BIOLOGICAL AND MOLECULAR STUDIES ON PLANT PARA-RETROVIRUSES ASSOCIATED WITH *DAHLIA* SPP. IN NATURAL AND MANAGED ECOSYSTEMS

By

SAHAR EID

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

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Plant Pathology

AUGUST 2010

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation/thesis of SAHAR EID find it satisfactory and recommend that it be accepted

_____

Hanu Pappu, Ph.D., Chair

_____

Alex Karasev, Ph.D.

_____

Amit Dhingra, Ph.D.

_____

Scot Hulbert, Ph.D.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude and appreciation to my major advisor Dr. Hanu Pappu for his incessant encouragement and guidance which enormously contributed to this work. Dr. Pappu was always present to listen to me and to give me advice and encouragement to be persistent to accomplish my goals. Grateful appreciation goes to Dr. Amit Dhingra, Dr. Scot Hulbert and Dr. Alex Karasev for their critical reviews for this work. I am grateful to Dr. Dayle Saar for providing the wild *Dahlia* species samples and to Dr. Amit Dhingra for his help to conduct Southern and Northern hybridization experiments in his laboratory.

Most of all, I am thankful for the help and support of my family back home in Lebanon, my fiancé's support from Canada, and my extended family here in Pullman including my advisor, the lab group, and my fellow graduate students in the department. I would like to extend my gratitude to the faculty and staff of the Department of Plant Pathology for all the help and support during my graduate studies at Washington State University. A special thank you to Dr. Tim Murray (former Department Chair) and Dr. Lori Carris (Graduate students coordinator) for arranging financial support and help to attend the 2008 and 2009 annual meetings of the American Phytopathological Society (Nellie Heald Memorial Scholarship, and Graduate School block grant).

I would like to thank the Samuel and Patricia Smith Endowment for Dahlia Virus Research, created by the American Dahlia Society for the financial support for this study, the USDA Northwest Nursery Crop Research Center and the Washington State Commission for Pesticide Registration for funding this work, and a part-time Ph.D. Graduate Research Assistantship from Agricultural Research Center, WSU.

BIOLOGICAL AND MOLECULAR STUDIES ON PLANT PARA-RETROVIRUSES ASSOCIATED WITH *DAHLIA* SPP. IN NATURAL AND MANAGED ECOSYSTEMS

ABSTRACT

By SAHAR EID, Ph.D.
Washington State University
AUGUST 2010

Chair: Hanu Pappu

Three distinct plant para-retroviruses belonging to the genus *Caulimovirus* were found to be associated with dahlia (*Dahlia variabilis)*: *Dahlia mosaic virus* (DMV), DMV-D10, and Dahlia common mosaic virus (DCMV). DMV-D10 was the most prevalent (94% of tested samples) followed by DCMV (48.5%) and DMV (23%) in cultivated dahlias during surveys conducted in 2007 and 2008. Biological studies using *Verbesina encelioides* as the assay plant showed that DMV and DCMV were transmitted by the aphid, *Myzus persicae*. Mechanical inoculation of *V. encelioides* showed that DMV and DCMV are transmitted 58% and 36%, respectively. DMV-D10 was not transmissible by either aphids or mechanical inoculation.

Selected wild *Dahlia* species (*D. coccinea, D. tenuicaulis, D. rupicola,* and *D. sherffii*) in their natural habitats from west-central Mexico were tested for the three caulimoviruses. Results showed that the wild dahlia species contained only DMV-D10 in 91% of the samples (n=56). The structure and genome organization of DMV-D10 from three wild species, *D. coccinea*

(D10-DC), *D. sherffii* (D10-DS) and *D. tenuicaulis* (D10-DT) were determined and compared to that of D10-US (DMV-D10) from cultivated species (*D. variabilis*). The complete ca. 7 kb dsDNA genomes of D10-DC, D10-DS, and D10-DT had the structure and organization typical of a *Caulimovirus* species and shared 89.3 to 96.6% amino acid sequence identity among various ORFs when compared to those of D10-US from *D. variabilis.*

DMV-D10 exists as an endogenous plant pararetorviral sequence in cultivated dahlias, contrary to DMV and DCMV which are typical members of the genus *Caulimovirus* and considered to exist as episomal molecules in infected cells. Genomic Southern hybridization showed differences of integration frequency between different dahlia cultivars. RT-PCR results indicated that both RNA transcripts (35S and 19S) of D10 could be present in infected plants. The importance of these integrated viral sequences in inducing disease and the conditions that influence the induction of expression of the viral genes and symptoms remain to be seen. The discovery of plant para-retroviruses in wild dahlia species in their natural habitats could potentially provide important insights into the possible emergence, co-existence and co-evolution of para-retroviruses and their host plants.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER SIX

LIST OF FIGURES

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

DEDICATION

This thesis is dedicated to my family for their continuous love and unconditional support

Also to my fiancé Rony Chamoun for his love and encouragement

I promise you that I will continue to work hard and make you proud of me all the time

CHAPTER ONE

BIOLOGICAL AND MOLECULAR STUDIES ON PLANT PARA-RETROVIRUSES ASSOCIATED WITH *DAHLIA SPP*. IN NATURAL AND MANAGED ECOSYSTEMS

GENERAL INTRODUCTION

Dahlia is a perennial plant and its discovery was first reported in 1577 by Francisco Hernandez from Mexico. Dahlia plants have been introduced to Europe from Mexico and spread at the beginning of 19th century. In 1791, Antonio Cavanilles was able to flower three plants and named them *Dahlias* in memory of Andres Dahl, Author of "Observationes Botanicae" (Damp, 1981). At the present time, the genus *Dahlia* consists of 35 species and are primarily from Mexico (Saar and Sørensen 2000, 2005; Saar 2002; Saar et al. 2002, 2003a; Hansen and Sørensen 2003). Dahlia plants are easily propagated by cuttings, tubers, and seeds (Dole and Wilkins, 1999). Newly registered cultivars are becoming available every year (The International Register of Dahlia Names 1969 et seq.) and this is due to cross and make hybrids. *Dahlia pinnata* named by Cavanilles doesn't exist at present while *D. variabilis* (named for its color similar to the scarlet dye obtained from the insect Cocchineal) is the cultivated dahlias worldwide (Saar et al., 2003b). *D. variabilis* is thought to be a cross between *D. coccinea* and at least one other species (Hansen and Hjerting, 1996; 2000; Lawrence 1931). A recent study by Saar et al. (2003b) placed *Dahlia* species in four sections: *Pseudodendron, Epiphytum, Entemophyllon*, and *Dahlia* and they showed that there are three distinct clades: the variable root clade, the core *Dahlia* clade and the section *Entemophyllon* as very well-defined clade based on the phylogenetic analysis using the internal and external transcribed spacer regions. The

cultivated dahlias plus the four wild species used in this research belong to three different sections and clades (Table 1) according to Saar et al. (2003b).

Table 1. Distribution of *Dahlia* species used in this research between clades and sections

Dahlia *spp*.	**Clade**	**Section**
D. variabilis	Core *Dahlia*	*Dahlia*
D. coccinea	Core *Dahlia*	*Dahlia*
D. tenuicaulis	Variable root	*Pseudodendron*
D. sherffii	Variable root	*Dahlia*
D. rupicola	Sect *Entemophyllon*	*Entemophyllon*

Besides the 35 species of *Dahlia* and the many thousands of named cultivars of *D. variabilis*, the chromosome numbers in *Dahlia* are also variable. The basic number of chromosomes or chromosome set (monoploid genome x) varies between 16, 17 and 18 while species of x =16 can have 2n = 32 or 64 (Gatt et al., 1998; 1999; Strother and Panero, 2001). The 2C-values (DNA content of haploid genome with chromosome number 2n) ranged from 3.30 pg in *D. dissecta* (2n = 34) to 9.62 pg in a *D. variabilis* (2n = 64) (Temsch et al., 2008). It was found that man-made interspecific hybrids had C-values close to those from their parents.

There are many challenges facing dahlia production but most important ones are viral diseases. These include *Tobacco streak virus* (*Ilarvirus*, RNA virus), *Tomato spotted wilt virus* (*Tospovirus*, RNA virus), *Cucumber mosaic virus* (*Cucumovirus*, RNA virus) and *Dahlia mosaic virus* (DMV, *Caulimovirus*, DNA virus) (Albouy, 1992). DMV is considered the most prevalent and important since symptoms can vary between cultivars and cases of asymptomatic infections

are common (Brunt, 1971; Albouy, 1995). The most characteristic symptoms are chlorotic spots and leaf malformation, mosaic, vein clearing, and systemic chlorosis.

DMV was first reported from Germany in 1928 by Brandenburg (Brunt, 1971). It belongs to the family *Caulimoviridae* and genus *Caulimovirus*. It has double-stranded DNA genome and six essential open reading frames coding for movement protein, aphid transmission factor, DNA binding protein, coat protein, reverse transcriptase, and inclusion body protein/ translational transactivator protein and a seventh open reading frame of unknown protein. The virus is composed of isometric particles, 48-50nm in diameter. The physical map of the DMV genomic DNA was reported by Richins and Shepherd in 1983. Two additional caulimoviruses DMV-D10 and DCMV were recently characterized and found to be associated with dahlia mosaic disease (Nicolaisen, 2003; Pappu et al., 2005; Pahalawatta et al., 2008a; Pappu et al., 2008). The dsDNA genomes of DMV-D10 and DCMV were cloned and sequenced and based on their genome organization, they belong to the genus *Caulimovirus* but based on their nucleotides and amino acid sequence homologies, they are considered distinct species (Fig. 1)

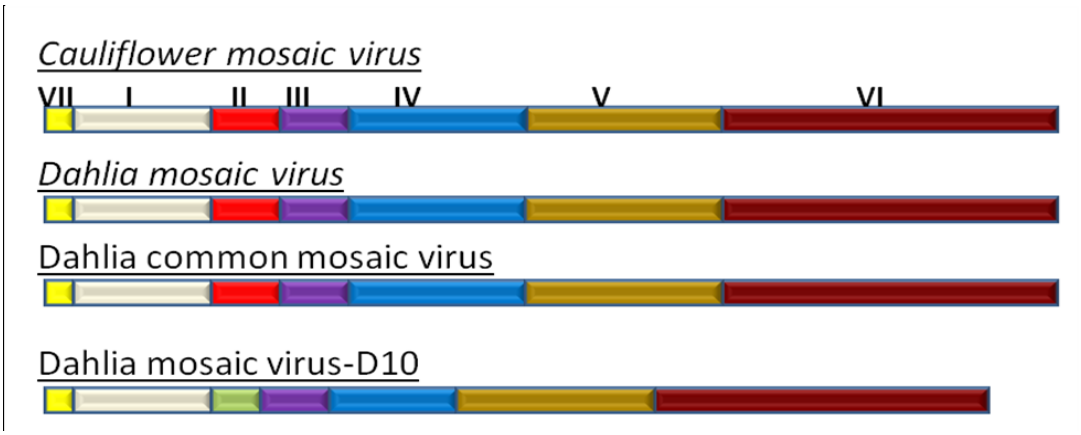


Fig. 1. Genome organization (shown as linear molecules) of the three caulimoviruses from dahlia compared to *Cauliflower mosaic virus* (type species of the genus *Caulimovirus*). Roman numerals represent the open reading frames (ORFs). VII – protein of unknown function; I- movement protein; II – aphid transmission factor; III- DNA binding protein; IV- coat protein; V – aspartic protease, reverse transcriptase, and ribonuclease H; VI – inclusion body / translational transactivator protein.

