATAAAGAAAAATCTAACGAGTTTTCCAAAACTTTATCTTCC 3900
ORF IV/V    T  N  Y  I  K  K  I  K  K  N  L  T  S  F  P  K  L  Y  L  P

DMV-D10    AAAAGAAAAAGAATTTCTTATTATAGAAACAGATGCTTCCAACGATTTTTGGGGAGGCGT 3960
ORF IV/V    K  E  K  E  F  L  I  I  E  T  D  A  S  N  D  F  W  G  G  V

DMV-D10    TCTAAAGGCAAAAACTGCTGACAAAGAAGAAGTATGCCGATACACTTCTGGTAGTTTTAA 4020
ORF IV/V    L  K  A  K  T  A  D  K  E  E  V  C  R  Y  T  S  G  S  F  K

DMV-D10    AACAGCCGAAAGGAATTACCACAGTAATGAAAAGAATTACTTGCGGTAAAAAACACTAT 4080
ORF IV/V    T  A  E  R  N  Y  H  S  N  E  K  E  L  L  A  V  K  N  T  I

DMV-D10    TTCCAAATTCTCCATTTATCTTACTCCGGTAAAATTCTTGGTTAGAACGGATAATAGAAA 4140
ORF IV/V    S  K  F  S  I  Y  L  T  P  V  K  F  L  V  R  T  D  N  R  N

DMV-D10    CTTCACTTATTTCCTCAAAACAAAAATTTCAGGAGATAACAAACAAGGTCGTCTTGTCAG 4200
ORF IV/V    F  T  Y  F  L  K  T  K  I  S  G  D  N  K  Q  G  R  L  V  R

DMV-D10    ATGGCAAATGTGGTTCTCCAGATACTCTTTTGATATTGAGCATCTTGAAGGATCAAAAAA 4260
ORF IV/V    W  Q  M  W  F  S  R  Y  S  F  D  I  E  H  L  E  G  S  K  N

DMV-D10    TGTGTTGGCTGACTGCCTACCACGCGACTTTCAAAATTAAAAAATTTTCTCTTCAGTTTA 4320
ORF IV/V    V  L  A  D  C  L  P  R  D  F  Q  N  *

DMV-D10    TCTCTTTCTTTCAAAAAACATTTTTTTTTAAGAAATAAGAGTTATGGAAGAAATTAAGGCG 4380
ORF VI                                               M  E  E  I  K  A

DMV-D10    TTACGCCTAAAGGAAAAAATTCTTTTAATAGAATTAAAAGCTATTCAGCAAAAAATAGCA 4440
ORF VI      L  R  L  K  E  K  I  L  L  I  E  L  K  A  I  Q  Q  K  I  A
```

26

```
DMV-D10   CTCTACGACGAAAGTCTAGAAACAACTGTCAATTCAGTACAGGAAACCGCGAGGTTACAA 4500
ORF VI    L  Y  D  E  S  L  E  T  T  V  N  S  V  Q  E  T  A  R  L  Q

DMV-D10   ACTGAATCTATTCCTTTACAAACGGCTAAAGGTAAAGAAATATCAAGCCCGTTGATCCCT 4560
ORF VI    T  E  S  I  P  L  Q  T  A  K  G  K  E  I  S  S  P  L  I  P

DMV-D10   GATGCTTTGGGAAAAAGCATGAAAAGGTCAAATGATGAACCACAATCATCATCTACCCCT 4620
ORF VI    D  A  L  G  K  S  **M**  K  R  S  N  D  E  P  Q  S  S  S  T  P

DMV-D10   GATGCTTTGGGAAAAAGCATGAAATGGTCAAATGATGAACCACAATCATCATCTAGTCCG 4680
ORF VI    D  A  L  G  K  S  **M**  K  W  S  N  D  E  P  Q  S  S  S  P

DMV-D10   GTTGCAAACGGCTCCGGTAAAGACTCATCAAATCCGTTGATGGCTGACAGTTTGCCAAAA 4740
ORF VI    V  A  N  G  S  G  K  D  S  S  N  P  L  **M**  A  D  S  L  P  K

DMV-D10   ACTGATAAGGTCGTTCTTAGACCAACTTATTCTCAAATCATGCAGGAACCCATGAAATCT 4800
ORF VI    T  D  K  V  V  L  R  P  T  Y  S  Q  I  **M**  Q  E  P  **M**  K  S

DMV-D10   AAATCAAGATTTTATGTGATCTTCGATGGAGAACACAGAGGCATTTATGAAGACTGGTCC 4860
ORF VI    K  S  R  F  Y  V  I  F  D  G  E  H  R  G  I  Y  E  D  W  S

DMV-D10   ATAGTCTCCAGATATGTCACAGGAACGTCTTGTCCGTTCAAGAAATTTGGGTCCTTGTTA 4920
ORF VI    I  V  S  R  Y  V  T  G  T  S  C  P  F  K  K  F  G  S  L  L

DMV-D10   CAAGCCCAACAGGAAGCTACTCGGTACTCAGCAGAGTTCGGGAAAAAAGAAATACCACTT 4980
ORF VI    Q  A  Q  Q  E  A  T  R  Y  S  A  E  F  G  K  K  E  I  P  L

DMV-D10   AGGATTACGGTTTTCTCTGAACCTCTAAAGCCTAAGAAAAAAGAAAGAATGGTAGAATTC 5040
ORF VI    R  I  T  V  F  S  E  P  L  K  P  K  K  K  E  R  **M**  V  E  F

DMV-D10   AAAAAAGAAAAAGACTTCTTAAAAAATATTAAAGAAGCTAAAGAAGAAGAAGAAGATAAT 5100
ORF VI    K  K  E  K  D  F  L  K  N  I  K  E  A  K  E  E  E  E  D  N

DMV-D10   GTTATTATCTCTTTAGACGAATTCAGACTAGTTTGGTCAAAAGCCAGATTATTAAGTCAT 5160
ORF VI    V  I  I  S  L  D  E  F  R  L  V  W  S  K  A  R  L  L  S  H

DMV-D10   CTCGACTTTGAGTCTGAACACATATACATAGAAGATAAAGCTACTAAAAGTTTAATTATC 5220
ORF VI    L  D  F  E  S  E  H  I  Y  I  E  D  K  A  T  K  S  L  I  I

DMV-D10   TTTTGTCCAGGTGCAAGTCCGGAATTAGTTTCTCTGGCATTTTATGTTGGTTTAACTAAA 5280
ORF VI    F  C  P  G  A  S  P  E  L  V  S  L  A  F  Y  V  G  L  T  K

DMV-D10   TACATATATCCTTCTGCTAATTTACTTGAAATTAGTCTGTTATCCGAAGGAATGAAGAAA 5340
ORF VI    Y  I  Y  P  S  A  N  L  L  E  I  S  L  L  S  E  G  **M**  K  K

DMV-D10   GCTATTAAAAACTTTCGAAGAAAGATTGCTGATGCTAGAGATGCAAATATTTTCATTAAA 5400
ORF VI    A  I  K  N  F  R  R  K  I  A  D  A  R  D  A  N  I  F  I  K

DMV-D10   TGCAACTCTACTCTACCAGATTGGTACCAAGGAAGATCATTTCCAAGCTACCACCATTTG 5460
ORF VI    C  N  S  T  L  P  D  W  Y  Q  G  R  S  F  P  S  Y  H  H  L

DMV-D10   GAGATTGGCATCGCCAAAACAAGAACAGTCGAACCGTCCAAGGTAATGGCTGAAAAATAT 5520
ORF VI    E  I  G  I  A  K  T  R  T  V  E  P  S  K  V  **M**  A  E  K  Y
```

27

```
DMV-D10   GAAGACCCTGAAAATTGGGTGGAAATAAGGTCAAGAGGATTCATGAAGATTCTCAACAAC 5580
ORF VI     E  D  P  E  N  W  V  E  I  R  S  R  G  F  M  K  I  L  N  N

DMV-D10   CTACAAAAAATAACAGCTGGAAGCTTTGTTAAAGTCAATTATTGCAGCACCAACTGCATA 5640
ORF VI     L  Q  K  I  T  A  G  S  F  V  K  V  N  Y  C  S  T  N  C  I

DMV-D10   ATTACCGGCCACAGTGGATCTCCTATTCCCAAGCACGAAAAGGAGTTACTGACAATAATG 5700
ORF VI     I  T  G  H  S  G  S  P  I  P  K  H  E  K  E  L  L  T  I  M

DMV-D10   GAGCGGCGAATTCTTCTCAATGAGACAGAAGTTGGCCCTACCACCAAAGAAAATCTTTGC 5760
ORF VI     E  R  R  I  L  L  N  E  T  E  V  G  P  T  T  K  E  N  L  C

DMV-D10   CAAAAGTTGCAGAGAATAATTGAAGATCATCAGTGCGCTCACTGTAAAGCATCTTCAACA 5820
ORF VI     Q  K  L  Q  R  I  I  E  D  H  Q  C  A  H  C  K  A  S  S  T

DMV-D10   GAAGAAACGTCTCTATCCAAAGACTGCTTTATGGATGAAGACAAATGAAAAGAGAACGTT 5880
ORF VI     E  E  T  S  L  S  K  D  C  F  M  D  E  D  K  *

DMV-D10   CAAAAATAACGGCGTCTACTTTTCAAAAAGAAGATAGAAGACAGCTCACCCTTCACATGG 5940

DMV-D10   GTATGTCGACTAAATACGTCAGGGATGACGAATTAGTAAGGGCTGTTATGGTCTTTTGTA 6000

DMV-D10   TAGATCGACTCCTAGTATATAAAGGAGTTAGATTTTCAATTTTAGATCATCGATCATCTA 6060

DMV-D10   GCCTAGAAAATACACTAGAAAAAATAAACCTCCCTCTCTTGTAAACACACAAATGTGTGA 6120

DMV-D10   GTAAACCCCTACCTTAAGGGGAGATATGTAAATTATCAAGAAGTGCTCCTCTTCGGAGGA 6180

DMV-D10   GAACAGTTATAATAAAATATTTCCTTTTTGTCTACCTCCTTCTATTCACTACCGTAGACG 6240

DMV-D10   TTAAACCTCCTTCATATGTCTACCCCCTCTTCACTAACGCAAAAGTTTAATCTTCTTGTC 6300

DMV-D10   TACCCCCTCTTCACTAACGCAAACGTTGTTCTGTAAAACCTGGAGCCCTGGAGCCGTCAC 6360

DMV-D10   GGAGTACCGCTTTGGCAGGTGGCCGTTTAAGGGAAACGAGTGACTAGGTTGGGTTGGTCC 6420

DMV-D10   CTTGCAGAACAGATAACGAATGCACAGCTTGATAGAGGACGATGAAGATTTTTGCACAAG 6480

DMV-D10   ATTGATAGAGGACGACGAAGGTAACACGACGGATGCACAGACTGATAAAGGACGACAATG 6540

DMV-D10   GTATTTTATTTAAATACTGTTAAAACACTTCAGGAGATTTACAGATTCTCTCCTTAAGAG 6600

DMV-D10   TTAAAAAACTCATAAGTTTGTGAACAAACAAGAAATTAAAAGGTTATTTTTATAATCTAG 6660

DMV-D10   AAATGACGATCAAAAGTAGATTATTCTACTTAAAAAATGATTTTTAAAATCATAGAAATT 6720

DMV-D10   CAAACGATTTTTATCGTTTACATGATTCATAAGAATCAAACAAAAACATGGTTGTTTTGA 6780

DMV-D10   ACCATTAAAAACAATGATTTTTTATAATCAATAAATTGATTTTCATAATCAATTAAAGAG 6840

DMV-D10   ATTTTATAATCTAAAGTATGACTATCCATACTTTACAACCCCTGCAAAGGGAAAAGTCTA 6900
ORF VII                            M  T  I  H  T  L  Q  P  L  Q  R  E  K  S  N

                                    28
```

```
DMV-D10    ACCGTAAAGGTTTGATAACCCGTACTCCGGGATTAAAGAGTCAAGATTATAATTTTATTA 6920
ORF VII     R   K   G   L   I   T   R   T   P   G   L   K   S   Q   D   Y   N   F   I   K

DMV-D10    AGTTAAAATTATTCCTGGACCACAAAATAAAAATGATTTTACAATCTGAAAAAGAGATTT 7020
ORF VII     L   K   L   F   L   D   H   K   I   K   **M**   I   L   Q   S   E   K   E   I   F