DMV-D10 shares limited sequence similarity with known caulimoviruses including DMV (Fig. 1) and it is the first caulimovirus reported to be integrated into the dahlia genome and thus exists as endogenous plant pararetroviral sequence (EPRS) (Pahalawatta et al., 2008b). Recently, two different systems of classification and nomenclature of endogenous viruses of the family *Caulimoviridae* have been proposed (Staginnus et al., 2009; Geering et al., 2010). They both agree on adding the term "endogenous" prior to virus species name and letter "e" to virus acronym to differentiate between integrated viral DNA and actively replicating virus. When no exogenous virus is known for the integrated viral DNA, Staginnus et al. (2009) proposed to add the plant species's (host) initials before the term EPRS. But Geering et al. (2010) suggested

including numerical code, a BAC address or a GenBank accession number until the genome sequence is available and then using specific code for locus of integration. We adopted the suggestion by Staginnus et al. (2009) for nomenclature and added the initials "Dv" for *Dahlia variabilis* before "EPRS" when referring to the endogenous DMV-D10 (DvEPRS).

DMV-D10 (DvEPRS) ORFII coding for an unknown protein without any homology to aphid transmission factor product of ORFII in caulimoviruses. In addition, DMV-D10 ORFIV (encoding for coat protein) was found to be truncated and in frame with ORFV (encoding for reverse transcriptase) (Pahalawatta et al., 2008a). Plant para-retroviruses (*Caulimoviridae*) replicate their genome via reverse transcription but integration into their host genome is not typical and they do not code for enzymes for integration and excision as retroviruses. However, endogenous pararetrovirus-related sequences (EPRS) are recently discovered in several plants including tobacco and other *Nicotiana* species (Jakowitsch et al., 1999; Lockhart et al., 2000; Gregor et al., 2004) , *Solanum tuberosum* (Hansen et al., 2005), *Musa* spp. (Harper et al., 1999; Ndowora et al., 1999; Geering et al., 2001; 2005a,b), *Petunia* spp (Richert-Poggeler et al., 2003), *Oryza* spp. (Kunii et al., 2004) and *Solanum lycopersicum* (Staginnus et al., 2007). It was shown that some EPRS (i.e. *Petunia vein clearing virus*, PVCV) could give rise to episomal particles and cause symptoms under certain conditions (Richert – Pöggeler et al., 2003). EPRS are not neutral components and they may contribute to virus evolution and pathogenicity. They may escape host control in interspecific hybrids and give rise to infectious viruses and they may participate in creating new viruses by recombination between integrants and invading viruses leading to probable new biological properties such as expanded host range or new symptoms (Hohn et al., 2008).

DMV can be transmitted by mechanical inoculation and aphid transmission. There were about 16 vector species reported by Brierley and Smith (1950) to transmit DMV whereas its principal vector is *Myzus persicae* (Sulz.). The presence of ORFII product is essential for viral transmission (Albouy, 1995). *Dahlia* is the natural host of DMV but Brunt (1971) showed that *Verbesina encelioides* and *Zinnia elegans* are the best choice as experimental hosts for DMV. There is a lack of information concerning these caulimoviruses evolution and virus-host relationships. Therefore, the overall goal of this study is to understand the biological and molecular basis of viral pathogenicity using the caulimoviruses complex and cultivated and wild *Dahlia* species as model system.

The specific objectives were:

1-Determine their incidence in cultivated and wild *Dahlia* species (Chapter 2 and 3).

2-Compare the biological properties of the three caulimoviruses associated with dahlia mosaic disease (Chapter 4).

3-Determine the genome structure and organization and characterization of these viruses in the wild species, and compare it to the sequences from cultivated species in order to understand the possible emergence, co-existence and co-evolution of these para-retroviruses and their host plants (Chapter 5).

4-Determine if the integrated DMV-D10 is expressed in cultivated species and exists as a functional integrant and the extent of its integration in *Dahlia variabilis* (cultivated species) genome (Chapter 6).

REFERENCES

Albouy, J. 1995. Dahlia, p.265-273. In: G. Lobenstein, R.H. Lawson, and A.A. Brunt (eds.). Virus and virus-like diseases of bulb and flower crops. John Wiley and Sons, Chichester, NY.

Albouy, J. and A. Mevel. 1992. Production of pathogen – tested Dahlia in France. Acta Horticulturae 325: 781-786.

Brierly, P. and F.F. Smith. 1950. Some vectors, hosts, and properties of dahlia mosaic virus. Plant Disease Reporter. 34: 363.

Brunt, A. A. 1971. Some hosts and properties of dahlia mosaic virus. Annals of Applied Biology 67: 357-368.

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

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

Gatt, M., H. Ding, K. Hammett, and B. Murray. 1998. Polyploidy and evolution in wild and cultivated Dahlia species. Annals of Botany 81:647–656.

Gatt, M, K. Hammett, and B. Murray. 1999. Confirmation of ancient polyploidy in Dahlia (Asteraceae) species using genomic *In situ* hybridization. Annals of Botany 84:39–48.

Geering, A.D.W, T. Scharaschkin, and P-Y. Teycheney. 2010. The classification and nomenclature of endogenous viruses of the family *Caulimviridae*. Archives of Virology 155: 123-131.

Geering, A.D., M.M. Pooggin, N.E. Olszewski, B.E. Lockhart, and J.E. Thomas. 2005a. Characterisation of Banana streak Mysore virus and evidence that its DNA is integrated in the B genome of cultivated Musa. Archives of Virology 150:787–796.

Geering, A.D., N.E. Olszewski, G. Harper, B.E. Lockhart, R. Hull, and J.E. Thomas. 2005b. Banana contains a diverse array of endogenous badnaviruses. Journal of General Virology 86:511–520.

Geering, A.D.W., N.E. Olszewski, G. Dahal, J.E. Thomas, and B.E.L Lockhart. 2001. Analysis of the distribution and structure of integrated banana streak virus DNA in a range of Musa cultivars. Molecular Plant Pathology 2:207–213.

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

Hansen, H.V. and P.D. Sørensen. 2003. A new species of Dahlia (Asteraceae, Coreopsideae) from Hidalgo State, Mexico. Rhodora 105:101–105.

Hansen, H.V. and J. P. Hjerting. 2000. The early history of the domestication of *Dahlia* (Asteraceae, Heliantheae) with emphasis on the period 1791–1836 including observations on taxonomy, chromosome numbers, biochemistry, biosystematics, ray colour inheritance, and ray colour designation within the genus. The Botanic Garden, University of Copenhagen, Denmark.

Hansen, H.V. and J. P. Hjerting. 1996. Observations on chromosome numbers and biosystematics in *Dahlia* (Asteraceae,Heliantheae) with an account on the identity of *D. pinnata, D. rosea* and *D. coccinea*. Nordic Journal of Botany 16: 445–455.

Hansen, C.N, G. Harper, and J.S. Heslop-Harrison. 2005. Characterization of pararetrovirus-like sequences in the genome of potato (*Solanum tuberosum*). Cytogenetic and Genome Research 110: 559-565.

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

Hohn, T., K.R. Richert- Pöggeler, C. Staginnus, G. Harper, T. Schwarzacher, C.H. Teo, P.Y. Teycheney, M.L. Iskara-Caruana, and R. Hull. 2008. Evolution of integrated plant viruses, p. 53-81. In: M.J. Roossinck (eds). Plant Virus Evolution. Springer-Verlag Berlin Heidelberg.

Jakowitsch, J., M.F. Mette, J. van der Winden, M.A. Matzke, and A.J.M. Matzke. 1999. Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. Proceedings of the National Academy of Science USA 96:13241-13246.

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

Lawrence, W. J. C. 1931. The genetics and cytology of *Dahlia variabilis*. Journal of Genetics 24: 257–306.

Lockhart, B. E., J. Menke, G. Dahal, and N.E. Olszewski. 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. Journal of General Virology 81:1579-1585.

Ndowora, T., G. Dahal, D. LaFleur, G. Harper, R. Hull, N.E. Olszewski, and B. Lockhart. 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 Disease 87: 945-948.

Pahalawatta, V., K. L. Druffel, and H. R. Pappu. 2008a. A new and distinct species in the genus *Caulimovirus* exists as an endogenous plant pararetroviral sequence in its host, *Dahlia variabilis*. Virology 376: 253-257.

Pahalawatta, V., K. L. Druffel, S. D. Wyatt, K. C. Eastwell, and H.R. Pappu. 2008b. Genome structure and organization of a novel and distinct species of genus *Caulimovirus* (Family *Caulimoviridae*) associated with dahlia mosaic. Archives of Virology 153: 733-738.

Pappu, H. R., K. L. Druffel, R. Miglino, and A. R. Van Schadewijk. 2008. Nucleotide sequence and genome organization of a new and distinct caulimovirus associated with dahlia mosaic. Archives of Virology 153:2145–2148.

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

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

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

Saar, D.E. and P.D. Sørensen. 2000. *Dahlia parvibracteata* (Asteraceae, Coreopsideae), a new species from Guerrero, Mexico. Novon 10:407–410.

Saar, D.E. and P.D. Sørensen. 2005. *Dahlia sublignosa* (Asteraceae): a species in its own right. Sida 21:2161–2167.

Saar, D.E. 2002. *Dahlia neglecta* (Asteraceae: Coreopsideae), a new species from Sierra Madre Oriental, Mexico. Sida 20:593–596.

Saar, D.E., P.D. Sørensen, and J.P. Hjerting. 2002. *Dahlia spectabilis* (Asteraceae, Coreopsideae), a new species from San Luis Potosi, Mexico. Brittonia 54:116–119.

Saar, D.E., P.D. Sørensen, and J.P. Hjerting. 2003a. *Dahlia campanulata* and *D. cuspidata* (Asteraceae, Coreopsideae): two new species from Mexico. Acta Botany Mexico 64:19–24.

Saar, D.E., N.O. Polans, and P.D. Sørensen. 2003b. A phylogenetic analysis of the genus *Dahlia* (Asteraceae) based on internal and external transcribed spacer regions of nuclear ribosomal DNA. Systematic Botany 28:627–639.

Staginnus, C., M.L. Iskara-Caruana, B. Lockhart, T. Hohn and K.R. Richert-Pöggeler. 2009. Suggestions for a nomenclature of endogenous pararetroviral sequences in plants. Archives of Virology 154: 1189-1193.

Staginnus, C., W. Gregor, M.F. Mette, C.H. Teo, E.G. Borroto-Fernandez, M.L.C. Machado, M. Matzke, and T. Schwarzacher. 2007. Endogenous pararetroviral sequences in tomato (*Solanum lycopersicum*) and related species. BioMed Central Genomics 7:24.