DMV-D10    TTATAGTCTCAGACCTCATCCGCAAT 7046
ORF VII     I   V   S   D   L   I   R   N
```

Figure 4. Complete nucleotide sequence of *Dahlia mosaic virus* – D10 showing the locations of

the open reading frames ORF I, III – VII. The numbers refer to the nucleotide positions.

a) ORF I
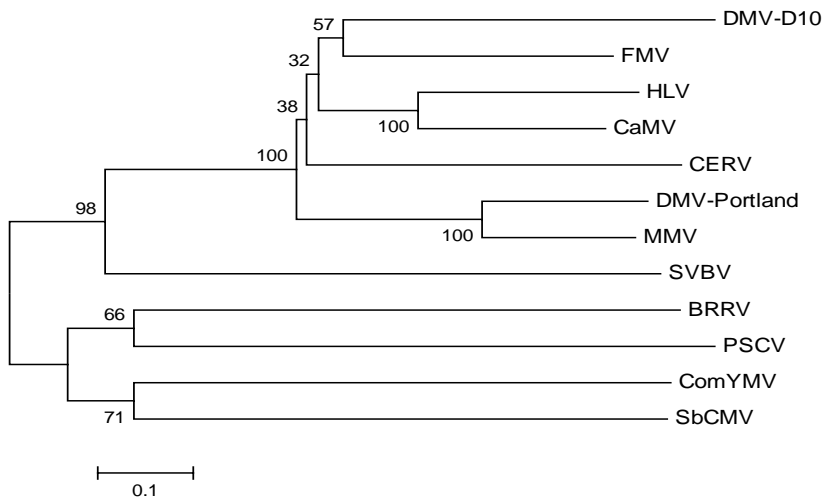
b) ORF III
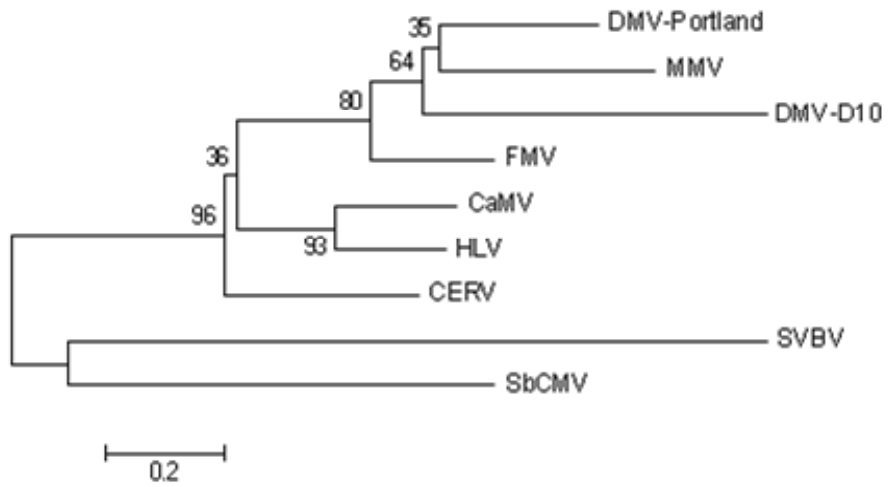
c) ORF IV
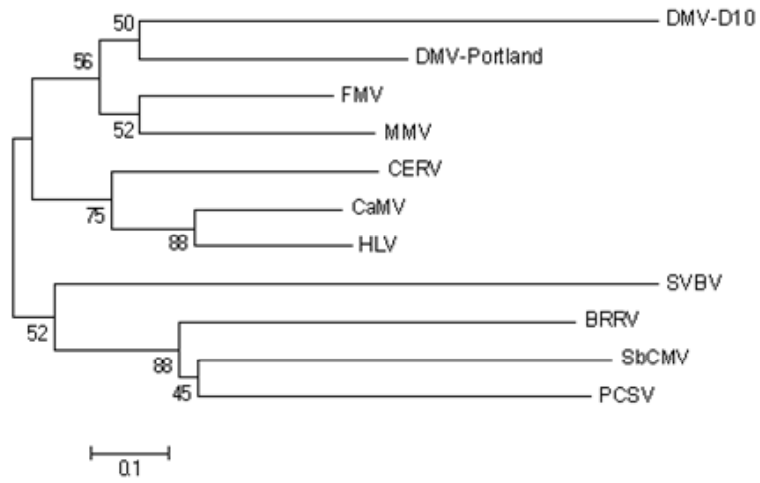
d) ORF V
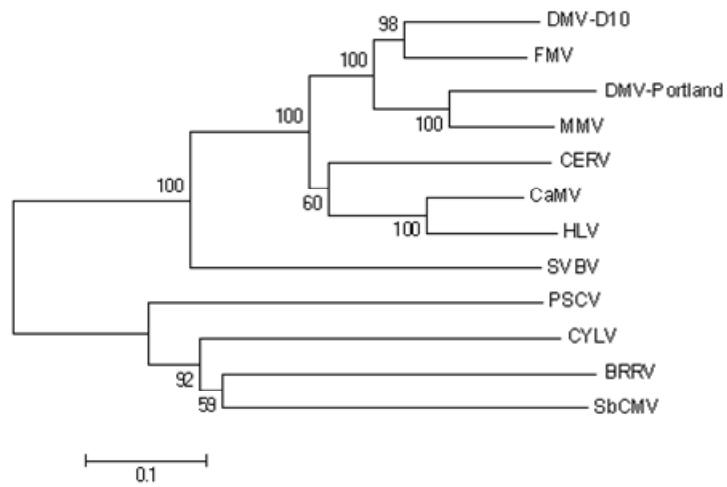
e) ORF VI
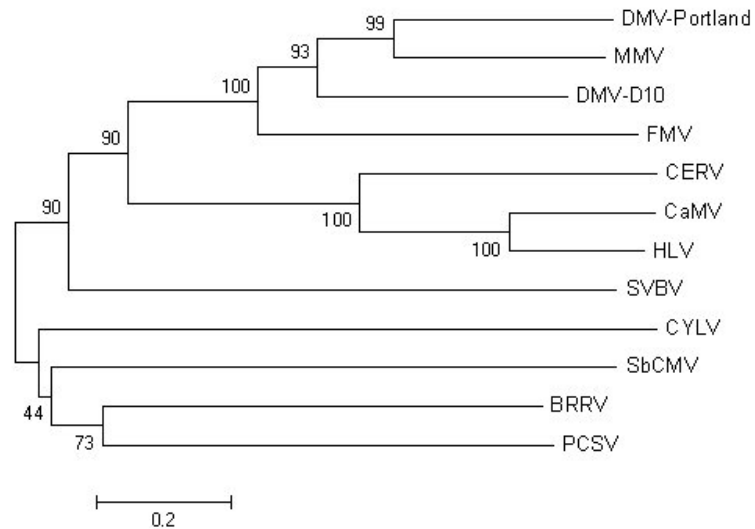
Figure 5. Phylograms drawn from Clustal W alignments of the different ORFs of selected members of the family *Caulimoviridae*. BRRV-*Blueberry red ringspot virus*, CaMV-*Cauliflower mosaic virus*, CERV-*Carnation etched ring virus*, CYLV- *Cestrum yellow leaf curling virus*, ComYMV- *Commelina yellow mottle virus*, DMV-*Dahlia mosaic virus*, 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 1000 iterations.

The ORF II of caulimoviruses encodes an aphid transmission factor (ATF).   A

32

homologous ORF was not identified in DMV-D10. A smaller ORF found in the DMV-D10 genome did not show any significant homology with those available in GenBank. Previous studies on caulimovirus ATF have identified amino acid sequence motif IXG, X being any amino acid, to be necessary for the interaction between the ATF and virus particles for aphid transmission (Schmidt *et al*., 1994). The IXG motif is present in the ATF of DMV-Portland, CaMV and *Carnation etched ring virus* (CERV) but absent in DMV-D10 ORF II.

ORF III of DMV-D10 potentially encodes a DNA binding protein (DNAb). The DNAb has a C-terminal proline rich domain that is considered to be necessary for its non-sequence-specific nucleic acid-binding activity (Fig. 6) (Jacquot *et al*., 1998; Mougeot *et al*., 1993). This ORF contains the coiled coil structure that is conserved among caulimoviruses at the N-terminus and was shown in other caulimoviruses to assemble as a tetramer to form a functional protein *in planta* (Stavolone *et al*., 2001). Amino acid sequence alignment of the N terminal domains with the predicted coiled coil region of DMV-D10 and other caulimoviruses is shown in Figure 7. Phylogenetic analysis of the DNAb of DMV-D10 DNAb formed a cluster with those of DMV-Portland, MMV and FMV (Fig 5b).

ORFs IV and V of other caulimoviruses occur as a single ORF in DMV-D10 due to the absence of a stop codon. The ORF IV of DMV-D10 consists of 100 aa representing a truncated viral coat protein (CP). DMV-D10 does not have a conserved motif identified in other caulimovirus CPs which include a RNA-binding domain that is consistent with a cysteine motif or "zinc finger" ($CX_2CX_4HX_4C$) (Hasegawa *et al*., 1989). Only part of this conserved motif (HX_4C) could be seen in the DMV-D10 putative ORF IV as HYANECP (aa 88-94). Furthermore, DMV-D10 CP also lacked a lysine rich core upstream of the cysteine motif that

was found in other caulimoviruses.

DMV-D10	P	N F N V G W P S L G S S K S	P	N A L T	W P
DMV-Portland	P	N W N V G N A S L G S S G N	P	D A L K	W P
CaMV	P	K A L T - W P F K A - - - -	P	- A - G	W P
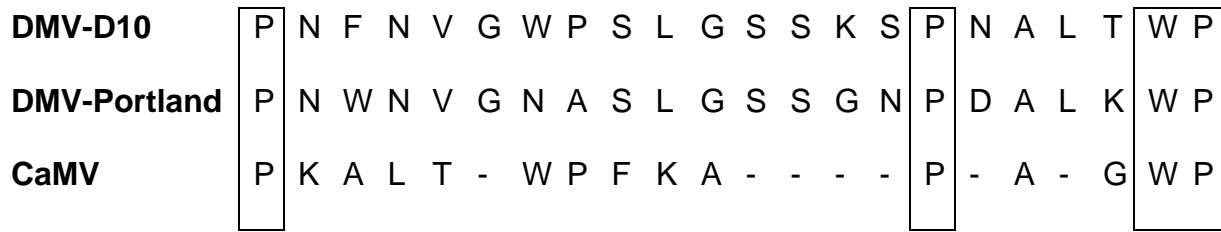
Figure 6. Alignment of C-terminal domains of *Dahlia mosaic virus* DMV ORF III with *Cauliflower mosaic virus* CaMV ORF III proline rich motif. The C-terminal domain contains serine, proline and basic residues, characteristic of the proline-rich DNA binding proteins. Identical residues are blocked.

```
                    a b c d e f g a b c d e f g a b c d e f g a b ---a b c d e f g a
DMV-D10           MS L L K L E E L I N K L I L K L D S L V S N E N ---L E V I A A K I
DMV-Portland      MT T I K D L A A L I A G L R E E I S K L K A E I T ---L E S L G A K I
CaMV              MA N L N Q I Q K E V S E I L S D Q K S M K S D I K ---L E T V A A K I
CERV              MN L A T I A S E I E V V K T N Q K T I E S K I D
FMV               MA A T L S E I R E L I Q S L T K T A N E I K A M L E
```

Figure 7. Alignment of N-terminal domains of ORF III of caulimoviruses. The seven positions of the heptad repeat are labeled a through g above the alignment. Positions a and d bold consist primarily of hydrophobic amino acids that form the helix interface, while the others are polar

34

residues.

Percentage amino acid sequence identity of DMV- D10 CP with those of other caulimovirus CPs was low (5.8-7.7%) and the DMV-D10 formed a clade with that of DMV-Portland (Fig 5c).

ORF V encodes a polyprotein containing all the motifs conserved in caulimovirus replicases. These motifs include aspartic protease, reverse transcriptase (RT) and ribonuclease H (RNase H). The amino acid sequence YVDTGASLC (aa 200-208) in DMV-D10 was similar to the putative protease domain reported in other caulimoviruses (Glasheen *et al*., 2002; Hasegawa *et al*., 1989; Torruella *et al*., 1989). A conserved amino acid sequence representing the RT domain of caulimoviruses was present in DMV-D10 as YVDDIIVF (aa 541-548) (Glasheen *et al*., 2002; Hasegawa *et al*., 1989). Amino acid sequences IIETDASNDFWG (aa 687-698), EEVCRYTSGSFK (aa 709-720) and EHLEGSKNVLADCL (aa 793-806) in DMV-D10 are homologous to the RNase H motifs reported in other caulimoviruses (Torruella *et al*., 1989). The amino acid sequence identity of DMV-D10 with other caulimovirus polyproteins ranged from 40% to 70%. The cladogram based on the protein sequences showed clustering of DMV-D10 with FMV while DMV-Portland formed a cluster with MMV (Fig. 5d).

The ORF-VI of DMV-D10 encodes the viral inclusion body protein (IB). Cluster relationships among the IBs of caulimoviruses showed clustering of DMV-Portland and MMV while DMV-D10 did not form a single cluster with any of the representative caulimoviruses (Figure 5e). An amino acid sequence motif, GLTKYIY, that is well conserved in other caulimoviruses was also seen in DMV-D10 (aa 303-309) (Glasheen *et al*., 2002; Hasegawa *et al*., 1989).

The ORF-VII of DMV-D10 encodes a 73 aa protein following a 990 nucleotide

intergenic region between ORF VI and ORF VII. A similar ORF is present in all known

caulimoviruses but functional proteins have not been found in infected plants (Wurch *et al*.,

1990). Furthermore, sequence comparison of this region of DMV-D10 with the corresponding

region of known caulimoviruses did not show any significant homologies.

The large intergenic region between ORFs VI and VII of DMV-D10 contains a putative

35S promoter homolog. Sequence homology between the intergenic regions of DMV-D10 and

other caulimoviruses is shown in Fig 8. The TATA box as well as the conserved residues, TC,

around the TATA box were identified (Fig. 8). The TGACG motif, present upstream of histone

genes and the CaMV 35S promoter (Fang *et al*., 1989) was also detected in DMV-D10.

However, a CAAT box that was found upstream of the TATA box in the 35S promoter of CaMV

was not found in the DMV-D10 sequence. Futhermore, the CaMV 35S RNA leader sequence is

reported to form a large stem-loop structure (Fig. 9a) (Fütterer *et al*., 1988; Hemmings-

Mieszczak, 1997). Formation of a similar stem-loop structure was detected in the DMV-D10

intergenic region homologous to the CaMV 35S RNA leader sequence (Fig. 9b).

DMV-D10   T C G A C - - - T C C T - A G T A T A T A A A G G A G T

CaMV   A A G A C C C T T C C T - - C T A T A T A A - G G A A G

FMV   A A G A A T - T C C C T - - C T A T A T A A - - G A A G

CERV   A G A T C C T C C C T G G C T A T A T A A A G G G A G

Figure 8. The TATA box in DMV-D10 and selected caulimoviruses.

a)

```
                              ___
                             /   \
                            |     |
                             \___/
                            . | |
                            G-C
                            A-U
                            A-U          STEM 1a
                            U   G
                            A-U
                            A-U
                  GUUGAGC ── A-UU
                U                  A      BL1a
                  GUACUCGCUU    U
                            U-A
                            G-C
                            G-C
                            G-C
                            A-U          STEM 1b
                            U-A
                            A-U
                            UUA
                          C     U        BL1b
                          UU    C
                            G-C
                            G-C
                            G-C          STEM 1c
                            A-U
                            U-A
                            U-A
                            A-UA
                          A     A        BL 1c
                            G-CC
                            G-C
                            G-C          STEM 1d
                            A-U
                            A-U
                            U-A
   7487                    A-U   7993
    |                      G-C    |
5'..AGAAUA AUG UGU GAG UAG UUUCCC      AAUA ..3'
         Short open reading frame    /    \
                          Shunt Donor    Shunt Acceptor
                            (7510)          (7989)
                            ─────────────────►
                              Ribosome Shunt
```

37

b)

```
                                    ___
                                   /   \
                                  |     |
                                   \___/
                                   |   |
                              ⠂⠄    U-A
                                    G-C
                                    A-U    STEM 1a
                                    C-G
                                    A-U
           G GAGGAG——— A-U A
           C                       A   BL1a
            U                       A
             U CUCCUC     G-C A
                          U-A
                          G-C    STEM 1b
                          A-U
                          A-U
                          G-C
                          A   A    BL1b
                          A   G
                          C-G
                          U-A
                           A
                          U-G
                          U-A
                          A-U
                          A-U    STEM 1c
                          A-U
                          U-A
                          G-C
                          U-AG
                          A   A   BL 1c
                          U   U
                          A-U
                          G-C
                          A-U
                          G-C
                          G-U    STEM 1d
                          G-C
                          G-C
                          A-U
                          A-U
   6106                   U-A
5'..ACACAA AUG UGU GAG UAA ACCCCUACC     GAGUUA ..3'
       Short Open Reading Frame      /       \
                              Shunt donor   Shunt acceptor
                                (6132)          (6596)
```

38

Figure 9. Stem- loop structures that are formed by the leader sequences of the RNA. a) Stem-loop within the leader sequence of the *Cauliflower mosaic virus* 35S RNA (Shabibi *et al.*, 2006). b) Predicted stem-loop within the leader sequence of the *Dahlia mosaic virus* – D10 large RNA. BL- bulge loop.

DISCUSSION

The genome organization of DMV-D10 suggested that it belongs to the genus *Caulimovirus*. Sequence comparisons and phylogenetic analyses revealed that it is a distinct species within the genus *Caulimovirus*. DMV-D10 was found to be the most prevalent DMV sequence in dahlias in the US (Pahalawatta *et al*., 2007b; Pappu *et al*., 2005). While the overall genome organization of DMV-D10 conformed to that of a species in genus *Caulimovirus*, major differences are the smaller CP and the absence of the gene coding for the ATF. Despite extensive sequencing of numerous DMV-D10 isolates from several parts of the country, the gene coding for the ATF could not be found in the genome. Absence of the gene coding for the ATF could be explained by the fact that DMV-D10 is seed transmitted at a very high frequency (Pahalawatta *et al*., 2007a). Moreover, dahlia is primarily vegetatively propagated and the loss of ATF may be an indication of the influence of cropping procedures on the evolution and selection of DMV. A similar loss of aphid transmission after consecutive mechanical inoculations was reported in *Cucumber mosaic virus* (CMV) (Ng & Perry, 1999). A study by Leclerc et al. (2001) showed that the C-terminal half of the CaMV CP is important for virus transmission through its interaction with the C-terminal of the ATF. Thus the importance of the C-terminal half of the CP is in the context of the involvement of the ATF in aphid transmission. In contrast, the N-terminus of the CP was reported to be critical for viral infection in CaMV (Leclerc *et al*., 1999). Further studies need to be carried out to determine whether this is a possible explanation for the presence of only the amino terminal part of the CP in DMV-D10. Whether this unique organization is a result of viral evolution brought about by continuous vegetative propagation of dahlias is of particular interest.

Studies on CaMV suggest that nucleic acid binding activity is the major function of the ORF III protein product and that it is related to aphid transmission via its interaction with ATF (Leh *et al*., 1999). However, other reports have shown that the ORF III protein product is essential for infectivity within the plant host (Jacquot *et al*., 1999). Therefore, it is possible that in DMV-D10, where ATF appears to be absent, the ORF III product potentially plays a role in virus infectivity in dahlia.

In CaMV, the 35S RNA that is formed when the host RNA polymerase II initiates transcription, functions as the replication intermediate that serves as the template for reverse transcription (Pfeiffer & Hohn, 1983). The approximately 604 bp leader sequence contains nearly 9 short open reading frames (sORFs) that range in size from 9-102 bp (Shababi *et al*., 2006). These sORFs in the CaMV 35S RNA are considered to be a hindrance to translation in the plant host. Eukaryotic ribosomes generally initiate translation at the first AUG at the 5' end of the transcript that is in a favorable context (Kozak, 1999). Reinitiation of transcription of a second cistron on the same transcript by the eukaryotic ribosome does not occur. In the CaMV 35S RNA leader sequence the sORFs are bypassed by the formation of a large stem-loop structure. This translation strategy of CaMV is referred to as the 'ribosomal shunt' (Ryabova *et al*., 2002). The stem-loop structure formed by the 35S RNA leader sequence functions by bringing together the shunt donor and acceptor, thereby facilitating the ribosome to bypass the central region of the leader sequence and reach the start codon of ORF VII. Formation of a similar stem-loop structure in the DMV-D10 intergenic region is an indication that the ribosomal shunt mechanism of translation is utilized by the virus.

DMV-D10 appears to be closely related to DMV-Portland and FMV when considering

41

the phylogenetic relationships based on sequence relationships of ORFs I, IV and V. Overall, the complete molecular characterization of the DMV-D10 genome provides evidence of the existence of at least two distinct caulimoviruses that are associated with dahlia mosaic and highlights the need for developing management options for these viruses.

REFERENCES

Albouy, J., Lemattre, M., Wang, W. C ., Mevel, A. 1992 . Production of pathogen tested dahlia in France. Acta Hort. 325: 781-784.

Altschul, F., Thomas L. M., Alejandro A. S., Jinghui, Z., Zheng, Z., M., W ., Lipman, D. J. 1997 . Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402.

Benson, D. A., Karsch-Mizrach, I., Lipman, D. J., Ostell, J ., Wheeler, D. L. 2005 . GenBank. Nucleic Acids Res. 33: D34-38.

Brierley, P ., Smith, F. F. 1950 . Some vectors, hosts and properties of dahlia mosaic virus. Plant Dis. Rep. 34: 363-370.

Brunt, A. A. 1971 . Some hosts and properties of dahlia mosaic virus. Annal App. Biol. 67: 357-368.

Fang, R. X., Nagy, F., Sivasubramaniam, S ., Chua, N. H. 1989 . Multiple cis regulatory elements for maximal expression of the Cauliflower mosaic virus 35S promoter in transgenic plants. Plant Cell 1: 141-150.

Fütterer, J., Gordon, K., Bonneville, J. M., Sanfaçon, H., Pisan, B., Penswick, J. R., Hohn, T. 1988. The leading sequence of *Cauliflower mosaic virus* large RNA can be folded into a large stem-loop structure. Nucleic Acids Res. 16:8377-8390.

Glasheen, B. M., Polashock, J. J., Lawrence, D. M., Gillet, J. M., Ramsdell, D. C., Vorsa, N .,

Hillman, B. I. 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.

Hasegawa, A., Verver, J., Shimada, A., Saito, M., Goldbach, R., Van Kammen, A., Miki, K.,
Kameya-Iwaki , 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.

Hemmings-Mieszczak, M., Steger, G., Hohn, T. 1997. Alternative structures of the *Cauliflower
mosaic virus* 35S RNA leader: implications for viral expression and replication. J. Mol.
Biol. 267: 1075-1088.

Jacquot, E., Geldreich, A., Keller, M ., Yot, P. 1998 . Mapping regions of the *Cauliflower
mosaic virus* ORF III product required for infectivity. Virology 242: 395-402.

Kozak, M. 1999. Initiation of translation in prokaryotes and eukaryotes. Gene 234: 187-208.

Kumar, S., Tamura, K ., Nei, M. 2004 . MEGA3: Integrated software for molecular evolutionary
genetics analysis and sequence alignment. Brief Bioinform. 5: 150-163.

Leclerc, D., Chapdelaine, Y ., Hohn, T. 1999 . Nuclear targeting of the cauliflower mosaic virus
coat protein. J. Virol. 73: 553-560.

Leclerc, D., Stavolone, L., Meier, E., Guerra-Peraza, O., Herzog, E ., Hohn, T. 2001 . The
product of ORF III in cauliflower mosaic virus interacts with the viral coat protein

through its C-terminal proline rich domain. Virus Genes 22: 159-165.

Leh, V., Jacquot, E., Geldrich, A., Hermann, T., Leclerc, D., Ceruttio, M., Yot, P., Keller, M ., Blanc, S. 1999 . Aphid transmission of Cauliflower mosaic virus requires the viral PIII protein. EMBO 18: 7077-7085.

Mougeot, J., Guidasci, T., Wurch, T., Lebeurier, G ., Mesnard, J. 1993 . Identification of C-terminal amino acid residues of Cauliflower mosaic virus open reading frame III protein responsible for its DNA binding activity. PNAS 90: 1470-1473.

Needleman, S. B ., Wunsch, C. D. 1970 . A general method applicable to the search for similarities in the amino acid sequences of two proteins. J. Mol. Biol. 48: 443-453.

Ng, J ., Perry, K. L. 1999 . Stability of the aphid transmission phenotype of Cucumber mosaic virus. Plant Pathol. 48: 388-394.

Nicolaisen, M. 2003 . Partial molecular characterization of Dahlia mosaic virus and its detection by PCR. Plant Dis. Rep. 87: 945-948.

Pahalawatta, V., Druffel, K ., Pappu, H. R. 2005 . Molecular characterization of Dahlia mosaic virus genome. Phytopathol. 95: S79.

Pahalawatta, V., Druffel, K ., Pappu, H. R. 2007a . Seed Transmission of Dahlia mosaic virus in Dahlia pinnata. Plant Dis. 91: 88-91.

Pahalawatta, V., Miglino, R., Druffel, K. L., Jodlowska, A., van Schadewijk, A. R ., Pappu, H. R. 2007b. Incidence and relative distribution of distinct Caulimoviruses Genus

Caulimovirus, Family *Caulimoviridae* Associated with Dahlia Mosaic in *Dahlia pinnata*. Plant Dis: In press .

Pappu, H. R., Wyatt, S. D ., Druffel, K. L. 2005 . DMV: Molecular detection and distribution in dahlia in the US. HortScience 40: 697-699.

Pfeiffer, P., Hohn, T. 1983. Involvement of reverse transcription in the replication of *Cauliflower mosaic virus*: a detailed model and test of some aspects. Cell 33:781-789.

Presting, G. G., Smith, O. P ., Brown, C. R. 1995 . Resistance to potato leaf roll virus in potato plants transformed with the coat protein gene or with vector control constructs. Phytopathol 85: 436-442.

Ryabova, L. M., Pooggin, M. M., Hohn, T. 2002. Viral strategies of translation initiation: ribosomal shunt and reinitiation. Prog. Nucleic Acid Res. Mol. Biol. 72: 1-39.

Richins, R. D., Scholthof, H. B., Shepherd, R. J. 1987. Sequence of figwort mosaic virus DNA (caulimovirus group). Nucleic Acids Res. 15: 8451-8466.

Richins, R. D ., Shepherd, R. J. 1983 . Physical maps of the genomes of Dahlia mosaic virus and Mirabilis mosaic virus - two members of the Caulimovirus group. Virology 124: 208-214.

Sambrook, J., Fritsch, E. F ., Maniatis, T. 1989 . Molecular cloning: A laboratory manual. New York: Cold Spring Harbor Laboratory.

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. PNAS 91: 8885-8889.

Shababi, M., Bourque, J., Palanichelvam, K., Cole, A., Xu, D., Wan, X. F., Schoelz, J. 2006. The ribosomal shunt translation strategy of *Cauliflower mosaic virus* has evolved from ancient long terminal repeats. J. Virol. 80: 3811-3822.

Stavolone, 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.

Stavolone, L., Ragozzino, A., Hohn, T. 2003 . Characterization of *Cestrum yellow leaf curling virus*: a new member of the family *Caulimoviridae*. J. Gen. Virol. 84: 3459-3464.

Thompson, J. D., Higgins, D. G ., Gibson, T. J. 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.

Torruella, M., Gordon, K ., Hohn, T. 1989 . Cauliflower mosaic virus produces and aspartic proteinase to cleave its polyproteins. EMBO 8: 2819-2825.

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.

Zuker, M., Mathews, D. H., Turner, D. H. 1999. Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide in RNA Biochemistry and

Biotechnology, J. Barciszewski and B.F.C. Clark, eds., NATO ASI Series, Kluwer

Academic Publishers.

CHAPTER THREE

INCIDENCE AND RELATIVE DISTRIBUTION OF DISTINCT CAULIMOVIRUSES (GENUS *CAULIMOVIRUS*, FAMILY *CAULIMOVIRIDAE*) ASSOCIATED WITH DAHLIA MOSAIC IN *DAHLIA PINNATA*

ABSTRACT

Dahlia mosaic, caused by *Dahlia mosaic virus* (DMV), is one of the most important viral diseases of dahlia. Molecular characterization of DMV showed the association of two distinct caulimoviruses (DMV-D10, DMV-Portland) and a D10-like sequence variant (DMV-Holland) with the disease. Using primers specific to these two viruses and the sequence variant, a polymerase chain reaction-based assay was used to determine their relative incidence in several dahlia samples from the USA and the Netherlands. Testing was done on samples collected in 2005 and 2006 in the USA and in 2006 in the Netherlands. Results indicated the predominance of DMV-D10 over DMV-Portland and DMV-Holland in both USA and the Netherlands. Using conserved regions of the viral genome, primers were designed and used to detect all three sequences. Results suggested that DMV-D10 is predominantly associated with dahlia mosaic but diagnostics should also include testing for DMV-Portland and DMV-Holland.

INTRODUCTION

Dahlia mosaic virus (DMV) of genus *Caulimovirus* and Family *Caulimoviridae* is one of the most important viruses of dahlia (*Dahlia pinnata*) in USA (Pappu *et al*., 2005) and several parts of the world (Albouy, 1995; Brunt, 1971). Symptoms caused by DMV include mosaic, chlorotic vein banding and leaf malformation (Pappu & Wyatt, 2003). DMV is geographically widespread but the natural host range is limited only to *Dahlia* spp. A survey by Pappu *et al*. (2005) using a PCR-based detection assay of dahlias collected from several states in USA showed a very high incidence of DMV in dahlias. Accurate and reliable detection of the virus is required for developing effective virus elimination programs. Based on the partial molecular characterization of the DMV genome, PCR-based detection of DMV was reported (Bogunov, 2006; Nicolaisen, 2003; Pappu *et al*., 2005).

As part of an ongoing study to determine the sequence variability of the DMV genome, we identified a distinct caulimovirus in dahlia, referred to as DMV-D10 (Pahalawatta *et al*., 2005) which was found to be distinct from another caulimovirus (DMV-Portland) (Richins & Shepherd, 1983). Partial genomic sequence representing the viral coat protein and reverse transcriptase genes of another isolate, referred to as DMV-Holland, has been reported (Nicolaisen, 2003). While our previous survey for the incidence of DMV-D10 showed its widespread occurrence in USA (Pappu *et al*., 2005), no information is available on the incidence of DMV-Portland. Moreover, no information is available on the incidence of DMV-Holland in Europe or in the USA. Due to the continuous movement of dahlia planting material within USA and between USA and Europe, the relative incidence of these viruses should be ascertained to devise effective and comprehensive detection and virus elimination strategies. In this study,

using primers specific to DMV-D10, DMV-Portland, and DMV-Holland, the relative distribution was determined in samples collected from several states in USA and from the Netherlands.

MATERIALS AND METHODS

Source of dahlia samples. One hundred and nineteen dahlia plant samples collected from commercial nurseries in California, Georgia, Montana, New Mexico, Oklahoma and dahlia varietal trial gardens in Maryland and Washington State in USA during 2005 and 2006 were used in the study. In addition, 41 samples from field-grown dahlia plants from a dahlia varietal trial garden in Lisse, the Netherlands collected in 2006 were also used. Samples were collected at random irrespective of the presence of symptoms suggestive of virus infection.

Total nucleic acid purification. Total nucleic acid was purified according to a modified Dellaporta procedure (Pappu *et al*., 2005; Presting *et al*., 1995). Approximately 0.1g (or a quarter-sized leaf piece) was used for DNA extraction.

Primers used in PCR. Several primer pairs encompassing different regions of the viral genomes were designed for the specific detection of the caulimoviruses associated with the disease. One set consisting of three primer pairs, designated DMV-D10, was designed for the specific detection of DMV-D10 (Table 1). A second set, designated DMV-Portland, consisting of six primer pairs (Richins & Shepherd, 1983) and a third set, designated DMV-Holland (Nicolaisen, 2003), were designed based on the sequences available in GenBank for the specific detection of DMV-Portland and DMV-Holland, respectively (Table 1). Two DMV-D10, three

DMV-Portland and two DMV-Holland primer pairs were used for exclusive detection of the two viruses and the sequence variant respectively, in the dahlia samples. A primer pair, Den518-F/Den1156C, that could detect DMV-D10, -Holland and -Portland was designed based on the conserved region of the 1171 bp sequence representing partial open reading frames of coat protein and the reverse transcriptase genes (Figure 1).

PCR amplification. Two microliters of 1:5 diluted total nucleic acid from above was used in each PCR reaction containing 2µl 10X PCR buffer (Invitrogen, Carlsbad, CA), 1.2µl 2.5mM dNTP's, 0.8µl 50mM $MgCl_2$, 0.6µl 20µM primer 1, 0.6µl 20µM primer 2, 13.6µl ddH_2O, and 0.2µl Taq DNA polymerase (Invitrogen). PCR amplification was performed in a DNA thermal cycler (BioRad, Hercules, CA) programmed for 3 min at 94°C for initial denaturation and 50 cycles each consisting of a denaturation step at 94°C for 30 sec, the required annealing temperature based on the primer pair used for 20 sec and an extension step at 72°C determined based on the size of the amplicon to be synthesized at the rate of 1000 bp of PCR per minute followed by a final extension for 7 min at 72°C. PCR reactions (7.5µl) were analyzed by agarose gel electrophoresis.

Cloning and sequencing. PCR products of expected size were cloned using pGEM-T (Promega, Madison, WI) or TOPO TA (Invitrogen) cloning kits. Some PCR products were sequenced directly after purification using QIAquick PCR purification kit (Qiagen, Valencia, CA). Nucleotide sequences were determined using the ABI Prism sequencing system at the Molecular Biology Core Laboratory of the Washington State University, Pullman, WA. Recombinant

plasmids were sequenced using M13 forward and reverse primers, whereas PCR products were

sequenced using the specific primer pairs that were initially used to obtain the PCR product.

Table 1. Primers used for the detection of *Dahlia mosaic virus* (DMV)-D10, DMV-Portland and DMV-Holland in *Dahlia pinnata.*

GenBank Accession	Primer pair	Sequence (5'-3')	T_m (°C)	Expected size (bp)
	DMV-D10			
	ORF1-START	ATGGATCGTAAAGATT	55	
	ORF1-END1	CTGTTTTTCTGTGTTTCTACTGG		900
	ORF4CP1-1431F[a]	TGCATAAAATGAGTTCTATC		
	ORF4CPI-1926C	TGAACTTGTTCATCATTATC	63	480
	ORF6-START	ATGGAAGAAATTAAGGCGT		
	ORF6-END1	TTGTCTTCATCCATAAAGCAG	65	1280
	DMV-Portland			
	KORF1-F	ATG AAT ATC TTA GAA AGG AA	55	966
AY291585	KORF1-R	CTT AAT CCT TAA GTT ATC AA		
	Kapht-F	ATG AGT AAT GCT TCA GCA A	61	504
AY291586	Kapht-R	TGA CCA TGG CTT CTA ACT GT		
	KDNAb-F	TGA CCA CCA TCA AAG ACT TA	64	363
AY291587	KDNAb-R	TCA TAC TGG AGG CCA TTT TA		
	Kcp-F	ATG GCC TCC AGT ATG AAA G	58	1476
AY291588	Kcp-R	TTA TTC TGT TCC TGA TGA TT		
	Kpoly-F	ATG TTG TCA CAA CAG ATG AT		
AY309479	Kpoly-R		55	1830
		ATT CTT TTT CAT TAC TAT GG		
	KIB-F	ATG GAG GAA GAA TTA AAA GC		
AY309480	KIB-R		56	1515
		TTA TAT AGG CAA GTC TTC AG		
	DMV-Holland			
	DenORF4-F	TCCAAGCAGACAGACAAACC	61	
AJ515906	DenORF4-R1	ATTGCATTTCATTTAGAACA		600
	DenORF4-F1	CAGCAAGAAACAGGAATTGA		
AJ515906	DenORF4-R	TTTAGCAGCTTCGACTGTAAA	65	600