Strother,J.L. and J.L. Panero. 2001. Chromosome studies: Mexican Compositae. American Journal of Botany 88:499–502.

Temsch, E.M., J. Greilhuber, K.R.W. Hammett, and B.G. Murray. 2008. Genome size in Dahlia Cav. (Asteraceae–Coreopsideae). Plant Systematic Evolution 276:157–166.

The International Register of *Dahlia* Names (1969 et seq.). Ninetheenth Supplement. The Royal Horticultural Society, London.

CHAPTER TWO

INCIDENCE OF MULTIPLE AND DISTINCT SPECIES OF CAULIMOVIRUSES IN DAHLIA (*DAHLIA VARIABILIS*)*

ABSTRACT

Dahlia mosaic is a serious disease affecting dahlias. In addition to the *Dahlia mosaic virus* (DMV) reported previously, two putative new caulimoviruses, tentatively designated as DMV-D10 and Dahlia common mosaic virus (DCMV) were recently reported from dahlia. To better understand their relative incidence in dahlia, a total of 213 samples were collected during 2007 and 2008 from several varieties of cultivated dahlia (*Dahlia variabilis*) in the US. Samples were tested for the three caulimoviruses using virus-specific primers in a polymerase chain reaction. Amplicons were cloned and sequenced to confirm the infection of dahlia with these viruses. Results showed that DMV-D10 was the most prevalent (94%) followed by DCMV (48.5%) and DMV (23%). Mixed infections were common and viruses were detected irrespective of symptom expression at the time of sampling. Two percent of the samples were not infected by any of the three tested caulimoviruses. Results suggest that caulimovirus infections are widespread in dahlia and highlight the need for testing and production of virus-free material to reduce their spread.

*Part of this research was published: Eid, S., K.L. Druffel, D. Saar, and H.R. Pappu. 2009. Incidence of multiple and distinct species of caulimoviruses in dahlia (*Dahlia variabilis*). HortScience 44:1498-1500

INTRODUCTION

Dahlia is a high value flower crop in the United States (Pappu et al., 2005) and several other countries in the world (Brunt, 1971; Albouy, 1995). The genus *Dahlia*, whose vernacular name is also "dahlia," is in the sunflower family, Asteraceae. Currently, 35 wild species are recognized (Saar and Sørensen, 2005) plus the cultivated forms, known as either *D. pinnata* or *D. variabilis*. There are many thousands of named cultivars of *D. variabilis* in the world (The International Register of Dahlia Names 1969 et seq.).

There are more than a dozen viruses that infect dahlia (Albouy, 1995). All of these viruses are RNA viruses except one DNA virus, *Dahlia mosaic caulimovirus* (DMV). DMV is widely prevalent in cultivated dahlias and causes an economically important disease referred to as dahlia mosaic (Brunt, 1971; Albouy, 1995). Symptoms of the dahlia mosaic disease include mosaic, vein clearing, vein-banding, leaf distortion, systemic chlorosis and overall stunting of the plant (Pappu and Wyatt, 2003; Pappu et al., 2005). Symptoms vary depending on the cultivar, and cultivars with few or no symptoms are common. DMV is a member of the genus *Caulimovirus*, family *Caulimoviridae*. In addition to DMV, which was characterized at the molecular level (Richins and Shepherd, 1983), recent studies have revealed the presence of at least two more caulimoviruses, tentatively referred to as DMV-D10, and Dahlia common mosaic virus (DCMV) (Nicolaisen, 2003; Pappu et al., 2008; Pahalawatta et al., 2008a). DMV-D10 shares limited sequence similarity with known caulimoviruses including DMV, though the genome organization clearly suggests that it is a member of *Caulimovirus* (Pahalawatta et al., 2008a). DMV-D10 was found to be widespread based on the surveys conducted in the US and Europe (Pappu et al., 2005; Pahalawatta et al., 2007). Moreover, Pahalawatta et al. (2008b) reported that the genome of DMV-D10 is integrated into the dahlia genome and the virus exists

as an endogenous plant pararetroviral sequence (EPRS). DCMV was found to be a distinct caulimovirus species based on sequence relationships with known caulimoviruses (Pappu et al., 2008).

Previous surveys for the caulimoviruses utilized partial sequences of DCMV (then referred to as DMV-Holland in Pahalawatta et al., 2007) and results showed differences in relative incidence of the three caulimoviruses in the US and the Netherlands. Following the complete genome characterization of DCMV, it becomes evident that there are at least three distinct caulimoviruses that are associated with dahlia mosaic. The aim of this study was to determine the incidence of the three distinct caulimoviruses in dahlia.

MATERIALS AND METHODS

Sample collection. A total of 213 dahlia (*D. variabilis*) samples were collected from display gardens in Washington (Tacoma trial garden) between June and September 2007 and during September 2008. A wide range of symptoms were noticed. These included mosaic, systemic chlorosis, leaf malformation, and stunting. Sampling was done randomly and irrespective of the presence of symptoms.

Total nucleic acid extraction. Total nucleic acids were extracted from dahlia samples using the modified Dellaporta procedure (Presting et al., 1995) as described in Pahalawatta et al. (2007).

Polymerase chain reaction (PCR) amplification. Primer pairs specific to each of the three virus species were used (Table 1). The quality of the extracted DNA was verified by PCR using a host- specific primer pair designed based on a portion of the internal transcribed spacer regions

of dahlia (ITS)(GenBank # AF165831.1). A 1:20 dilution of the total nucleic acid extract was used in PCR. Each 20 µL PCR reaction contained 2 µL of the total nucleic acid extract, 1X PCR buffer (20 mM Tris, pH 8.4 and 500 mM KCl), 150 µM dNTP mix, 2 mM $MgCl_2$, and 0.6 pmole each of sense and antisense primers, 12.7 µL sterile H_2O, and 0.1 µL Taq DNA polymerase. The amplification was performed in a DNA thermal cycler (BioRad, Hercules, CA) with the following parameters: Fifty cycles at 94 °C for 30 s, the required annealing temperature (Table 1) based on the primer pair used for 20 s, and 72 °C extension step determined based on the size of the amplicon to be synthesized (at the rate of 1000 bp /min), preceded by an initial incubation of 94 °C for 4 min, and terminated by a final incubation at 72 °C for 7 min. PCR products (7.5 µL) were analyzed by agarose gel electrophoresis (1.2%) in 0.5 X TAE (Tris-acetate-EDTA) buffer.

Cloning and sequencing of the amplicons. At least one amplicon for each of the virus species was cloned into the TOPO vector (Invitrogen, Carlsbad, CA). Recombinant plasmids were sequenced at the Washington State University sequencing core facility. Sequences were compared with the genomic sequences of the three viruses to confirm the identity of the amplicons.

Table 1. Primers used to detect *Dahlia mosaic virus* (DMV), Dahlia common mosaic virus and DMV-D10.

	Primer Pairs[a]	Sequence (5'-3')	Annealing temperature	Extension time	Expected size (bp)
Partial internal transcribed spacer region, dahlia	ITS F2/R	AATGCGATACTTGGTGTGAA TCGAAGCATCATAACAAGAC	59°C	30 s	340
DMV-D10	Orf 6 End/Start	ATGGAAGAAATTAAGGCGT TTGTCTTCATCCATAAAGCAG	60°C	1 min 30 s	1280
	Orf 6 1299-1300	CCAGGTGCAAGTCCGGAATTA TAGGGCCAACTTCTGTCTCAT	60°C	40 s	513
	Orf 4 1926c/1431F	TGCATAAAATGAGTTCTATC TGAACTTGTTCATCATTATC	53°C	30 s	480
	Orf 1 Start /end	CTGTTTTTC TGTGTTTCTACTGG ATGGATCGTAAAGATT	50°C	1 min	900
	Orf 1 416/1070	TTACCACTTCTAACAAAAGG TCCATACATGTCTACTTTTTCG	60°C	40 s	600
DMV	Kapth F/R	ATGAGTAATGCTTCAGCAA TGACCATGGCTTCTAACTGT	56°C	30 s	504
DCMV	Den CP F/R	GGATCCTCATTCTGAGTCTTCGTCTTC CATATGGCCACCCAAATGACC	59°C	50 s	1517

[a]primer sequences for DCMV and DMV-D10 were from Pahalawatta et al, 2007, 2008b. DMV- and ITS-specific primers were designed from GenBank accessions # AY291586 and AF165831, respectively.

RESULTS AND DISCUSSION

Of the 193 symptomatic and asymptomatic dahlia samples collected from a dahlia display garden in 2007, DMV-D10 was found in 94% (Fig. 1) of the samples. Specific amplicons of expected size were obtained from each of the primer pair used and the sequence results verified their identity. DCMV was the second most frequent virus, after DMV-D10, found during the survey (Fig. 2). Mixed infections were common. The frequency of incidence of DMV alone and associated with other viruses was low compared to DMV-D10 and DCMV and it ranged between 17% and 26% (Fig. 2). The occurrence of mixed infections containing all three caulimoviruses, DMV-D10, DMV, and DCMV, was from 8.82% to 15.63% in 2007 (Fig. 2). The results of the 20 samples collected in September 2008 also showed the relatively higher incidence of DMV-D10 and DCMV compared to that of DMV (Fig. 2).

Symptoms associated with virus infection included mosaic, mottling, vein chlorosis and necrosis and in some plants resulted in stunting. There was no noticeable correlation between the detection of the virus and the symptom expression as 38-41% of DMV-D10 infected samples and 40-44% of DCMV infected samples showed symptoms (Fig. 3). However, plants were frequently infected with more than one virus, and the majority of samples were asymptomatic but were found to be positive by PCR. DMV-D10 incidence was higher (> 50 %) in asymptomatic samples, compared to DCMV and DMV (Fig. 3).

Results of this study showed the widespread occurrence of DMV-D10 in cultivated dahlias. Additionally, testing showed that DCMV was more prevalent than DMV, and mixed infections were common. This is supported by earlier studies (Pappu et al., 2005; Pahalawatta et al., 2008a). Infection of dahlia plants by DMV-D10 was often asymptomatic (>50 % of tested samples), but sometimes associated with symptoms (mentioned above). The high incidence of

DMV-D10 could be due to the fact that this virus was shown to exist as DvEPRS (endogenous pararetroviral sequence) (Pahalawatta et al., 2008b) and for this reason different primer sets to detect DMV-D10 were used. It was reported that EPRS are not always neutral and they contribute to pathogenicity as shown in the case of *Tobacco vein clearing virus* (TVCV) and *Petunia vein clearing virus* (PVCV) (Richert-Poggeler et al., 2003; Lockhart et al., 2000; Harper et al., 1999; Ndowora et al., 1999).

The other two caulimoviruses, DCMV and DMV, do not appear to be integrated into the dahlia genome and instead they seem to exist as episomal elements typical of other known species in the genus *Caulimovirus* (Pahalawatta et al., 2008b). DCMV was the second most prevalent caulimovirus after DMV-D10, followed by DMV in incidence with 48.5 % and 23%, respectively (Fig. 1). In a previous survey by Pahalawatta et al. (2007) of dahlias in the US and the Netherlands, DCMV and DMV infections were 31.9% and 21.8% respectively. Mixed infections were found to be common, whereas the frequency of incidence of DMV and DMV-D10 (9.38%) together was low compared to DMV-D10 and DCMV (43.75%) (Fig. 2), and 15.6% of the samples had all three viruses.