```
                                    →
D10        GTATAGATTTACTCCTTGGAAATAACTTTTGTCAATTATATGGGCCATTCACCCAATGGA 1977
Portland   GAATAGATATCCTTTTAGGAAATAATTTCTGCCAACTATATAATCCATTCATACAATGGG 408
Holland    GAATTGACATCCTCATAGGAAATAATTTCTGTAGACTCTACAACCCTTTCATTCAATGGG 572
           * **  **   * **    * ********  ** **    * * **   ** ****  ******
               GAYWTMCTYHTWGGAAATAA[a]

D10        TTGACAGAATATCTTTTCATCTTAATCAAGAAAT------GATTTTGATCAAAAAGATAA 2031
Portland   TTGACCGGATTGCTTTTCACAAAAATGAAGAGAT------AATTCTTGTACCTAAAGTAC 462
Holland    AAGATAGAATTGCATTTCATTTAGAACAAGGAACCAAACAAATCCTAGTTCCTAAACAAA 632
            **  * **  * *****     *  *** *       **  *  *     **  *

D10        CAAAAGCTTTTCAAATTGGAAAAACTGGTTTTCTTGAATCCATGAAAAAAGATTCAAAAA 2091
Portland   GAAAAGCCATGAGAAAAGGATATGAAGGTTTCCTTAAAACCATGGAAAGAGATTCAAAAT 522
Holland    GAAAAGCAATAAAAAGAGGTTATCAAGGATTCCTTGAAACCATGAAGAAAGAGTCAAACA 692
            ******   *   ** **   *   ** ** *** ** ***** * * *** *****

D10        CTAAACAAATTCCGGGTACCAATATAACCCAAGAAGTTATAAAACCGGAAA--GATTTTT 2149
Portland   CTCAGCCTACCCCAGGCACCAACATTACCCAAGAGGTAATTGATGAAGAAAACAACTTCA 582
Holland    CTCAACCACTGCCAGGTACGAACATCACCAAAGAAGCAGTAGAAGATGAAAGCAACTTCA 752
           ** * *       ** ** ** ** ** *** **** *   *  *   ****   * **

D10        TC-TTGAAATACAGAGATATCAAAAAATTGAAGATCTACTTGAAAAAGTATGTTTTGAAA 2208
Portland   TCATTGATATCTCCAGATATCAAGAGATTCAAGAACTACTTAAGAGAGTATGTTCAGAAA 642
Holland    TCATAGATATCTCAAGATATCAAGAGATTCAAAACTTACTTGAAAAAGTTTGTTCTGAAA 812
           ** * ** **   ********* * *** ** *  ***** * * *** ****  ****

D10        ACCCTATTGACCCAGAAAAGTCCAAATACTGGATGAATGCCTCAATAGAATTAATAGATC 2268
Portland   ATCCTATTGATCCTCAGAAGTCTAAAGGATGGATGACGGCATCAATAAAACTTGCTGACC 702
Holland    ATCCTTTAGACCCAGCCAAATCCAAAGCGTGGATGAAAGCATCAATAAAACTAGCTGATC 872
            * *** * ** **    ** ** ** ***  ***  ** ****** ** *    ** *

D10        CCAAAACCGTGGTTAGGGTAAAACCCATGAAATATAGTCCACAGGATCGCGAAGAATTTG 2328
Portland   CAAAGTCAGTAGTTAAGGTGAAACCAATGGTTTATTCACCCACAGGATAGAAAGGAATTCG 762
Holland    CAAAATCAGTAGTCAGAGTCAAACCTATGGTTTACTCACCAGAAGACAGAAAAGAGTYTG 932
           * **  * ** ** * *  ** *****  ***    **   ***  *   *  * ** *

D10        GTAAGCAAATCAAAGAATTACTTGATTTAAAATTAATTATTCCAAGTAAATCTCCTCACA 2388
Portland   AAATCCAAATTAAGGAACTCCTTAATCTTAAGGTAATAATTCCAAGTAAATCCCAACACA 822
Holland    AGATTCAAATTAAGGAACTCCTTGACCTAAAAGTCATAGAACCCAGCRAATCTCAACACA 992
            *  *****  ** *** * *** *   * ** * * **   ** **    * ****

D10        TGTCTCCAGCTTTTCTTGTAGAAAACGAAGCAGAAAAAAGGAGAGGAAAAAAGCGCAGGG 2448
Portland   TGTCCCCTGCTTTTCTCGTTGAAAAGAAGCAGAAAAAACGACGTGGCAAGAAAAGAATGG 882
Holland    TGYCACCCGCCTTTTTGGTTGAAAAGGAAGCGGAAAAAAGAAGAGGAAAAAAGAGAATGG 1052
           ** * ** ** *** * ** **** ***** ****** *  * ** ** **   * * **

D10        TTGTAAATTATAAGGCAATTAATGCAGCTACCAAAGGAGACAGCCATAATCTTCCTTGTA 2508
Portland   TTGTTAACTATAAAAAGTTAAATGAAGTCACCATTGGTGATTCTCACAACCTTCCTAACA 942
Holland    TCGTTAACTACAAAAAGTTAAACGAAGTCACCATTGGTGATTCTCACAACCTTCCCCACA 1112
           * ** ** ** **       * ** * **   **** ** **   ** ** ***** *