The results from testing the cultivated dahlias highlight the importance of screening dahlia propagative material for these three viruses before using them in vegetative propagation, especially that presence of symptoms is not a criterion for virus infection diagnosis. At present, PCR-based detection assays are the most reliable methods for detecting viruses associated with dahlia mosaic. Reducing the spread and impact of these viruses in dahlia requires production of virus-free stock, and virus interception and elimination tactics should take into account the diversity of caulimoviruses extant in dahlia.

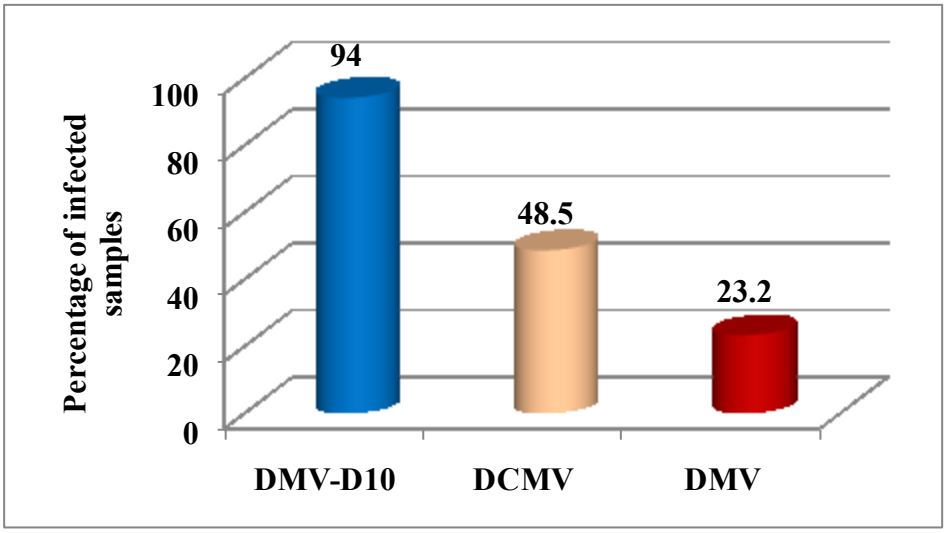


Fig. 1. Percent incidence of Dahlia mosaic virus (DMV), DMV-D10, and Dahlia common mosaic virus (DCMV) in cultivated dahlia. Virus-specific primers were used in a polymerase chain reaction assay. A total of 193 samples were tested.

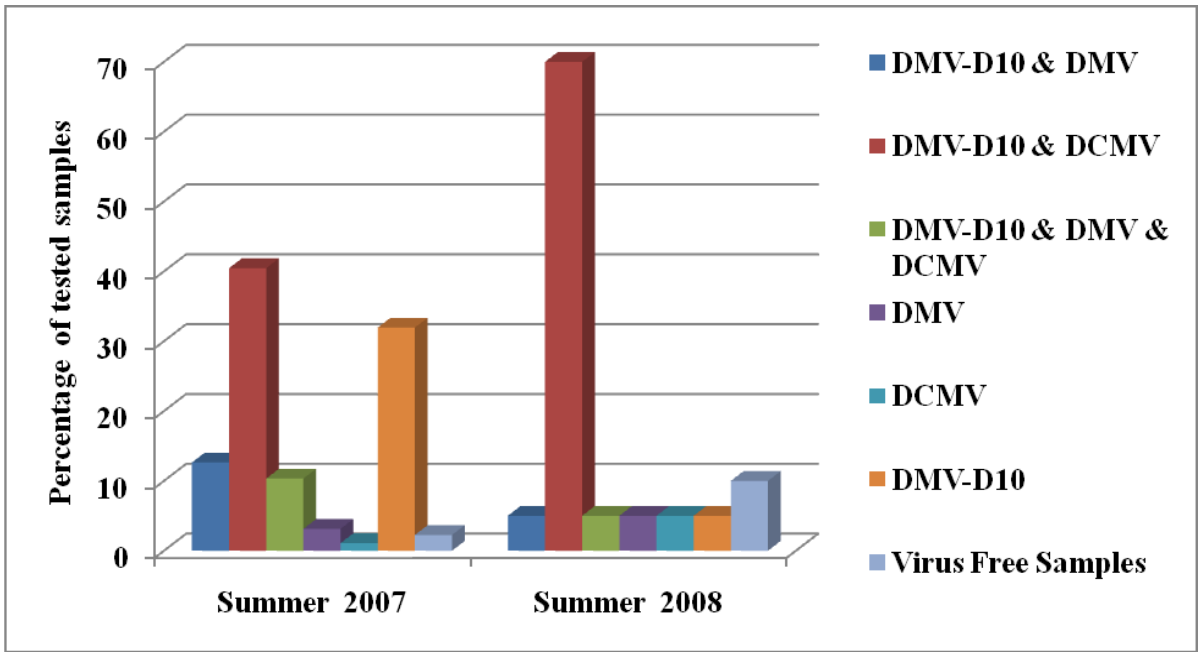


Fig. 2. Relative incidence (percent) of Dahlia mosaic virus (DMV), DMV-D10, and Dahlia common mosaic virus (DCMV) in 213 samples of cultivated dahlias tested in 2007 and 2008.

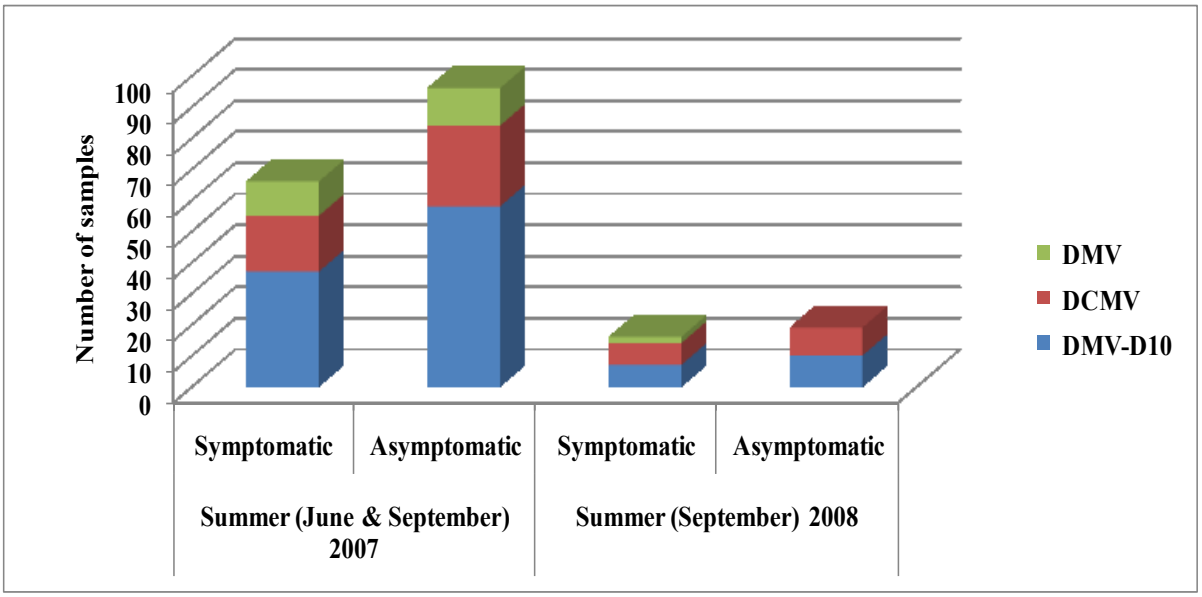


Fig. 3. Number of symptomatic (mosaic, mottling, vein chlorosis and necrosis, with stunting) and asymptomatic plants collected in June and September 2007 and September 2008).

REFERENCES

Albouy, J. 1995. Dahlia, p.265-273. In: G. Lobenstein, R.H. Lawson, and A.A. Brunt (eds.). Virus and virus-like diseases of bulb and flower crops. John Wiley and Sons, Chichester, NY.

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

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

Lockhart, B.E., J. Menke, G. Dahal, and N. E. Olszewski. 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. Journal General Virology 81: 1579-1585.

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

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

Pahalawatta, V., K. L. Druffel, S. D. Wyatt, K. C. Eastwell, and H.R. Pappu. 2008a. Genome structure and organization of a novel and distinct species of genus *Caulimovirus* (Family *Caulimoviridae*) associated with dahlia mosaic. Archives of Virology 153: 733-738.

Pahalawatta, V., K. L. Druffel, and H. R. Pappu. 2008b. A new and distinct species in the genus *Caulimovirus* exists as an endogenous plant pararetroviral sequence in its host, *Dahlia variabilis*. Virology 376: 253-257.

Pahalawatta, V., R. Miglino, K. L. Druffel, A. Jodlowska, A. R. Van Schadewijk, and H. R. Pappu. 2007. Incidence and relative prevalence of distinct caulimoviruses (Genus *Caulimovirus*, Family *Caulimoviridae*) associated with dahlia mosaic in *Dahlia variabilis.* Plant Disease 91:1194-1197.

Pappu, H. R., K. L. Druffel, R. Miglino, and A. R. Van Schadewijk. 2008. Nucleotide sequence and genome organization of a new and distinct caulimovirus associated with dahlia mosaic. Archives of Virology 153:2145–2148.

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

Pappu, H. R. and S. D. Wyatt. 2003. Dahlia mosaic virus. A description of symptoms on dahlia. APSnet Image of the Week. http://apsnet.org/online/archive/2003/IW000032.asp.

Presting, G. G., O. P. Smith, and C. R. Brown. 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.

Richert-Poggeler, K. R., F. Noreen, T. Schwarzacher, G. Harper, and T. Hohn. 2003. Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia. European Molecular Biology Organization Journal 22: 4836-4845.

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

Saar, D.E. and P. D. Sørensen. 2005. *Dahlia sublignosa*: a species in its own right. SIDA Contributions to Botany 21: 2161-2167.

The International Register of *Dahlia* Names (1969 et seq.). Ninetheenth Supplement. The Royal Horticultural Society, London.

CHAPTER THREE

PLANT PARA-RETROVIRAL SEQUENCES IN WILD DAHLIA SPECIES IN THEIR NATURAL HABITATS IN MEXICAN MOUNTAIN RANGES

ABSTRACT

The central mountain ranges of Mexico are home to the greatest diversity for genus *Dahlia*. First described in 1791, the genus has 35 recognized wild species in addition to the cultivated forms, known as either *D. pinnata* or *D. variabilis*. Majority of the wild dahlia species were found in the Mexican mountain ranges. The family *Caulimoviridae* consists of several plant-infecting viruses that are considered plant para-retroviruses as these viruses exist in RNA phase in their hosts but utilize a virally-coded reverse transcriptase to synthesize and package double-stranded DNA in progeny virions. Three distinct caulimoviruses, *Dahlia mosaic virus*, DMV-D10, and Dahlia common mosaic virus, were found to be associated with cultivated dahlia (*D. variabilis*). To better understand the incidence of these para-retroviruses, selected wild *Dahlia* species in their natural habitats from west - central Mexico were tested for the three caulimoviruses. Virus species-specific primers and PCR were used followed by cloning and sequencing of the amplicons. Results showed that the wild dahlia species in their natural habitat contained DMV-D10. Viral sequences were found in 91% of the samples (n=56) representing four different wild species. The gene coding for the movement protein of DMV-D10 from *D. coccinea* was cloned and sequenced. Sequence comparisons showed divergence of this gene when compared to that of DMV-D10 from cultivated dahlias. The discovery of plant para-retroviruses in wild dahlia species in their natural habitats suggests a possible emergence, co-existence and co-evolution of para-retroviruses and their host plants.