                                  ←
D10        TGCAGGAATTATTAACCCCTCTTAAGAGGAAAAACCATTTTTTCTACTTTCGACTGCAAGA 2568
Portland   TGCAAGAGTTAATTACTCTTCTTAGAGGAAAATCTATCTTTAGCAGTTTTGACTGTAAAA 1002
Holland    TGCAAGAGTTAATCACTCTGCTCAGAGGAAAAACCATCTTTAGCAGCTTCGACTGTAAA- 1171
           **** ** ***  * **  *   * ********* * ** ***   * ** *****  **
                   CTCCTTTTWGRTARAAA[b]
```

55

Figure 1. Nucleotide sequence alignment of the reverse transcriptase gene of DMV-D10 (unpublished), DMV-Portland (Genbank accession AY309479) and DMV-Holland (Genbank accession AJ515906) using CLUSTAL W (Thompson *et al*., 1994). * - identical nucleotides. Boxed regions denote the sequence used to design degenerate primer Den518-F[a] and Den1156C[b]. Degenerate primer sequences are given in bold.

Sequence analysis. Sequences were compiled using PC-based AlignX (Vector NTI Suite 9, Informax Inc, Bethesda, MD) or SeqAid (Rhoads & Roufa, 1985) software. Multiple alignments were generated using CLUSTAL W (Thompson *et al*., 1994). Sequences of DMV available in GenBank were used for comparisons. Primers were designed by using Primer Designer version 2.0 (Scientific and Educational Software, Cary, NC).

RESULTS

Each of the primer pairs used produced a specific amplicon of expected size. Amplicons resulting from using each of the primer pairs were sequenced to verify their identity. A majority of the dahlia leaf samples collected in USA (87.4%) and the Netherlands (97.6%) produced amplicons of expected size with the DMV-D10 primers. Results of PCR using the DMV-Portland specific primers showed fewer positives, with 21.8% of the USA samples and 53.6% of the samples from the Netherlands producing amplicons of expected size. In contrast, 97.6% of the samples from the Netherlands produced amplicons of expected size with DMV-Holland primers, whereas only 31.9% of the USA samples tested positive (Table 2). Only 6.7% of the samples from USA tested positive for all three viruses, whereas 36.6% of samples tested from the Netherlands were positive for all three. All samples positive with DMV-Portland were also positive for either DMV-D10 or DMV-Holland. The percentage of USA samples positive for both DMV-D10 and DMV-Holland was 28.6% whereas, 21.0% of the samples were positive for both DMV-D10 and DMV-Portland. A similar trend was seen in the samples from the Netherlands, where 41.6% of the samples were positive for both DMV-D10 and DMV-Holland,

57

while none of the samples were positive for only DMV-D10 and DMV-Portland primers. These results indicate that the most prevalent DMV sequences in dahlia in both USA and the Netherlands were those amplified by DMV-D10 primers.

Primers based on the conserved region among DMV-D10, DMV-Portland and DMV-Holland (Figure 1) were capable of amplifying all three viral sequences associated with dahlia mosaic (Figure 2).

DISCUSSION

Overall, the data indicate that there is a higher incidence of DMV-D10 in USA, whereas in the Netherlands both DMV-D10 and DMV-Holland appear to be prevalent. It is not surprising that more samples from the Netherlands were positive for DMV-Holland since the sequence from which our primers were derived was obtained from a Dutch isolate (Nicolaisen, 2003). These results also highlight the importance of testing dahlia planting material for viruses associated with dahlia mosaic prior to export to other countries in the light of our findings that DMV-D10, DMV-Holland and DMV-Portland sequences seem to be present at a different frequency in dahlias from USA than in the dahlia samples in the Netherlands.

Table 2. Percentage of dahlia samples positive for caulimoviruses associated with dahlia mosaic.

Primer set	USA.	Netherlands
DMV-D10	87.4	97.6
DMV-Portland	21.8	53.6
DMV-Holland	31.9	97.6
DMV-D10+DMV-Portland	21.0	43.9
DMV-D10+DMV-Holland	28.6	97.6
DMV-Portland+DMV-Holland	6.7	17.1
DMV-D10+DMV-Portland + DMV-Holland	6.7	36.6
Negative	4.2	0
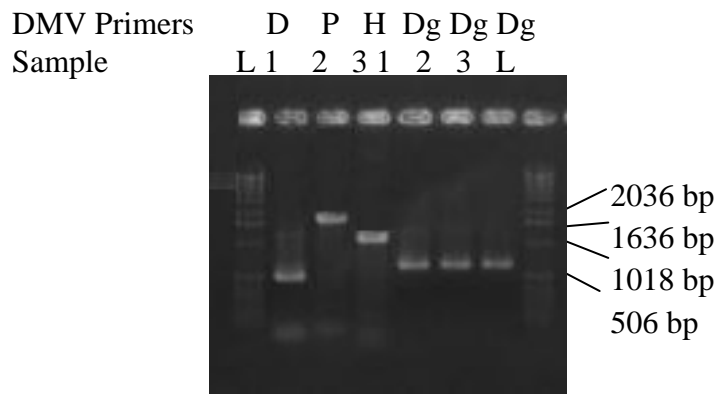
Figure 2. Agarose gel electrophoresis of PCR products from DNA extracts of three dahlia samples amplified with primers specific to *Dahlia mosaic virus* (DMV)-D10 (D), DMV-Portland (P), DMV-Holland (H) and degenerate (Dg) primers that could amplify all three caulimoviruses associated with dahlia mosaic. L – 1 kb ladder.

Pairwise comparisons of amino acid sequences representing each open reading frame of DMV-D10 (Pahalawatta *et al.*, 2005) and DMV-Portland (Genbank accessions AY309480, AY309479, AY291588, AY291587, AY291586, AY291585) showed that the sequence identity between DMV-D10 and DMV-Portland ranged from 47% to 73%. This divergence suggests that there are at least two distinct caulimoviruses associated with dahlia mosaic. The third primer pair representing DMV-Holland amplifies a part of the coat protein gene and the reverse transcriptase gene (Nicolaisen, 2003). The ca. 1.2 kb sequence of DMV-Holland has 59.1% sequence identity with the corresponding sequence in DMV-D10 and 41.2% identity with that of DMV-Portland. Thus, it appears that there are at least two and quite possibly three distinct caulimoviruses associated with dahlia mosaic. However, complete genome characterization of DMV-Holland is needed to establish whether DMV-Holland is a distinct caulimovirus infecting dahlia. Our results also indicate that the co-existence of divergent caulimoviruses in a given dahlia plant is common. However, at present, there is no successful infectivity assay to correlate mosaic symptoms to a particular virus/sequence variant in dahlias.

DMV infection often remains latent and as a result many infected but asymptomatic plants and propagating material are distributed which might have contributed to its widespread distribution in the USA (Pappu *et al.*, 2005). At present, PCR-based detection assay is the most reliable method of detecting DMV in dahlias. Considering the diversity in caulimoviruses associated with dahlia mosaic disease and the prevalence of mixed infections, use of primers specific to a single virus is not adequate for disease diagnosis as it may lead to false negatives for other caulimoviruses that may be present.

REFERENCES

Albouy, J. 1995. Dahlia. 265-273 in: Virus and virus-like diseases of bulb and flower crops. G. Lobenstein, R. H. Lawson, and A. A. Brunt. eds. John Wiley and Sons, Chichester, New York.

Bogunov, Y. V. 2006. Identification of *Dahlia mosaic virus* with molecular biological methods. Mol. Biol. 40: 184-185.

Brunt, A. A. 1971. Some hosts and properties of dahlia mosaic virus. Annal. Appl. Biol. 67: 357-368.

Nicolaisen, M. 2003. Partial molecular characterization of *Dahlia mosaic virus* and its detection by PCR. Plant Dis. 87: 945-948.

Pahalawatta, V., Druffel, K., Pappu, H. R. 2005. Molecular characterization of *Dahlia mosaic virus* genome. Phytopathol. 95: S79.

Pappu, H. R., Wyatt, S. D., Druffel, K. L. 2005. DMV: Molecular detection and distribution in dahlia in the US. HortScience 40: 697-699.

Pappu, H. R., Wyatt, S. D. 2003. Dahlia mosaic virus. A description of symptoms on dahlia. APSnet Image of the Week. Online/2003/IW000032.asp.

Presting, G. G., Smith, O. P., Brown, C. R. 1995. Resistance to potato leaf roll virus in potato plants transformed with the coat protein gene or with vector control constructs. Phytopathol. 85: 436-442.

Rhoads, D. D., Roufa, D. S. 1985. Emetine resistance of Chinese hamster cells: structure of

    wild type and mutant ribosomal protein S14 messenger RNA species. Mol. Cell. Biol.

    5:1655-1659.

Richins, R. D., Shepherd, R. J. 1983. Physical maps of the genome of *Dahlia mosaic virus* and

    Mirabilis mosaic virus – two members of the Caulimovirus group. Virology 124: 208-

    214.

Thompson, J. D., Higgins, D. G., Gibson, T. J. 1994. CLUSTAL W: Improving the sensitivity

    of progressive multiple sequence alignment through sequence weighting, positions-

    specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680.

CHAPTER FOUR

SEED TRANSMISSION OF DAHLIA MOSAIC VIRUS IN DAHLIA PINNATA

ABSTRACT

Dahlia is an economically important ornamental crop in the USA and several other countries in the world. Among the viral diseases that affect dahlia, *Dahlia mosaic virus* (DMV) is considered to be the most widespread and to have the greatest impact on flower production. Using grow-out tests followed by PCR-based testing of the seedlings, dahlia seed obtained from three different sources were shown to contain DMV. Additionally, the distribution of DMV in various parts of the dahlia seed was determined by PCR. Grow-out tests revealed a high rate of seed transmission. DMV was detected in cotyledons. The virus was also detected in pollen collected from infected plants. In addition to vegetative propagation, seed-borne infection could be contributing to the spread of DMV in dahlia. Use of virus-free seed and vegetative material would result in reduced incidence of the disease.

INTRODUCTION

Dahlia mosaic virus (DMV), a member of the genus *Caulimovirus*, family *Caulimoviridae* is an economically important viral pathogen of dahlias. First reported in *Dahlia pinnata* from Germany, the disease is prevalent in *Dahlia spp.* and has become one of the most dominant viruses in the floriculture/ornamental industry (Brunt, 1971; Pappu *et al*., 2005). The

disease is characterized by symptoms of mosaic, chlorotic vein banding, and leaf malformations but symptoms vary depending on dahlia cultivar (Pappu & Wyatt, 2003). Dahlia cultivars with few or no symptoms are considered to be important reservoirs of virus. DMV is geographically widespread but the natural host range is limited to only *Dahlia* spp (Albouy, 1995). Previous studies on DMV transmission have resulted in identification of 16 aphid species, including *Aphis fabae, Myzus persicae* and *Macrosiphum euphorbiae* (Brierley & Smith, 1950), as vectors transmitting the virus in a non-persistent manner. A virus-encoded aphid transmission protein is reported to be necessary for viral transmission (Albouy, 1995). DMV was also shown to be transmitted by mechanical inoculation (Brierley & Smith, 1950; Brunt, 1971). Previous studies on DMV report that the virus is not transmitted by seed (Brierley & Smith, 1950; Brunt, 1971). However, there have been no studies to confirm this definitively. Partial molecular characterization of the DMV genome has been reported (GenBank accessions AY309480, AY309479, AY291588, AY291587, AY291586, AY291585) (Nicolaisen, 2003; Pahalawatta *et al.*, 2005) and detection of DMV based on PCR and real-time PCR is now possible (Bogunov, 2005; Nicolaisen, 2003; Pappu *et al.*, 2005).

A survey by Pappu *et al.* (2005), using a PCR-based detection assay, of dahlias collected from several states in the USA, revealed a very high incidence of DMV in dahlias. Subsequently, numerous dahlia plants and seeds from different locations within the USA were tested with the PCR-based detection assay. The majority of plants and seeds tested positive for the presence of DMV, suggesting the occurrence of seed-borne viral infection. However, there have been no systematic studies to unequivocally demonstrate seed-borne viral infection or the potential of the virus to be seed transmitted. The objective of this study was to investigate seed

65

transmission of DMV in dahlia, including whether DMV could be pollen-borne, under conditions that eliminate the possibility of introduction of the virus by aphids.

MATERIALS AND METHODS

Seed sources. Three commercially available seed sources of *Dahlia* spp. from Kentucky, Tennessee and Colorado were used in this study. A total of twenty seeds from each seed source were used. Ten seeds were used for virus detection in seed coat and cotyledon and the other ten seeds were used for grow-out tests and testing for the virus in the resulting seedlings. In addition, five seeds each from two of the seed sources were grown in an insect-proof greenhouse. Plants were analyzed by PCR and inspected visually for development of DMV symptoms.

DNA extractions. Ten seeds from each source were soaked in distilled water for 2 to 4 hrs until the seed coat softened. Seeds were then blotted dry and the seed coats were separated from the cotyledons using forceps. Total nucleic acid was extracted from the separated seed coats and cotyledons using the modified Dellaporta method (Pappu *et al*., 2005; Presting *et al*., 1995). In addition to the Dellaporta method, DNA was also extracted from the seed coats using the Qiagen DNeasy plant kit (Qiagen, Valencia, CA). The remaining ten seeds from each seed source were allowed to germinate on moistened filter paper in covered Petri dishes (100x15 mm) to avoid exposure to insect vectors. After 9 to 10 days when the first true leaves started to emerge, total DNA was extracted from the seedlings as described above. Four weeks after planting, a single leaf from each plant grown in the greenhouse was also used for total DNA extraction.

Detection of DMV by PCR. DMV isolate D10, originally detected in an infected dahlia plant, was used to clone and sequence the viral genome. Based on the sequence information (Pahalawatta *et al.*, 2005), primer pairs were designed for specific detection of DMV in dahlias. Three primer pairs representing three open reading frames (ORFs) of the DMV genome were used to detect the virus in the seed, seedling and leaf samples (Table 1). Six additional primer pairs designed based on DMV sequences available in the GenBank (Accessions AY309480, AY309479, AY291588, AY291587, AY291586, AY291585) were also used. The PCR reaction included a final volume of 2µl 10X PCR buffer (Promega, Madison, WI) and a final concentration of 0.15mM dNTP's, 2mM MgCl$_2$ and 0.6µM of each primer in a total reaction volume of 20µl. PCR amplification was performed in a DNA thermal cycler (BioRad, Hercules, CA) programmed for 3 min at 94°C for initial denaturation and 50 cycles each consisting of 30 sec at 94°C, 20 sec at annealing temperature and 1 min extension per 1000 bp product at 72°C, followed by a final extension for 7 min at 72°C. Each PCR reaction included a DNA sample from a DMV-infected plant and a water control. The PCR products were analyzed by agarose gel electrophoresis in 1X TAE buffer. PCR products representing each ORF were cloned and sequenced to verify the identity of amplified fragments.

Table 1. *Dahlia mosaic virus* (DMV) specific primer pairs used for DMV-D10 detection by polymerase chain reaction

Primer pair	Sequence (5'-3')	T_m (^{o}C)	Expected size (bp)
ORF1-START	ATGGATCGTAAAGATT	55	900
ORF1-END1	CTGTTTTTCTGTGTTTCTACTGG		
ORF4CP1-1431F	TGCATAAAATGAGTTCTATC	63	480
ORF4CPI-1926C	TGAACTTGTTCATCATTATC		
ORF6-START	ATGGAAGAAATTAAGGCGT	65	1280
ORF6-END1	TTGTCTTCATCCATAAAGCAG		

RESULTS

DMV was detected in the cotyledons of all seeds analyzed with the three primer pairs specific for DMV-D10. The PCR reaction with primer pair ORF1-START/ORF1-END1, representing the first ORF, produced an amplicon of 900 bp.   Similarly, the primer pair ORF4CP1-1431F/ORF4CP1-1926C, representing ORF4, produced an amplicon of 480 bp, and the primer pair ORF6-START/ORF6-END1, representing ORF6, produced an amplicon of 1280 bp. All amplicons produced were within the expected size range for DMV. When the Dellaporta method was used for DNA extraction the virus could not be detected in the seed coat except in two cases where seed coats from two seeds were positive for DMV using the primers derived from ORF 4 of DMV-D10(Table 2 and Fig. 1). However when the DNA extracts from the seed coat, using the Qiagen DNeasy kit, were used in the PCR reactions using the same primer pairs the virus was detected. All seedlings obtained from seeds germinated on filter paper also tested positive for the presence of the virus with each of the three primer pairs (Table 2 and Fig. 2). Similarly, DNA extracts from 4- week-old leaf samples, taken from the plants grown in the

greenhouse, were also positive for DMV when tested by PCR although the plants were asymptomatic (Fig. 3a). All of these plants developed mosaic symptoms, to varying degrees, 5 to 7 weeks after planting (Fig. 3b). The six primer pairs designed based on DMV sequences available in the GenBank failed to produce any amplicons.

Table 2. Detection of *Dahlia mosaic virus*-D10 by polymerase chain reaction in seeds and seedlings from three different commercial seed sources. DNA was extracted using the modified Dellaporta method (Pappu *et al*., 2005; Presting *et al*., 1995).

Dahlia seed source	PCR primer pair	No. positive for DMV/ No. tested		
		Seed coat	Cotyledon	Seedlings[a]
Tennessee	ORF1-START ORF1-END1	0/10	10/10	10/10
	ORF4CP1-1431F ORF4CP1-1926C	0/10	10/10	10/10
	ORF6-START ORF6-END1	0/10	10/10	10/10
Kentucky	ORF1-START ORF1-END1	0/10	10/10	10/10
	ORF4CP1-1431F ORF4CP1-1926C	0/10	10/10	10/10
	ORF6-START ORF6-END1	0/10	10/10	10/10
Colorado	ORF1-START ORF1-END1	0/10	10/10	10/10
	ORF4CP1-1431F ORF4CP1-1926C	1/10	10/10	10/10
	ORF6-START ORF6-END1	0/10	10/10	10/10

[a] Seedlings were tested when 9 to 10-days-old and asymptomatic.
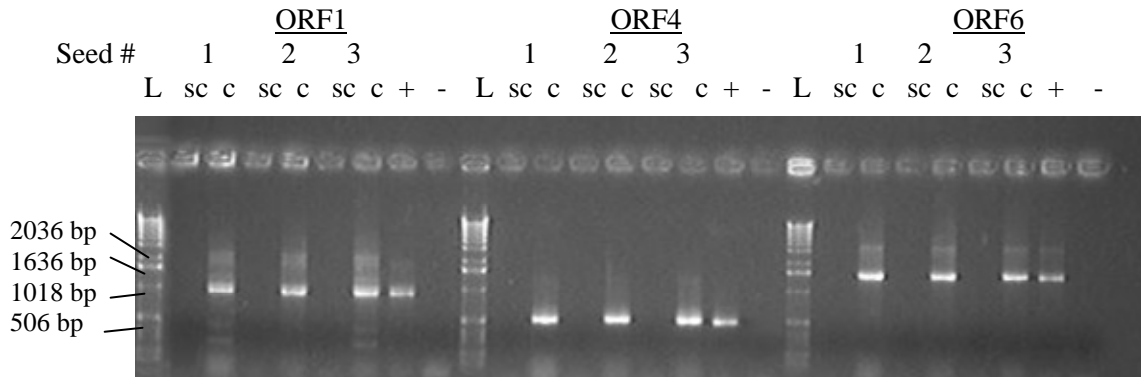
Figure 1. Agarose gel electrophoresis of PCR products from DNA extracts of dahlia seed coats and cotyledons amplified with *Dahlia mosaic virus* specific primer pairs. DNA was extracted using the modified Dellaporta method (Pappu *et al*., 2005; Presting *et al*., 1995). Representative sample of PCR products for DNA extracts of dahlia seeds amplified with ORF1-START/ORF1-END1 (ORF1), ORF4CP1-1431F/ORF4CP1-1926C (ORF4) and ORF6-START/ORF6-END1 (ORF6). sc- seed coat, c – cotyledon, + - positive control, - negative control, L- 1 kb DNA ladder.
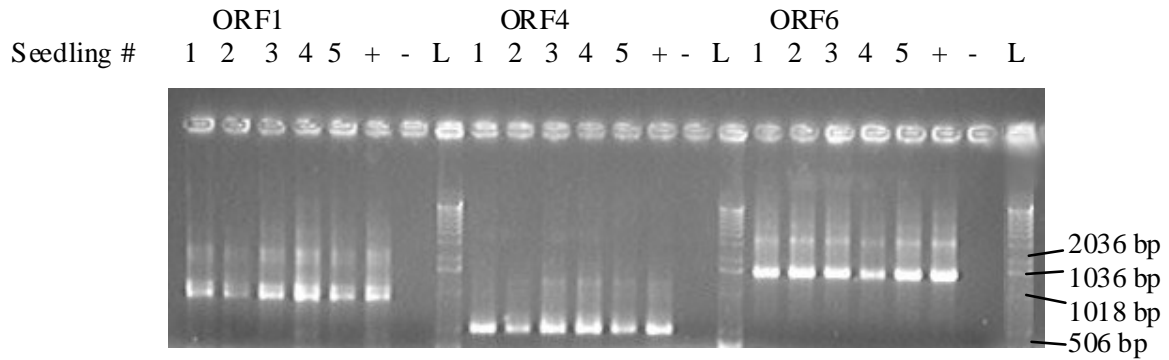
Figure 2. Agarose gel electrophoresis of PCR products from DNA extracts of 9 to10-day-old dahlia seedlings germinated in covered Petri dishes and tested with *Dahlia mosaic virus* specific primer pairs. ORF1 = ORF1-START/ORF1-END1, ORF4 = ORF4CP1-1431F/ORF4CP1-1926C, ORF6 = ORF6-START/ORF6-END1, + - positive control, - negative control, L- 1 kb DNA ladder.

A)

B)

Figure 3. Seedlings positive for *Dahlia mosaic virus* (DMV) by PCR. A) Asymptomatic four week-old DMV-infected dahlia seedling grown from seed under aphid-free conditions; B) DMV-infected seedling displaying symptoms (mosaic and chlorosis) at 5 to 7 weeks after planting. The seedling was grown from seed under aphid-free conditions.
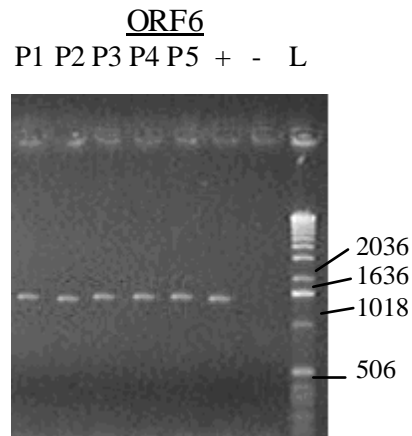
Figure 4. Agarose gel electrophoresis of PCR products, from DNA extracts of pollen collected from virus-infected dahlias, amplified with *Dahlia mosaic virus* specific primer pair ORF6-START/ORF6-END1. + - positive control, - negative control, L- 1 kb DNA ladder.

DISCUSSION

Our results indicate that DMV is seed-transmitted in dahlia. All the seeds analyzed using the PCR-based assay tested positive for DMV demonstrating the seed-borne nature of the virus. Generally, whole seed viral assays are considered unsuitable in estimating the rate of seed transmission because viral presence in the seed coat has no correlation to virus transmission from seeds to seedlings (Gillaspie *et al*., 1993; Maury *et al*., 1987). In our study dahlia seed coats and cotyledons were separated and individually tested for DMV. When the Dellaporta method was used for DNA extraction, the results showed that the viral infection is associated primarily with the cotyledons and rarely with the seed coat. However when the Qiagen DNeasy kit was used

for DNA extraction viral infection was detected in the seed coat as well. It is highly probable that the contradictory results are due to the purity of the DNA extracts obtained using the two different techniques. It is possible that the seed coat contained inhibitors to PCR that were remove by the DNeasy spin column. Due to the small size of the dahlia seeds it was not possible to separate the embryo axis from cotyledon. Seed transmission was confirmed by detecting DMV in very young seedlings (9 to 10-day-old) that were germinated in covered Petri dishes which excluded the possibility of virus transmission by aphid. We also found that DMV could be amplified in total DNA extracts of pollen collected from dahlia flowers of infected plants (Fig. 4). Furthermore, viruses transmitted through seed are reported to have certain common characteristics (Johansen *et al*., 1994) such as mechanical transmissibility, narrow host range among annual plants and vector transmission in a non-persistent manner which are also exhibited by DMV. Another caulimovirus, *Cauliflower mosaic virus*, has been reported to be seed-borne due to the occurrence of the virus in the seed testa (Tomlinson & Walker, 1973) although no subsequent transmission to seedlings was recorded. The frequency of seed transmission for plant viruses varies from 0 to 100% and rarely exceeds 50% for most viruses (Mink, 1993). It was reported that plant pararetrovirus DNA might integrate into host chromosomes (Jakowitsch *et al*., 1999; Schoelz *et al*., 2005). Research is in progress to determine if the high frequency of detection of DMV sequences in dahlia could be due to such an event.

The high rate of seed transmission of DMV combined with delayed symptom development, both observed in this study, may be partly responsible for the widespread occurrence of the virus in the US (Pappu *et al*., 2005). Four-week-old dahlia plants tested positive for DMV but symptoms were not evident until 1 to 3 weeks later. This demonstrates the

need for testing young plants for DMV since asymptomatic infections contribute to inadvertent distribution of virus-infected material and its subsequent propagation.

The six primer pairs designed from the DMV sequences available in GenBank did not produce any amplicons. This could be due to the limited sequence homology between the D10 isolate used in this study and those available in GenBank. Also, amino acid sequence similarity between D10 and those available in GenBank ranged from 49% to 73% for the three ORFs represented by the amplicons obtained using D10-specific primers. Preliminary sequence data obtained in our laboratory and reports (Nicolaisen, 2003) suggest that there are at least three distinct populations potentially associated with dahlia mosaic. Therefore it is highly probable that the DMV sequences in GenBank are not represented in the DMV isolates used in this study.

One of the biggest setbacks to detecting DMV and eliminating diseased plants is the propagation of infected plants and continued production of plants with asymptomatic infections. The results of this study demonstrate the potential role of seed transmission in the spread of DMV. While the majority of the dahlia propagation is through tubers, control options should include use of virus-free seed to reduce the impact of DMV.

REFERENCES

Bogunov, Y. V. 2006. Identification of *Dahlia mosaic virus* with molecular biological methods. Mol. Biol. 40:184-185.

Brierley, P., Smith, F. F. 1950. Some vectors, hosts and properties of dahlia mosaic virus. Plant Dis. Rep. 34:363-370.

Brunt, A. A. 1971. Some hosts and properties of dahlia mosaic virus. Anal. App. Biol. 67:357-368.

Gillaspie, J. A. G., Hopkins, M. S., Pinnow, D. L. 1993. Relationship of cowpea seed-part infection and seed transmission of blackeye cowpea mosaic potyvirus in cowpea. Plant Dis. 77:875-877.

Jakowitsch, J., Mette, M. F., van der Winden, J., Matzke, M. A., Matzke, A. J. M. 1999. Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. Proc. Natl. Acad. Sci. USA 96:13241-13246.

Johansen, E., Edwards, M. C., Hampton, R. O. 1994. Seed transmission of viruses: Current perspectives. Annu. Rev. Phytopathol. 32:363-386.

Lobenstein, G., Lawson, R. H., Brunt, A. A. 1995. Viruses and virus-like diseases of bulb and flower crops. Pages 266-267. John Wiley and Sons. Chichester, NY.

Maury, Y., Bossennec, J. M., Boudazin, G., Hampton, R., Pietersen, G., Maguire, J. 1987. Factors influencing ELISA evaluation of transmission of pea seed-borne mosaic virus in infected pea seed: seed-group size and seed decortication. Agronomie 7:225-230.

Mink, G. I. 1993. Pollen- and seed-transmitted viruses and viroids. Annu. Rev. Phytopathol. 31:375-402.

Nicolaisen, M. 2003. Partial molecular characterization of *Dahlia mosaic virus* and its detection by PCR. Plant Dis. 87:945-948.

Pahalawatta, V., Druffel, K., Pappu, H. R. 2005. Molecular characterization of *Dahlia mosaic virus* genome. Phytopathology 95:S79.

Pappu, H. R., Wyatt, S. D. 2003. Dahlia mosaic virus. A description of symptoms on dahlia. APSnet Image of the Week. Online/2003/IW000032.asp.

Pappu, H. R., Wyatt, S. D., Druffel, K. L. 2005. DMV: Molecular detection and distribution in dahlia in the US. HortScience 40:697-699.

Presting, G. G., Smith, O. P., Brown, C. R. 1995. Resistance to potato leaf roll virus in potato plants transformed with the coat protein gene or with vector control constructs. Phytopathology 85:436-442.

Schoelz, J. E., Wiggins, B. E., Balaji, B., Palukaitis, P. F., Squires, J. 2005. Integration of *Cauliflower mosaic virus* into the genome of *Arabidopsis thaliana* through a mechanism involving homologous recombination. XIIIth Intl. Cong. of Virology. July 23-28. San Francisco, California.

Tomlinson, J. A., Walker, V. M. 1973. Further studies on seed-transmission in the ecology of some aphid-transmitted viruses. Ann. App. Biol. 73:292-298.

CHAPTER FIVE

HITCHHIKING BY A CAULIMOVIRUS: AN ENDOGENOUS PLANT PARARETROVIRAL SEQUENCE OF A NEW AND DISTINCT *CAULIMOVIRUS*

ABSTRACT

Family *Caulimoviridae* consists of several genera of plant viruses that share a unique genome structure, organization and replication strategy among known plant viruses. Considered as plant pararetroviruses, these viruses exist as RNA in their hosts but code for a reverse transcriptase to synthesize and package dsDNA in progeny virions. Viruses in certain genera such as *Badnavirus*, *Cavemovirus, Petuvirus* in this family were shown to integrate their genomic sequences into their host genomes and exist as endogenous pararetroviral sequences. However, members of the genus *Caulimovirus* remained to be an exception and are believed to exist only as episomal elements during the virus infection cycle. We present evidence that the DNA genome of a new and distinct species in the genus *Caulimovirus* is integrated into its host genome, dahlia (*Dahlia pinnata*). Using cloned genes of the virus as probes, Southern blot hybridization of total plant DNA from dahlia seedlings showed the presence of viral DNA in the host DNA whereas, viral DNA was not detected in other known hosts of DMV. This is the first report of natural integration of a caulimovirus genome into its host and existence of a caulimovirus species as an endogenous pararetroviral sequence in a host plant.

INTRODUCTION

Integration of viral DNA into bacterial and animal host chromosomes is a common occurrence. In contrast, there are very few reported cases of integration of viral sequences in the plant genome. The integrated viral sequences that have been described so far were all derived from viruses with DNA genomes or from viruses that have a DNA phase in their replication cycle. These include the single-stranded DNA geminivirus (Ashby *et al*., 1997; Bejarano *et al.,* 1996) and the double- stranded DNA caulimoviruses (Harper *et al.,* 2002). The family *Caulimoviridae*, consisting of six genera, contains plant viruses that replicate by reverse transcription. Four viruses representing three of the genera belonging to family *Caulimoviridae* are reported to have integrants in the host plant genome. These include *Banana streak virus* (BSV) in *Musa* spp. (Ndowora *et al.,* 1999; Harper *et al.,* 1999), *Petunia vein clearing virus* (PVCV) in *Petunia* spp. (Harper *et al.,* 2003; Richert-Pöggeler *et al.,* 2003; Richert-Pöggeler & Shepherd 1997), *Tobacco vein clearing virus* (TVCV) in *Nicotiana* spp. (Gregor *et al.,* 2004; Jakowitsch *et al.,* 1999; Lockhart *et al.,* 2000) and *Rice tungro bacilliform virus* (RTBV) in *Oryza* spp. (Kunnii *et al.,* 2004) genomes. Although integration events were observed for these viral sequences, the viruses involved do not encode an integrase function. The integration of viral sequences is hypothesized to have occurred by illegitimate recombination (Jakowitsch *et al.,* 1999; Kunnii *et al.,* 2004; Lockhart *et al.,* 2000).