INTRODUCTION

Dahlia (*Dahlia variabilis*) is an economically important ornamental crop in the US and several parts of the world. The genus *Dahlia*, whose vernacular name is also "dahlia," is in the sunflower family, Asteraceae. Currently, 35 "wild" species are recognized (Saar and Sørensen, 2005) plus the cultivated forms, known as either *D. pinnata* or *D. variabilis*. There are over 2,000 named cultivars of *D. variabilis* in the US. The geographic range for the genus *Dahlia* is in Mexico's in Sierra Madre Occidental region (Fig. 1). Many of the *Dahlia* spp have very limited ranges; several are known from only one or two populations (Saar, 1999). The exceptions are *D. australis,* which occurs at least as far south as southwestern Guatemala, and *D. coccinea* and *D. imperialis*, which have been reported throughout Central America into northern South America. Plant size varies from the small *D. tenuis* and *D. scapigera* that only average three to six decimeters in height on slender stems, to tall arborescent species such as *D. tenuicaulis*, to *D. macdougallii*, which grows from under the mosses and ferns covering tree trunks and produces long shoots that sprawl across the canopy branches of tropical hardwoods.

There are more than dozen viruses that have been found to infect dahlia (Albouy, 1995). All, but one, are RNA viruses. *Dahlia mosaic caulimovirus* (DMV), DNA virus, is widely prevalent in cultivated dahlias and causes an economically important disease referred to as dahlia mosaic (Brunt, 1971; Albouy, 1995). Based on the molecular characterization, DMV is considered as a distinct species in the genus *Caulimovirus* (Richins and Shepherd, 1983). The most characteristic symptoms of the dahlia mosaic disease include mosaic, vein clearing, vein-banding, leaf distortion, systemic chlorosis, flower breaking, and overall stunting of the plant. Symptoms vary depending on the cultivar, and cultivars with few or no symptoms are common.

In addition to DMV, we have recently shown that there are at least two more distinct caulimoviruses associated with dahlia. Tentatively referred to as Dahlia common mosaic virus (DCMV) and DMV-D10 (Pahalawatta et al., 2007; Pappu et al., 2008), these two new caulimoviruses showed significant sequence divergence compared to that of DMV (Pappu et al., 2008; Pahalawatta et al., 2008a). Both DCMV and DMV-D10 were found to be widespread based on the surveys conducted in the US and Europe (Eid et al., 2009; Pahalawatta et al., 2007; Pappu et al., 2005).

Of the three caulimoviruses found in dahlia, we obtained evidence that DMV-D10 is integrated into the dahlia genome and thus exists as an endogenous plant pararetrovirus (DvEPRS) (Pahalawatta et al., 2008b). While much is known about the plant para-retroviruses in cultivated plant species, their incidence in natural habitats of wild species not influenced by modern agricultural practices remains to be investigated. Using wild *Dahlia* collected in their native habitats as a model system to study plant-pararetrovirus ecology and evolution could provide insights into the evolutionary pathways, and relationships of pararetroviral sequences with their host plants.

MATERIALS AND METHODS

Sampling site and procedure. Wild dahlia species were collected from the slopes of the Sierra Madre Occidental including the states of Nayarit, Sinaloa, Chihuahua and Durango in western Mexico during August and early September of 2006 (Fig. 1; Fig. 2). Plants were collected irrespective of the presence of any virus-induced symptoms. Fifty-six samples that included four wild dahlia species (*D. coccinea, D. rupicola, D. tenuicaulis,* and *D. sherffii*) were tested for

DMV, DCMV and DMV-D10 incidence. Each sample was from a separate population. The tissues were dried and stored at -80°C.

Total nucleic acid extraction. Total nucleic acids were extracted from dahlia leaf material using the modified Dellaporta procedure (Pappu et al., 2005; Presting et al., 1995). To avoid contamination, each sample was treated with care. The sample tube was opened and approximately weighing 20-40mg of leaf tissue and placed in a new tube. To avoid possible contamination, weighing and extraction procedures were conducted in a room where no DMV work had been done previously.

Polymerase chain reaction amplification. Polymerase chain reaction (PCR) was performed by two different persons in two different periods of time to determine the relative incidence of the three caulimoviruses. Primer pairs specific to each of the virus species were used (Fig. 3; Table 1). The quality of extracted DNA was controlled by PCR using host specific primer set designed on internal transcribed spacer region of dahlia (ITS).

A 1:20 dilution of the total nucleic acid extract was used in PCR. Each 20 µl PCR contained 2 µl of the total nucleic acid extract, 1X PCR buffer (20 mM Tris, pH 8.4 and 500 mM KCl), 150 µM dNTP mix, 2 mM $MgCl_2$, 0.6 pmole each of sense and antisense primer, 12.7 µl sterile H_2O, and 0.2 µl GoTaq® Flexi DNA Polymerase (Promega, Madison, WI) . The amplification was performed in a DNA thermal cycler (BioRad, Hercules, CA) with the following parameters: 50 cycles at 94 °C for 30 sec, the required annealing temperature (Fig. 3; Table 1) based on primer pair used for 20 sec, and 72 °C extension step determined based on the

size of the amplicon to be synthesized at the rate of 1000 bp /min, preceded by an initial incubation of 94 °C for 4 min, and terminated by a final incubation at 72 °C for 7 min. PCR products (7.5 µl) were analyzed by agarose gel electrophoresis (1.2%) in 0.5 X TAE (Tris-Acetate-EDTA) buffer. The identity of amplicons was verified by cloning and sequencing. Amplicons were cloned into pGEM-T (Promega, Madison, WI). Nucleotide sequences were determined using the ABI Prism Sequencing System at the Molecular Biology Core Laboratory of the Washington State University, Pullman, WA, and compared to the available sequences in the GenBank.

Sequence analysis. DNA fragments were assembled into contigs and the contigs were assembled to give complete sequences using ContigExpress (Vector NTI Suite 9.0.0, Informax Inc, Bethesda, MD). Phylogenetic analysis was done using CLUSTALW version 1.83 (Thompson et al., 1994) and MEGA 4 (Tamura et al., 2007).

Fig. 1. Wild *Dahlia* spp. in their natural habitat "Sierra Madre Occidental". Left: *Dahlia coccinea*; Right: *Dahlia sherffii*. Photos © D. E. Saar.

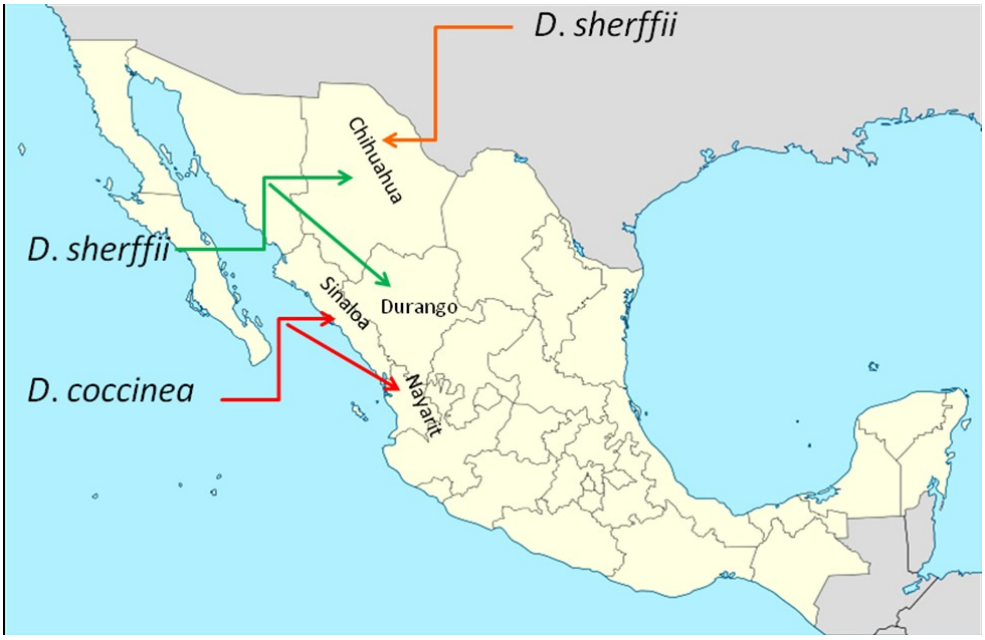


Fig. 2. Map of Mexico showing the occurrence of the four wild *Dahlia* spp. The sampling was done from Nayarit, Sinaloa, Chihuahua and Durango states. Modified from http://commons.wikimedia.org/wiki/File:Mexico_States_maps.svg

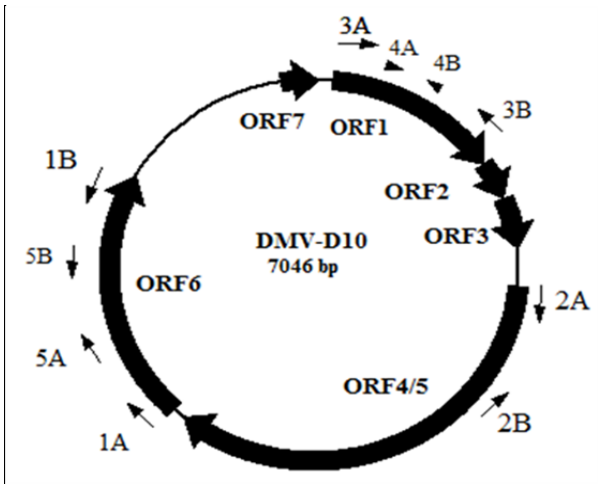


Fig. 3. Schematic of genome organization of Dahlia mosaic virus -D10 showing the relative positions (sense A; antisense B) of each ORF-specific primer pair used in this study.