Of the six genera in the family *Caulimoviridae*, genus *Caulimovirus* has seven viruses that include the most extensively studied *Cauliflower mosaic virus*. There are no reports of integrated viral sequences representing any of the viruses in the genus *Caulimovirus*. We present

79

evidence that the DNA genome of a new and distinct species in the genus *Caulimovirus* is integrated into its host genome, dahlia (*Dahlia pinnata*). *Dahlia mosaic caulimovirus* (DMV) is an important viral pathogen of dahlia (*Dahlia pinnata*) in the US and several parts of the world. DMV is a member of the family *Caulimoviridae,* genus *Caulimovirus* with a circular double stranded-DNA genome of approximately 7 kb. The physical map of a caulimovirus (DMV-Portland) genome associated with dahlia mosaic was reported (Richins & Shepherd 1983), and its genomic sequence representing six open reading frames is available in GenBank. DMV-Portland appears to be a typical species of genus *Caulimovirus* based on its genome organization. Studies on the incidence of DMV-Portland in dahlia samples from the US and Europe revealed 21% and 53% of the respective samples tested to be infected (Pahalawatta *et al.,* 2007b). The symptomatology, propagative hosts and the role of various aphid species in DMV-Portland transmission have been reported (Albouy, *et al.,* 1992; Brierley & Smith; 1950; Brunt, 1971). However, our surveys of dahlias to determine the incidence of DMV revealed the overwhelming presence of another caulimovirus, designated DMV-D10 that was found to be distinct from DMV-Portland (Pahalawatta *et al.,* 2007b; Pahalawatta *et al.,* 2005; Pappu *et al.,* 2005) . A partial sequence representing a D10 like sequence variant, DMV-Holland (Nicolaisen, 2003) has also been reported.

The most characteristic symptoms of the disease include mosaic and vein-banding accompanied by stunting and leaf distortion. Symptoms vary depending on the cultivar, and cultivars with few or no symptoms are common. A survey of dahlia leaf samples, using a PCR-based assay, collected from different dahlia cultivars from the US, the Netherlands, Czech Republic and New Zealand detected the presence of DMV-D10 specific sequences in almost

80

100% of the samples tested irrespective of symptom expression (Pahalawatta *et al.,* 2007b). Further tests revealed the presence of DMV-D10 sequences in all plant parts tested including leaves, roots, seeds, flower petals and pollen. We also detected the presence of DMV-D10 sequences in tissue culture-derived dahlia plants. Our studies have also shown that DMV-D10 is transmitted through seed to 100% of the progeny plants (Pahalawatta *et al.,* 2007a). Studies on the genome structure and organization of DMV-D10 indicated that the virus differs from known caulimoviruses in that the aphid transmission factor was absent and only a partial coat protein was present. Results of these initial studies indicated potential integration of DMV-D10 viral sequences into the plant host genome.

MATERIALS AND METHODS

Plant material and DNA extractions: Total genomic DNA was extracted from dahlia seedlings, germinated in covered Petri dishes to avoid exposure to the aphid vector, which has been previously reported to transmit the virus in a non- persistent manner. To determine the presence of the virus in dahlia seeds, seeds were soaked in water and the seed coat was separated from the cotyledons. The separated seed coats and cotyledons were used for DNA extraction. Genomic DNA was also extracted from dahlia roots. In addition, total genomic DNA was extracted from leaves of *Zinnia elegans*, *Verbesina encelioides* and potato plants. Total genomic DNA extractions were made using the DNeasy plant mini kit (Qiagen Inc, Valencia, CA) according to manufacturer's instructions. DNA was quantified using the fluorescent DNA quantitation kit with Hoechst 33258 dye (Bio-Rad, Hercules, CA). Fluorescence readings were taken on a Spectra Max M2 fluorometer (Molecular Devices Corp., Sunnyvale, CA). Dahlia

seedling DNA extracts were digested overnight at 37°C with EcoRI.

Southern blot hybridization

Probe preparation: DMV-D10 specific primer pairs V416 (5'-TTA CCA CTT CTA ACA AAA GG- 3')/ C1070 (5'- TCC ATA CAT GTC TAC TTT TTC GG -3'); ORF1-START (5'-ATG GAT CGT TAA AGA TT -3')/ ORF1-END1 (CTG TTT TTC TGT GTT TCT ACT GG -3') and ORF4 CP1-1431F (5'- TGC ATA AAA TGA GTT CTA TC -3')/ ORF4 CP1-1926C (5'-TGA ACT TGT TCA TCA TTA TC -3') were used in PCR amplification using DNA extracts from DMV-D10 infected dahlia plants as template. The PCR reaction conditions were same as described in previous chapters. The PCR products were purified using the Qiaquick PCR purification kit (Qiagen Inc, Valencia, CA). Purified PCR products (25ng) were [32] P labeled using the Prime-a-gene® labeling kit (Promega, Madison, WI).

Spin column purification of the labeled probe: The bottom of a 1.0 ml tuberculin syringe was plugged with siliconized glass wool. Syringe was loaded with Sephadex G-50-300 equilibriated with STE buffer (10 mM Tris pH 8.0, 100 mM NaCl, 1mM EDTA pH 8.0). Syringe was set into a 15 ml conical tube and centrifuged at 1600 x g (Beckmann Coulter centrifuge; Ti 70.1 rotor) for 4 min. Eluent was collected in a de-capped Eppendorf tube. Sephadex was added to the column up to 0.9 ml, topped twice with 0.1 ml STE and centrifuged at 1600 x g for 4 min. Labeled DNA in STE (0.1 ml) was added to the column and centrifuged again at 1600 x g for 4 min. Purified labeled DNA was collected in an Eppendorf tube. One microliter of unspun

labeled probe solution was pipette into 9µl of water and the radioactivity recorded. Similarly, 1µl of spun eluent , which contains only labeled DNA, was added to 9µl water and the radioactivity recorded again. The two values were used to determine the incorporation percentage of the radioactivity.

Gel electrophoresis: DNA extracts (10-20µg) were loaded on 0.8% agarose gels. Electrophoresis was performed using 1X TAE buffer running at 20V overnight. Gels were stained with ethidium bromide and photographed. After photographing each gel was washed in 0.25N HCl for 15 min to depurinate the DNA to allow for the transfer of high molecular weight DNA fragments. Next, the gels were washed in deionized water followed by solution T (0.4M NaOH containing O.6M NaCl) on a shaking platform for 15 min.

Alkaline Southern transfer: A 30x15 cm glass dish with a 15x15 cm glass plate supported in the center was used to set up the transfer apparatus. The glass dish was filled with 1L solution T. A Whatman 3MM filter paper (30x15 cm) was used as a wick. The filter paper was soaked in solution T and draped over the glass plate with the two edges dipped in solution T. Six pieces of Whatman 3MM filter paper and one piece of Hybond N+ membrane (GE Healthcare Bio-Sciences Corp, Piscataway, NJ) were cut to the size of the gel. Three pieces of filter paper were soaked in solution T and placed on the glass plate. The agarose gel was placed on the filter papers, wells facing down. The Hybond N+ membrane was soaked in sterilized deionized water and placed on the gel. The remaining three filter papers were soaked in solution T and placed on top of the Hybond N+ membrane. A glass rod was used to remove air bubbles between each

layer of filter paper. A 12 cm wad of paper towels was placed on the filter papers and a 250g

weight was kept on the paper towels. Plastic wrap was used to cover the entire area of the glass

dish leaving only the surface with the gel exposed. DNA was transferred overnight onto the

membrane. The next day the transfer apparatus was disassembled and the membrane air dried on

filter paper for 2 h at 37°C.

Prehybridization: Membrane was soaked in 500ml 2X SSC [100ml 20X SSC (3M NaCl, 0.3M

sodium citrate, pH 7.0) in 900ml ddH$_2$O] 2X for 1 min and the excess moisture blotted off with

filter paper. Next the membrane was soaked in 30ml prehybridization solution (18ml deionized

water, 6ml 5X HSB, 3ml Denhardt's III, 3ml freshly boiled salmon sperm DNA 5mg/ml). 5X

HSB was an aqueous solution of 3M NaCl, 100mM PIPES (1,4-Piperazinediethanesulfonic

acid), 25mM Na$_2$EDTA with a pH of 6.8 adjusted with 4M NaOH. Denhardt's III was an

aqueous solution of 2% bovine serum albumen fraction V, 2% Ficoll-400, 2% Polyvinyl

pyrrolidone, 10% SDS and 1% Na$_4$P$_2$O$_7$. 10H$_2$O. Prehybridization was done overnight in an

orbital shaker (60 rpm) at 65°C.

Hybridization:

The labeled probe was added into a 50ml polypropylene tube containing 8.5ml deionized water,

3.0ml 5X HSB, and 1.5ml Denhardt's III solution. Contents were prewarmed to 65°C. The tube

was placed in a boiling water bath for 7 min to denature. The membrane was removed from the

prehybridization solution allowed to drip and placed in the hybridization solution and allowed to

shake (60 rpm) overnight at 65°C.

Stringency washes: Membrane was washed for 1 min in rinse solution A (2X SSC, 1% SDS). The solution was poured out, replaced and shaken at 75 rpm on an orbital shaker for 30 min at 65°C. Next, rinse solution A was replaced with rinse solution B (1X SSC, 1% SDS) and shaken at 75 rpm for 30 min at 65°C. Finally, rinse solution B was replaced with rinse solution C (0.2X SSC, 1% SDS) and shaken at 75 rpm for 30 min at 65°C. Excess moisture on the membrane was blotted with filter paper and the membrane was wrapped in saran wrap.
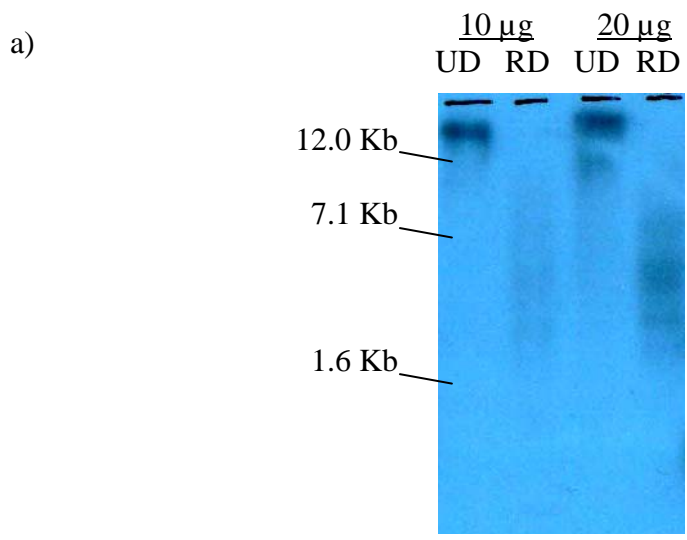
Autoradiography: The wrapped membrane was placed in a cassette with photographic paper (Kodak ™ X-Omat AR, 20X25 cm) and an intensifying screen. The cassette was closed firmly and kept at -70°C for 4-7 days depending on the level of radioactivity prior to developing the film.

RESULTS

Hybridization was observed with both undigested and restriction endonuclease-digested dahlia genomic DNA extracted from seedlings, seed coat as well as roots using DMV-D10 specific probes (Fig. 1a & b). The hybridization signal for DNA extracts from dahlia seed coat was very weak, possibly due to the difficulty of obtaining quality genomic DNA. However, agarose gel electrophoresis of PCR products amplified using DMV-D10 specific primer pairs and dahlia seed coat extracts as template indicated the presence of DMV-D10 sequences in dahlia seed coats (Fig. 2). None of the hybridization signals were seen to occur at ~7 kb, which is expected for genomic DNA of DMV-D10. In contrast, all of the hybridization signals from

the undigested DNA occurred at very high molecular weights which is expected for dahlia genomic DNA (Fig. 1a & b). No hybridization was detected with genomic DNA samples from hosts susceptible to DMV or from unrelated plant taxa (Fig. 1c).

a)

10 µg   20 µg
UD RD UD RD

12.0 Kb

7.1 Kb

1.6 Kb

b)

C DS SC R P

12 Kb

7 Kb

0.6 Kb

87

c)

C DS V Z

12 Kb

7 Kb

0.48 Kb
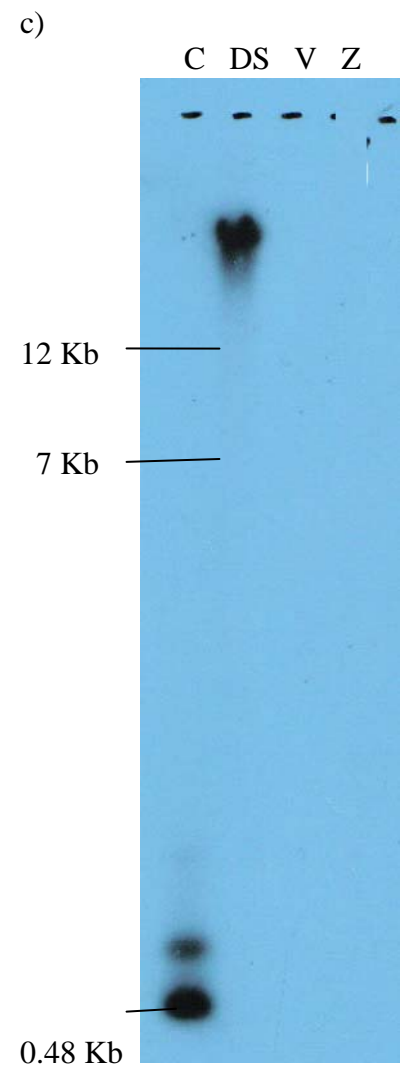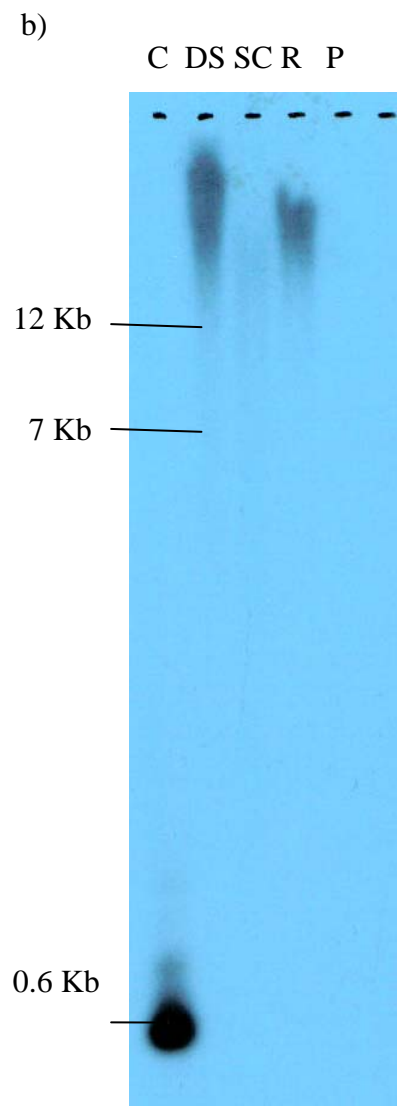
Figure 1. Southern blot hybridization using (a) probe representing open reading frame I (ORF I) of *Dahlia mosaic virus* (DMV)- D10. Total genomic DNA extracts from dahlia seedlings UD – undigested; RD – digested with EcoRI (b) probe representing ORF IV of DMV-D10. Total genomic DNA extracts from DS- dahlia seedlings; SC – seed coat; R – root; P – potato; C – control (PCR product used to produce the labeled probe). (c) probe representing open reading frame I (ORF I) of *Dahlia mosaic virus* (DMV)– D10. Total genomic DNA extracts from DS- dahlia seedlings; V – Verbesina; Z – Zinnia.
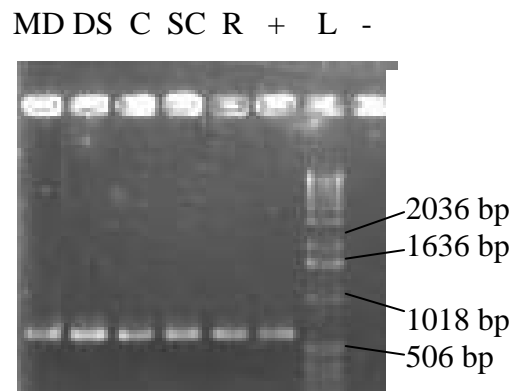
Figure 2. Agarose gel electrophoresis of PCR products from DNA extracts of dahlia plant components. PCR reaction was carried out using primer pairs representing open reading frame (ORF) I of *Dahlia mosaic virus* (DMV)- D10. MD – mature leaves; DS - seedlings; SC – seed coat; c – cotyledons; R – roots; L – 1Kb ladder; + - positive control; - negative control.

DISCUSSION

Several lines of evidence including the presence of D10 sequences in nearly 100% of the samples tested from several parts of the world, 100% seed and pollen transmission, combined with the Southern analysis support the conclusion that this caulimovirus exists as an endogenous pararetrovirus sequence. It is important to differentiate between endogenous and episomal viral sequences since the latter results in symptom expression whereas the endogenous viral sequences may or may not be capable of causing infection and symptom expression. It remains to be seen if D10 exists, in addition to as an endogenous pararetroviral sequence, as an episomal element in infected plants. In our studies on the genome structure and organization of DMV-D10, we were unable to detect the presence of a gene that is homologous to the aphid transmission factor that is coded by all known species in the genus *Caulimovirus*. Furthermore, the D10 genome coded for a partial coat protein. Therefore, it is possible that DMV-D10 exists primarily as an endogenous sequence. However, dahlia plants positive only for DMV-D10 sequences exhibited a range of disease symptoms ranging from mild to severe mosaic. Therefore it is highly probable that DMV-D10 may exist as infectious episomal genomes. Studies on other integrated viruses have shown that the expression of episomal infections and the degree of severity of symptoms is correlated with environmental conditions such as day length and temperature (Harper *et al.,* 2002). Similar environmental factors may play a role in the expression of infectious episomal DMV-D10 genomes.

Studies on *Musa,* and *Oryza* genomes have resulted in the identification of the junctions of virus sequence integration (Ndowora *et al.,* 1999; Kunnii *et al.,* 2004) due the availability