Table 1. Primers used to detect the three distinct caulimoviruses in dahlia

	Primer	Sequence (5'-3')	Annealing temperature (°C)	Expected size (bp)
DMV-D10	1A	ATGGAAGAAATTAAGGCGT	60	1280
	1B	TTGTCTTCATCCATAAAGCAG		
	2A	TGCATAAAATGAGTTCTATC	53	480
	2B	TGAACTTGTTCATCATTATC		
	3A	CTGTTTTTC TGTGTTTCTACTGG	50	900
	3B	ATGGATCGTAAAGATT		
	4A	TTACCACTTCTAACAAAAGG	60	600
	4B	TCCATACATGTCTACTTTTTCG		
	5A	CCAGGTGCAAGTCCGGAATTA	60	513
	5B	TAGGGCCAACTTCTGTCTCAT		
DMV	Kapth F	ATGAGTAATGCTTCAGCAA	56	504
	Kapth R	TGACCATGGCTTCTAACTGT		
DCMV	Den CP F	GGATCCTCATTCTGAGTCTTCGTCTTC	59	1517
	Den CP R	CATATGGCCACCCAAATGACC		
Dahlia ITS[a]	ITS F	AATGCGATACTTGGTGTGAA	59	340
	ITS R	TCGAAGCATCATAACAAGAC		

[a] Internal transcribed spacer region of dahlias

RESULTS AND DISCUSSION

Leaf samples from four wild dahlia species were tested for three caulimoviruses using virus-specific primers in a PCR assay conducted by two different persons as mentioned in Materials and Methods and the outcome was similar. Results showed that the wild dahlia species contained DMV-D10. The four wild species that were tested for DMV, DCMV and DMV-D10 represented each of the three major clades within the genus *Dahlia* (Saar et al., 2003). There was no evidence of infection by DMV and DCMV. DMV-D10 was found in 91% of the wild dahlia samples (n=56) tested, and the rest were found to be virus-free (Fig. 2; Table 2). The virus-free populations of *D. sherffii* were collected from Chihuahua while those found infected with DMV-D10 were from both Chihuahua and Durango (Fig. 2).

Primer pairs specific to each of the viral genes were used to determine the relative distribution of the viral genes of DMV-D10 in the wild *Dahlia* spp. The ORF VI coding for the inclusion body protein was found at the highest frequency (100 %, detected in 51 out of 51 infected samples), followed by the movement protein gene, ORF I (80%, detected in 41 out 51 infected samples), and the reverse transcriptase gene, ORF IV-V (45%, detected in 23 out of 51 infected samples) (Table 3). To further investigate the potential sequence divergence of this caulimoviral genome in wild dahlia species compared to that of cultivated dahlias, the gene coding for the movement protein (ORF I) was cloned and sequenced. Pairwise comparisons at nucleotide and amino acid levels with the corresponding gene of DMV-D10 from cultivated dahlias (Pahalawatta et al., 2008b) showed 93% and 89 % identity at nucleotide and amino acid levels, respectively (Fig. 4a,b). Phylogenetic analysis showed that the deduced amino acid sequences of the ORF I (encoding the movement protein) from wild and cultivated dahlias formed one cluster within the genus *Caulimovirus* (Fig. 5a,b).

The nature, extent and timing of the spread of DMV-D10 from wild species to cultivated plants or vice versa is not clear since the ORFII (aphid transmission factor) is not found in the cultivated plants infected with DMV-D10 (Pahalawatta et al., 2008b). It is well known that plant viruses co-evolved with wild plants in their centers of origin, before they were domesticated, which make them the main source of virus diversity (Lovisolo et al., 2003). Investigating the occurrence of viruses in natural settings gained significant attention in recent year and raised the possibility of future emergence of such viruses as crop pathogens (Melcher et al., 2006). MacClement and Richards (1956) found that total annual infection was about 10% of herbaceous annual and perennial plants, and these plants were infected by one or more viruses as indicated by mechanical inoculation to test plants. Gibbs et al. (2000) reported *Diuris virus* Y (genus *Potyviru*s) and *Pterostylis blotch virus* (genus *Tospovirus*) in plants of *Diuris orientis* and *Pterostylis* spp. in their natural habitats in Australia. Barley yellow dwarf viruses (BYDVs) are prevalent in natural grasslands (Garret et al., 2004) and also were detected in herbarium specimens (Malmstrom et al., 2007). While these findings mostly dealt with RNA viruses, our study, for the first time, confirmed the existence of reverse transcribing, DNA viral elements in wild plant species.

Molecular characterization and phylogenetic studies of viral genomes from wild and cultivated plant species may provide important clues about the relation between "source" and "sink", as was the case in *Cotton leaf crumple virus* (Brown, 2002). Malmstrom et al. (2007) used the historical virus sequences of BYDVs to determine rough time estimates of relevant phylogenetic events. Similarly, Bousalem et al. (2003) analyzed the diversification of viral sequences to reveal the course of virus invasion in *Dioscorea* spp. Studies such as these on plant –virus interaction in natural and managed ecosystems provide insight into plant virus

ecology and the human influence on plant virus diversification and spread within natural ecosystems (Malmstrom et al., 2007; Malmstrom and Melcher, 2008)

Our findings of viral sequences in plant species in their centers of diversity show the possible co-evolution of reverse transcribing elements and plant para-retroviruses with their plant hosts. Further studies on correlating the incidence of viral sequences with the species and population distribution of wild dahlias would provide new avenues of research into the evolutionary pathways of plant-associated pararetroviruses.

Table 2. Incidence of caulimoviruses in wild dahlia species

Dahlia spp.	Samples infected -DMV-D10 [a]	Samples free of viral sequences
D. coccinea	33 ／34	1／34
D. rupicola	1／1	0／1
D. sherffii	16／20	4／20
D. tenuicaulis	1／1	0／1

[a] Number of plants per species infected/total number of samples tested.

Table 3. Frequency of occurrence of Dahlia mosaic virus (DMV)-D10 ORFs

	DMV-D10 Primers corresponding to ORF I, ORF IV and ORFVI			
Dahlia spp.	ORFI,ORFIV and ORFVI[a]	ORFVIand ORFI[a]	ORFVI and ORFIV[a]	ORFVI[a]
D. coccinea	17／ 34	14／34	0／ 34	2／ 34
D. rupicola	0／ 1	0／ 1	1／ 1	0 / 1
D. sherffii	5／ 20	4／ 20	0／ 20	7／ 20
D. tenuicaulis	0／ 1	1／ 1	0／ 1	0／ 1

[a]Number of plants per species showed amplification (s) with primers pair of ORFI , ORFIV and ORFVI combinations /total number of samples tested.

```
D.variabilis (     1) atggatcgtaaagatttatttaatataaataaagatgattcggaagaatc
D.coccinea   (     1) ..........t.........g..t.......................

D.variabilis (    51) tttcgatcaggaaaccactgaagtaataaactttagtgaaacaaacaaag
D.coccinea   (    51) a...a.c..a....tt.................c.......acg.a..t.

D.variabilis (   101) gatttgtatcagactacatgattaattctgattatttagaacaaatcatg
D.coccinea   (   101) .t....c......t..................c.cc.....a....t...

D.variabilis (   151) aagcttaagctaaagcttgatacaaaacaggtttttaatcaacctagtaa
D.coccinea   (   151) ........at....a.....c........a..........g....a..g

D.variabilis (   201) tatacagagattagtttcaaaagctttctctagaaaaaataatatctttt
D.coccinea   (   201) .t.....................a.......................

D.variabilis (   251) attgctttaatactgaagaattgtcagtagatataaaagatactacaggt
D.coccinea   (   251) ..................a.........................a...c..

D.variabilis (   301) gaattgtatttaccacttctaacaaaaggagaaatagccagaagacttct
D.coccinea   (   301) ...g.............c.............................

D.variabilis (   351) gactattaaaccagaattaagaaaaaccatgaatatggtgcacatcggag
D.coccinea   (   351) .............................................

D.variabilis (   401) cagtaaaaatccttctgaaggcacagttcagagatggaattaacttcccg
D.coccinea   (   401) ...................t.a.........................

D.variabilis (   451) ataaaaatggctttagttgataacagaattatcaacaggcaagacgctct
D.coccinea   (   451) .............................................

D.variabilis (   501) actcggagcagttcaaggaaatttagcatacggtaaatttatgtttactg
D.coccinea   (   501) ...a..ca.....................................g...

D.variabilis (   551) tttatcctaaatttgcattacatcgagattcaaaagatttcgataaaacc
D.coccinea   (   551) ........................................a......

D.variabilis (   601) ttaagtttcatacatcagtgcgaaaggactgacctcatggaaccaggtaa
D.coccinea   (   601) c............................c....................

D.variabilis (   651) caaagtatttacgattaattatttaatttcgtatgctttgacaaatagta
D.coccinea   (   651) ......t.....................................

D.variabilis (   701) ctcattcaattgagtataaagaaaaggagagtataacacttgatgatgta
D.coccinea   (   701) ..........aa.............a...........ta....a......

D.variabilis (   751) ttctcagaaataggtactgtcgaaggaagcaagttcgctgaaccttctca
D.coccinea   (   751) ..............a..........a...t...................

D.variabilis (   801) gatacaagaaaattgggcgattgatattgctcgagagaaacaaactctag
D.coccinea   (   801) .t....g..............a....c........a.............

D.variabilis (   851) gatttcaacctagaaatagttttacaggaatccccttacaaataggcgag
D.coccinea   (   851) ...a................a.............t.........a.a..a

D.variabilis (   901) tccagtagaaacacagaaaaacaacttattcggtctatgtccgaaaaagt
D.coccinea   (   901) ......---.............................a...........

D.variabilis (   951) agacatgtatggagaaatccttagagaattgattgggaaatga
D.coccinea   (   948) ............................a.............
```

A

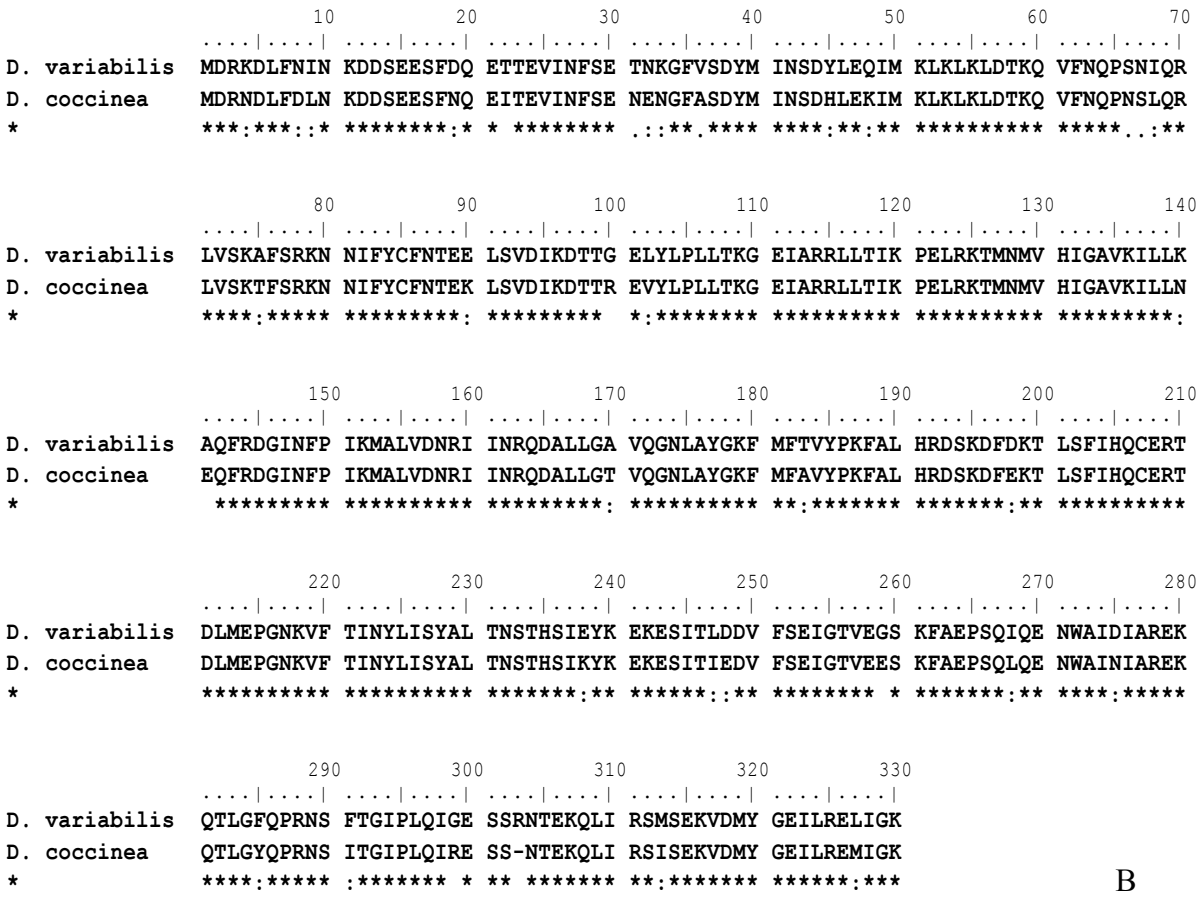
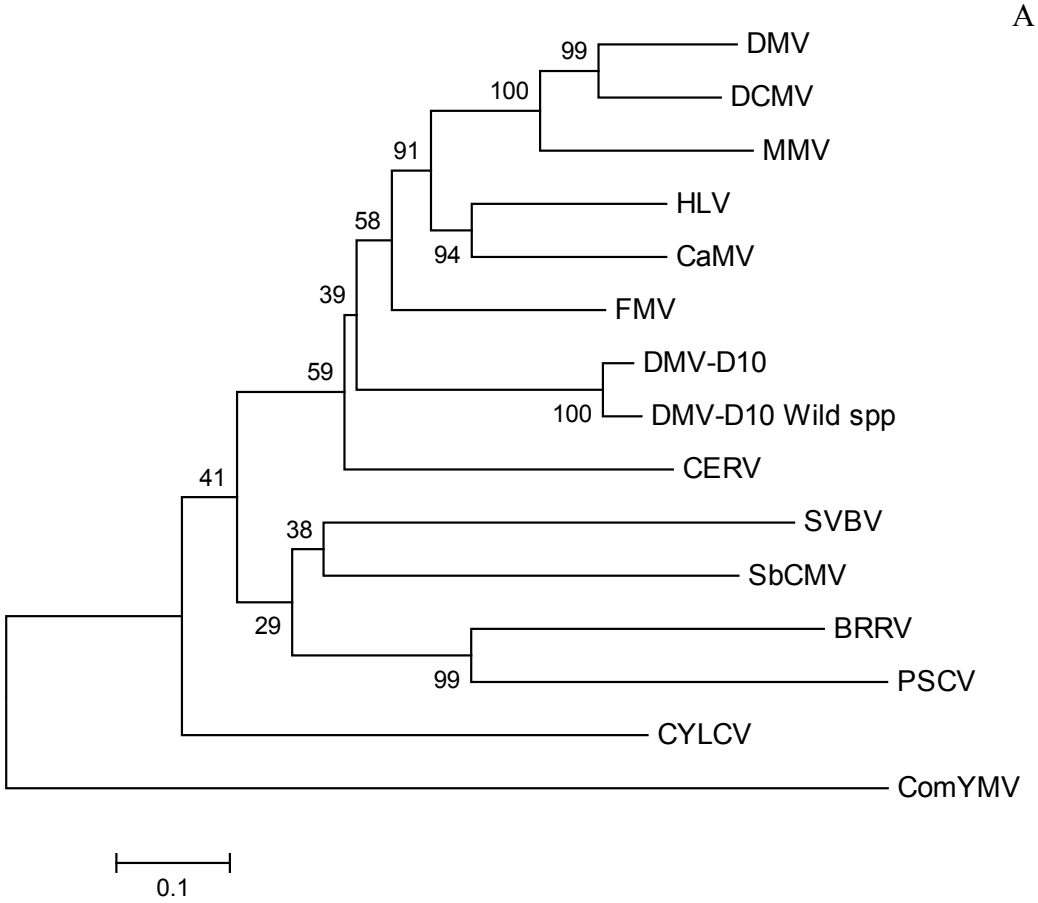


Fig. 4. Pair-wise alignment of a) nucleotide and b) deduced amino acid sequences of the open reading frame I of Dahlia mosaic virus (DMV) -D10 from wild dahlia, *Dahlia coccinea* and cultivated dahlia, *D. variabilis* (identical nucleotides or amino acid residues in a given position were denoted by asterisks).

A


99
DMV
100
DCMV
91
MMV
HLV
58
94
CaMV
39
FMV
59
DMV-D10
100
DMV-D10 Wild spp
CERV
41
38
SVBV
SbCMV
29
BRRV
99
PSCV
CYLCV
ComYMV
0.1

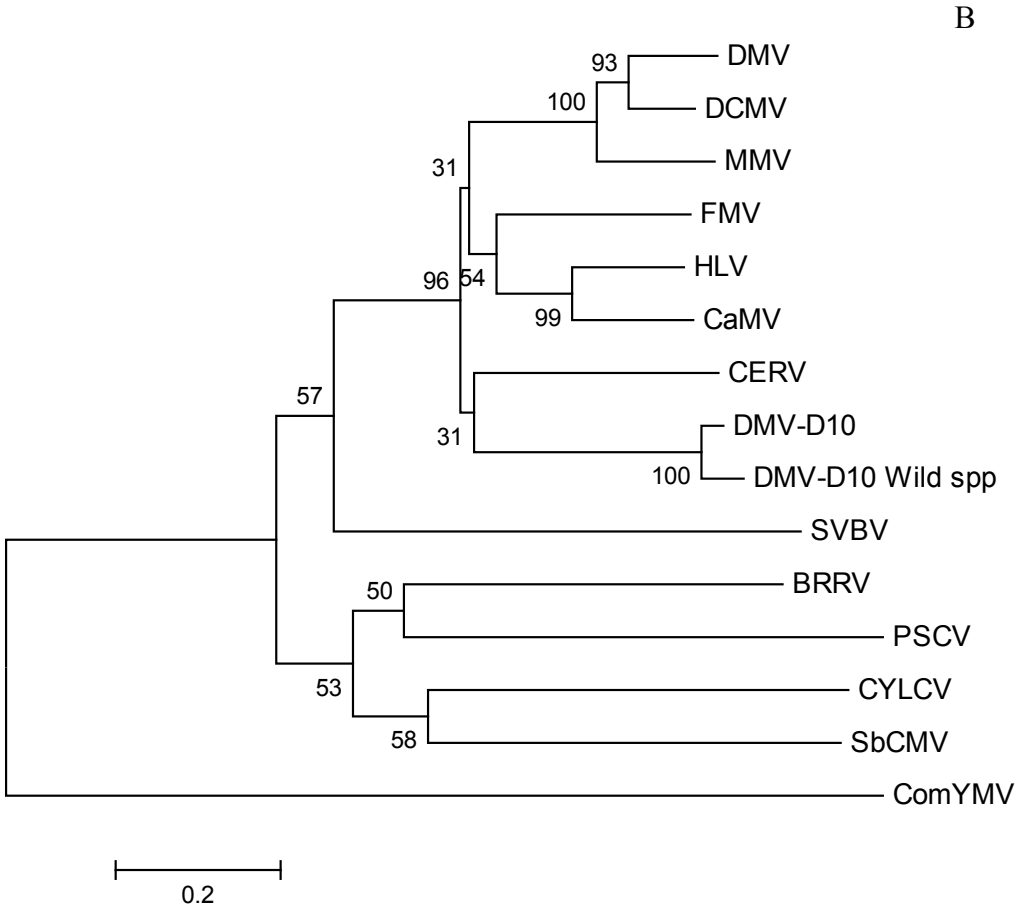


Fig. 5. Phylogenetic analysis of Dahlia mosaic virus - D10 from cultivated (DMV-D10) and wild dahlia *D. coccinea* (DMV-D10 wild spp) collected from the mountain ranges of Mexico. Cluster dendrograms were based on (a) nucleotide and (b) amino acid sequences of the movement protein gene (ORF I) of selected members of the family *Caulimoviridae*. BRRV-*Blueberry red ringspot virus*, CaMV-*Cauliflower mosaic virus*, CERV-*Carnation etched ring virus*, CYLCV- *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.

REFERENCES

Albouy, J. 1995. Dahlia, p.265-273. In: G. Lobenstein, R.H. Lawson, and A.A. Brunt (eds.). Virus and virus-like diseases of bulb and flower crops. John Wiley and Sons, Chichester, NY.

Bousalem, M., S. Dallot, S. Fuji, and K.T. Natsuaki. 2003. Origin, world-wide dispersion, bio-geographical diversification, radiation and recombination: an evolutionary history of *Yam mild mosaic virus* (YMMV). Infection, Genetics and Evolution 3: 189-206.

Brown, J. .K. 2002. The molecular epidemiology of begomoviruses, P. 279-315. .In: J.A. Kahn, and J. Dijkstra (eds.). Plant Viruses as Molecular Pathogens. The Haworth Press, Oxford.

Brunt, A. A. 1971. Some hosts and properties of dahlia mosaic virus. Annals of Applied Biology 67: 357-368.

Eid ,S., K.L. Druffel, D. Saar, and H.R. Pappu. 2009. Incidence of multiple and distinct species of caulimoviruses in dahlias (*D. variabilis*). HortScience 44:1498-1500.

Garrett, K.A., S.P. Dendy, A.G. Power, G.K. Blaisdell, H.M. Alexander, and J.K. McCarron. 2004. Barley yellow dwarf disease in natural populations of dominant tallgrass prairie species in Kansas. Plant Disease 88: 574.

Gatt, M., K.R.W. Hammett, and B.G. Murray. 2000. Molecular phylogeny of the genus *Dahlia* Cav. (Asteraceae, Heliantheae- Coreopsidinae) using sequences derived from the internal transcribed spacers of nuclear ribosomal DNA. Botanical Journal of the Linnean Society 133**:** 229–239.

Gibbs, A., A. Mackenzie, A. Blanchfield, P. Cross, C. Wilson, E. Kitajima, M. Nightingale, and M. Clements. 2000. Viruses of orchids in Australia; their identification, biology and control. Australian Orchid Review 65**:** 10-21.

Lovisolo, O., R. Hull, and O. Rosler. 2003. Co-evolution of viruses with hosts and vectors and possible paleontology. Advances in Virus Research 62: 325–379.

MacClement, W.D., and M.G.Richards. 1956. Virus in wild plants. Canadian Journal of Botany 34: 793-797.