of genomic libraries for these plant hosts. However similar resources are not yet available for the dahlia genome. Clonal propagation of vegetative tissues, stem cuttings and bulbs have been considered to be the primary reason for the spread of viruses in dahlias. The most effective means for managing viruses in vegetatively propagated plants is through testing and elimination of virus infected material and use of virus-free stock for propagation. Meristem tip culture was used to eliminate viruses from infected stock. However, in instances where the viral genome is integrated into the host genome, a similar approach is not feasible. Under these circumstances, knowledge of the environmental conditions that trigger symptom development will play an essential role in developing virus management strategies.

REFERENCES

Albouy, J., Lemattre, M., Wang, W. C., Mevel, A. 1992. Production of pathogen tested dahlia in France. Acta Hort. 325: 781-784.

Ashby, M. K., Warry, A., Bejarano, E. R., Khashoggi, A., Burrell, M., Lichtenstein, C. 1997. Analysis of multiple copies of geminiviral DNA in the genome of four closely related Nicotiana species suggest a unique integration event. Plant Mol. Biol. 35: 313-321.

Bejarano, E. R., Khashoggi, A., Witty, M., Lichtenstein, C. 1996. Integration of multiple repeats of geminiviral DNA into the nuclear genome of tobacco during evolution. Proc. Natl. Acad. Sci. 93: 759-764.

Brierley, P., Smith, F. F. 1950. Some vectors, hosts and properties of dahlia mosaic virus. Plant Dis. Rep. 34: 363-370.

Brunt, A. A. 1971. Some hosts and properties of dahlia mosaic virus. Anal. App. Biol. 67: 357-368.

Gregor, W., Mette, M. F., Staginnus, C., Matzke, M. A., Matzke, A. J. M. 2004. A distinct endogenous pararetrovirus family in Nicotiana tomentosiformis, a diploid progenitor of polyploid tobacco. Plant Physiol. 134: 1191-1199.

Harper, G., Hull, R., Lockhart, B., Olszewski, N. 2002. Viral sequences integrated into plant genomes. Annu. Rev. Phytopathol. 40: 119-136.

Harper, G., Osuji, J. O., Heslop-Harrison J. S., Hull, R. 1999. Integration of banana streak badnavirus into the *Musa* genome: molecular and cytogenetic evidence. Virology 255:207-213.

Harper, G., Richert-Pöggeler, K. R., Hohn, T., Hull, R. 2003. Detection of petunia vein-clearing virus: model for the detection of DNA viruses in plants with homologous endogenous pararetrovirus sequences. J. Virol. Methods 107:177-184.

Jakowitsch, J., Mette, M. F., van der Winden, J., Matzke, M. A., Matske, A. J. M. 1999. Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. Proc. Natl. Acad. Sci. USA 96: 13241-13246.

Kunii, M., Kanda, M., Nagano, H., Uyeda, I., Kishima, Y., Sano, Y. 2004. Reconstruction of putative DNA virus from endogenous rice tungro bacilliform virus-like sequences in the rice genome:implications for integration and evolution. BMC Genomics 5:80.

Lockhart, B. E., Menke, J., Dahal, G., Olszewski, N. E. 2000. Characterization and genomic analysis of tobacco vein clearing virus, a plant pararetrovirus that is transmitted vertically and related to sequences integrated in the host genome. J. Gen. Virol. 81:1579-1585.

Mette, M. F., Kanno, T., Aufsatz, W., Jakowitsch, J., van der Winden, J., Matzke, M. A., Matzke, A. J. M. 2002. Endogenous viral sequences and their potential contribution to heritable virus resistance in plants. EMBO J. 21**:** 461-469.

Ndowora, T., Dahal, G., LaFleur, D., Harper, G., Hull, R., Olszewski, N. E., Lockhart, B. 1999. Evidence that badnavirus infection in *Musa* can originate from integrated pararetroviral sequences. Virology 255: 214-220.

Nicolaisen, M. 2003. Partial molecular characterization of *Dahlia mosaic virus* and its detection by PCR. Plant Dis. 87:945-948.

Pahalawatta, V., Druffel, K ., Pappu, H. R. 2007a . Seed Transmission of Dahlia mosaic virus in Dahlia pinnata. Plant Dis. 91: 88-91.

Pahalawatta, V., Druffel, K., Pappu, H. R. 2005. Molecular characterization of *Dahlia mosaic virus* genome. Phytopathology 95:S79.

Pahalawatta, V., Miglino, R., Druffel, K. L., Jodlowska, A., van Schadewijk, A. R ., Pappu, H. R. 2007b. Incidence and relative distribution of distinct Caulimoviruses Genus *Caulimovirus*, Family *Caulimoviridae* Associated with Dahlia Mosaic in *Dahlia pinnata*. Plant Dis: In press.

Pappu, H. R., Wyatt, S. D., Druffel, K. L. 2005. DMV: Molecular detection and distribution in dahlia in the US. HortScience 40:697-699.

Richert-Pöggeler, K. R., Noreen, F., Schwarzacher, T., Harper, G., Hohn, T. 2003. Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia. EMBO J. 22**:** 4836-4845.

Richert-Pöggeler, K. R., Shepherd, R. J. 1997. Petunia vein clearing virus: a plant pararetrovirus with core sequences for an integrase function. Virology 236: 137-146.

Richins, R. D., Shepherd, R. J. 1983. Physical maps of the genome of *Dahlia mosaic virus* and *Mirabilis mosaic virus* – two members of the Caulimovirus group. Virology 124: 208-214.

CHAPTER FIVE

GENERAL CONCLUSIONS

Dahlia mosaic virus (DMV) has been reported to be one of the most widespread viruses affecting dahlias. The disease caused by the virus on dahlias is characterized by symptoms of mosaic and vein banding accompanied by stunted plant growth, affects the flower quality and hence the market value. Previous research on *Dahlia mosaic virus* has resulted in identifying partial genomic sequence, referred to here as DMV- Portland and DMV-Holland, with little sequence homology. DMV-Portland has also been determined to be transmitted mechanically as well as by aphids. Several plant species including *Zinnia elegans* and *Verbesina encelioides* have been reported to be suitable propagation hosts although *Dahlia* species are the natural hosts for the virus. Furthermore, the sequence of DMV-Portland closely resembles the genome organization of *Cauliflower mosaic virus* (CaMV), in containing 7 open reading frames (ORFs) coding for proteins with similar functions. Subsequent research on the disease revealed the presence of another viral sequence referred to as DMV-D10.

In this study, the complete sequence of DMV-D10 was determined and used to elucidate the genome structure and organization of the virus. The circular double-stranded DNA genome (7046 bp) of DMV-D10 had a genome organization very similar to that of DMV- Portland and CaMV. However, the genome of DMV-D10 differed from the above two viruses in the absence of an aphid transmission factor (ATF) and a complete coat protein. It is possible that the loss of an ATF is an indication of evolutionary adaptation of the virus due to continuous vegetative propagation. On the contrary it is possible that eventhough the DMV-D10 amino acid sequence

94

does not show any homology to ATFs in other caulimoviruses, that it plays a role in aphid transmission. Complementation studies involving the ATF of DMV-Portland and ORF II of DMV-D10 will be useful in determining whether ORF II of DMV-D10 has a functional role in aphid transmission. It is also possible to use aphids to determine the transmission characteristics of DMV-D10. The partial coat protein in DMV-D10 is unique due to the absence of a zinc finger motif ($C-X_2-C-X_4-H-X_4-C$). This motif is a well-conserved sequence among retroviruses. It is also present in pararetrovirus coat proteins and has been reported to be essential for replication. It is hypothesized that CaMV utilizes the zinc finger domain to mediate encapsidation of the pregenomic RNA leader prior to reverse transcription. Studies have shown that the zinc finger motif is absent from mature CaMV particles and thus to have no function after reverse transcription and assembly. Lack of the zinc finger domain in DMV-D10 maybe an indication of the presence of an alternate method of virion assembly or the absence of an encapsidated virus. Phylogenetic analysis comparing the different ORFs of viruses in Family *Caulimoviridae* revealed that DMV-D10 is grouped with viruses belonging to Genus *Caulimovirus* and is closely related to DMV-Portland and *Figwort mosaic virus* (FMV). According to the 8th report of the International Committee on Taxonomy of Viruses (ICTV), one of the species demarcation criteria in Genus *Caulimovirus* is differences in polymerase (reverse transcriptase + ribonuclease H) nucleotide sequences of more than 20%. The percentage identity of the DMV-D10 sequence with that of DMV-Portland and FMV was 54.4% and 70.1% respectively. Thus we can categorically state that DMV-D10 is a new virus species in the Genus *Caulimovirus* affecting dahlias.

Our studies revealed the high incidence of mixed infections in dahlia plants. However

potential interactions between DMV-D10, DMV-Portland and DMV-Holland that may affect symptom expression in dahlia plants were not examined. Furthermore the correlation between dahlia varieties and symptom expression also needs to be considered under conditions of single and multiple infections to reveal possible synergistic interactions between the three viruses.

The ubiquitous nature of DMV-D10 in samples from different states in the US as well as from the Netherlands, Czech Republic and New Zealand and the presence of DMV-D10 sequences in dahlia leaves, flower petals, pollen and seed instigated the study to determine whether DMV-D10 exists as an endogenous pararetrovirus. Southern blot hybridization studies revealed the presence of DMV-D10 sequences in total genomic DNA extracts from dahlia seedlings but not in other hosts reported to be susceptible to the virus. Endogenous virus sequences have been reported for other pararetroviruses including *Banana streak virus, Petunia vein clearing virus, Tobacco vein clearing virus* and *Rice tungro bacilliform virus*. However this study is the first instance in which a virus belonging to Genus *Caulimovirus* has been reported to be introgressed into the host genome. At present, PCR- based assays are considered to be the most reliable method of testing for the viruses associated with dahlia mosaic. However, in the presence of endogenous viral sequences, PCR based on total genomic DNA templates will result in the detection of integrated sequences as well as episomal virus particles. It is important to be able to distinguish between episomal virus and endogenous viral sequences especially in plant material that is used for breeding purposes. Purification of virus particles from the total plant genomic DNA extract prior to PCR can be used to determine the presence/absence of episomal virus particles. Furthermore, multiplex PCR using both virus specific and dahlia specific primer pairs will help in determining the purity of the virus extract as well as the presence/absence of

96

virus particles in the plant. Another technique may be to employ Southern blot hybridization using virus specific probes. However, the challenge is to design a convenient technique that is capable of making the distinction between endogenous and episomal virus sequences, that can be implemented in commercial settings. Multiple attempts to mechanically inoculate dahlia, zinnia and verbesina plants using DMV-D10 proved unsuccessful. This suggests that the virus is most probably present in the plant as endogenous sequences instead of episomal infections. This phenomenon is also a possible explanation for the absence of an aphid transmission factor as well as the zinc finger motif since the functions of the protein products in virus transmission and encapsidation will be redundant. However, dahlia plants positive only for DMV-D10 sequences exhibited a range of disease symptoms ranging from mild to severe mosaic. Therefore it is highly probable that DMV-D10 also exists as infectious episomal genomes. Aphid transmission tests as well as graft transmission studies will assist in determining the presence of episomal virus particles. In this study we used primers in PCR that could potentially amplify the 7 kb DMV-D10 sequence to test for episomal virus without any success. However the presence/absence of episomal virus genomes cannot be determined on the basis of the PCR results alone due to limitations in Taq polymerase, purity of template DNA and other unknown inhibitory factors.

At present, information on the dahlia genome is very limited. Unlike in other host integrated viruses, where the site/s of virus integration into host genome have been detected due to the availability of plant host genomic libraries, the site/s of DMV-D10 integration in dahlia genomes is not known. Cytological studies using DMV-D10 specific fluorescent probes hybridized to dahlia chromosomes can be utilized to determine the dahlia chromosome/s in

which DMV-D10 sequences are integrated. It may be possible to isolate the exact position of DMV-D10 sequence integration by using a DMV-D10 specific primer coupled with a random hexamer in PCR by screening the resulting amplicons for the presence of plant sequences. It would also be of interest to determine when such an integration event occurred during the evolution of dahlia species and DMV. Molecular genetic studies on wild dahlia species may provide important clues to these questions.

Clonal propagation of vegetative tissues, stem cuttings and bulbs in the case of dahlia, has been considered to be the primary reason for the spread of viruses in dahlias. Meristem tip culture has been used to eliminate disease from propagation stock that is infected. However in the event that DMV-D10 is integrated into the host genome a similar approach is not feasible. Testing for viral sequences and the use of virus-free stock would be the most effective methods of limiting the spread of DMV-D!0. However, control of DMV-Portland and DMV-Holland may also depend on control of the aphid population where applicable. It has been observed that many of the infected dahlia plants are asymptomatic. One approach that maybe used is to determine whether specific environmental conditions trigger the appearance of symptoms. A similar phenomenon has been observed in other host integrated viruses where certain stress conditions have resulted in the appearance of symptoms. Therefore it is possible that cultural conditions play a role in the appearance of symptoms. Knowledge of the environmental conditions that trigger symptom development may be useful in maintaining flower quality.