Malmstrom, C.M., and U.Melcher. 2008. The emerging field of plant virus ecology. Phytopathology 98 (Suppl. 6): 189.

Malmstrom, C.M., R. Shu, E.W. Linton, L.A. Newton, and M.A. Coo. 2007. Barley yellow dwarf viruses (BYDVs) preserved in herbarium specimens illuminate historical disease ecology of invasive and native grasses. Journal of Ecology 95: 1153-1166.

Melcher, U., V. Grover, V. Muthukumar, F. Zhang, and O. Blinkova. 2006. Methods for surveying plants in natural settings for viruses. Phytopathology 96 (Suppl. 6): 77.

Pahalawatta, V., K.L. Druffel, and H.R. Pappu. 2008b. A new and distinct species in the genus *Caulimovirus* exists as an endogenous plant pararetroviral sequence in its host, *Dahlia variabilis*. Virology 376: 253-257.

Pahalawatta, V., K.L. Druffel, S.D. Wyatt, K.C. Eastwell, and H.R. Pappu. 2008a. Genome structure and organization of a novel and distinct species of genus *Caulimovirus* (Family *Caulimoviridae*) associated with dahlia mosaic. Archives of Virology 153: 733-738.

Pahalawatta, V., K.B. Druffel, and H.R. Pappu. 2007. Incidence and relative prevalence of distinct caulimoviruses (Genus *Caulimovirus*, Family *Caulimoviridae*) associated with dahlia mosaic in *Dahlia variabilis*. Plant Disease 91: 1194-1197.

Pappu, H.R., K.L. Druffel, R. Miglino, and A.R. Van Schadewijk. 2008. Nucleotide sequence and genome organization of a new and distinct caulimovirus associated with dahlia mosaic. Archives of Virology 153: 2145–2148.

Pappu, H.R., S.D. Wyatt, and K.B. Druffel. 2005. Dahlia mosaic virus: Molecular detection and distribution in dahlia in the US. HortScience 40**:** 697-699.

Presting, G.G., O.P. Smith, and C.R. Brown. 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.

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

Saar, D.E. 1999. A phylogenetic analysis of the genus *Dahlia* (Asteraceae), DeKalb, Illinois, USA: Northern Illinois University, PhD thesis.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: 627-639.

Saar, D.E., and P.D. Sørensen. 2005. *Dahlia sublignosa*: a species in its own right. SIDA Contributions to Botany 21: 2161-2167.

Sherff, E.E. 1955. *Dahlia*. Compositae-Heliantheae-Coreopsidinae. North American Flora 2: 45–59.

Sørensen, P.D. 1969. Revision of the genus *Dahlia* (Compositae, Heliantheae-Coreopsidinae). Rhodora 71: 309-365, 367-416.

Sørensen, P.D. 1970. The *Dahlia*: An early history. Arnoldia 30: 121-138.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596-1599.

Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research 22: 4673-4680.

CHAPTER FOUR

BIOLOGICAL STUDIES ON THREE DISTINCT CAULIMOVIURSES IN DAHLIA (*DAHLIA VARIABILIS*)

ABSTRACT

Three distinct caulimoviruses were found to be associated with dahlia mosaic disease, *Dahlia mosaic virus* (DMV), DMV-D10, and Dahlia common mosaic virus (DCMV). DMV-D10 was shown to exist as an endogenous plant pararetorviral sequence in cultivated dahlias, whereas DMV and DCMV are typical members of the genus *Caulimovirus* and believed to exist only as episomal particles. Biological properties including mechanical and aphid transmission were investigated for DMV, DMV-D10 and DCMV using *Verbesina encelioides* as the indicator host. DMV and DCMV were transmitted by *Myzus persicae* at the rate of 50 and 73%, respectively. DMV and DCMV were mechanically transmissible to *V. encelioides*. The most effective buffer for mechanical transmission was phosphate buffer with nicotinic acid (1%) or sodium bisulfite (5%) with transmission rates of 58% and 36% for DMV and DCMV, respectively. DMV-D10 was not transmissible by either aphids or mechanical transmission. Availability of aphid and mechanical transmission protocols would facilitate further biological studies on caulimoviruses associated with *Dahlia*.

INTRODUCTION

Dahlia mosaic disease is one of the most prevalent and damaging viral diseases affecting *Dahlia variabilis* (Albouy, 1995; Pappu et al., 2005; Eid et al., 2009; Pearson et al., 2009). The occurrence of mosaic and stunting of dahlias was recorded since 1911 in the US (Albouy, 1995).

Symptoms of the dahlia mosaic disease include mosaic, vein clearing, vein-banding, leaf distortion, systemic chlorosis and overall stunting of the plant (Pappu and Wyatt, 2003; Pappu et al., 2005). Three distinct caulimoviruses are associated with dahlia mosaic disease: *Dahlia mosaic virus* (DMV) (Richins and Shepherd, 1983), Dahlia common mosaic virus (DCMV) (Nicolaisen, 2003; Pappu et al., 2008) and DMV-D10 (Pahalawatta et al., 2008a). They are members of the genus *Caulimovirus*, family *Caulimoviridae*, and possess a double-stranded (ds) DNA genome of ca. 7-8 kb (Pahalawatta et al., 2008a; Pappu et al., 2008; Richins and Shepherd, 1983). Members of this genus are considered plant para-retroviruses as these viruses exist in RNA phase in their hosts but code for a reverse transcriptase to synthesize and package dsDNA in progeny virions. The genome organization of DMV, DMV-D10 and DCMV is typical of members of the genus *Caulimovirus*, and consists of six major ORFs (ORFs I-VI) and one minor ORF (ORF VII). DMV-D10 however, shows divergence from typical known caulimoviruses including DMV and DCMV. Besides the lack of homology of ORFII encoding the aphid transmission factor of caulimoviruses, DMV-D10 ORFIV (encoding for coat protein) was found to be truncated and fused in frame with ORFV (encoding for reverse transcriptase) (Pahalawatta et al., 2008a). Pahalawatta et al. (2008b) reported that the genome of DMV-D10 is integrated into the dahlia genome and the virus exists as an endogenous plant pararetroviral sequence (DvEPRS). Recent surveys showed the prevalence of DMV-D10 followed by DCMV and DMV (Pappu et al., 2005; Pahalawatta et al., 2007; Eid et al., 2009). Despite their widespread occurrence, these viruses are considered to have limited host range. Natural infection is found only in dahlia species (Albouy, 1995; Brunt, 1971).

Previous report on mechanical transmission of what was considered to be DMV, showed that members of *Compositae*, *Solanaceae*, *Chenopodiaceae* and *Amaranthaceae* could be

infected with DMV (Brunt, 1971). *Zinnia elegans* and *Verbesina encelioides* were reported to be useful indexing hosts for DMV by Brierly (1951). Brunt (1971) confirmed that *V. encelioides* was the best propagating host. Aphid transmission of DMV by its principal vector, *Myzus persicae* (Sulz.) has been briefly described by Brierley and Smith (1950). Due to the diversity of caulimoviruses extant in dahlia and the lack of biological studies on DCMV and DMV-D10, this study was conducted to gain a better understanding of the characteristics of DCMV and DMV-D10 in comparison with DMV by aphid and mechanical transmission to *V. encelioides*.

MATERIALS AND METHODS

Virus source and transmission by *Myzus persicae*. Virus-infected dahlia plants were maintained in insect-proof cages and grown in a greenhouse with 20-24°C, and photoperiod of 16h. Mixed infection of DMV-D10 with either DCMV or DMV or presence of DMV-D10 alone was confirmed by PCR using species specific primers (Table 1). Non-viriluferous aphids (*Myzus persicae*) were reared on cabbage (*Brassica oleracea*) grown from seed under controlled conditions (20-24°C, photoperiod of 16h) in an insect proof cage. The colony was started from females feeding on *Physalis* plants provided by Dr. Eigenbrode (University of Idaho, Moscow, ID). The transmission assays were repeated twice and performed by allowing the aphids (20-30 aphids /leaf clip cage/plant) to acquire the virus from infected dahlias for 48h (Fig. 1). Aphids were then transferred to 2-4 leaf stage *V. encelioides* seedlings grown in 3 liter pots containing potting soil mixture. Aphids were allowed an inoculation access period of 72 h and then killed by spraying with Impidacloprid.

Mechanical inoculation. Different buffers were evaluated for their effect on transmission efficiency. Dahlia leaves infected with either DMV, or DCMV, or DMV-D10 were ground in 0.1M phosphate buffers (pH 7.0) with either 5% sodium bisulfite, or 0.03M sodium diethyldithiocarbamate, or 1% nicotinic acid or just distilled water. Sap extracts were used to rub-inoculate Carborundum-dusted leaves of *V. encelioides* seedlings at 2-4 leaf stage. The transmission assays were repeated twice for each buffer used. Inoculated plants were kept in a greenhouse (20-24°C, photoperiod of 16h).

Determination of virus presence by PCR. Virus transmission was confirmed by polymerase chain reaction (PCR). Total nucleic acids were extracted from *V. encelioides*-infected and healthy samples using the modified Dellaporta procedure (Presting et al., 1995) as described by Pahalawatta et al. (2007). Primer pairs specific to each of the three virus species were used (Table 1). A 1:20 dilution of the total nucleic acid extract was subjected to PCR using 2 µL template in a 20 µL total volume reaction containing 1X PCR buffer (20 mM Tris, pH 8.4 and 500 mM KCl), 150 µM dNTP mix, 2 mM $MgCl_2$, and 0.6 pmole each of sense and antisense primers, 12.7 µL sterile H_2O, and 0.1 µL Taq DNA polymerase. The thermal profile consisted of an initial denaturation step at 94°C for 4 min, followed by 50 cycles consisting each of three steps, 94°C for 30 s, the required annealing temperature (Table 1) based on the primer pair used, for 20 s, and 72 °C extension step determined based on the size of the amplicon to be synthesized (at the rate of 1000 bp /min), with a final incubation at 72 °C for 7 min. PCR products (7µL) were analyzed by agarose gel electrophoresis (1.2%) in 0.5 X TAE (Tris-acetate-EDTA) buffer.

Table 1. Primers used for each of the three caulimoviruses [*Dahlia mosaic virus* (DMV), Dahlia common mosaic virus (DCMV), and DMV-D10] to test plants for presence of virus.

Primers	Sequence (5'-3')	Annealing Temp °C	Expected size
DMV-D10 ORFI F	ATGGATCGTAAAGAT T	50	900 bp
DMV-D10 ORFI R	CTG TTT TTC TGT GTT TCT ACT GG		
DMV ORFI F	ATG AAT ATC TTA GAA AGG AA	50	939 bp
DMV ORFI R	CTT AAT CCT TAA GTT ATC AA		
DCMV CP F	GGATCCTCATTCTGAGTCTTCGTCTTC	59	1517 bp
DCMV CP R	CATATGGCCACCCAAATGACC		

Fig. 1. Aphids in leaf clip cage and on verbesina-inoculated leaf.

RESULTS AND DISCUSSION

DMV-D10 was not transmissible by aphids contrary to DMV and DCMV which were efficiently transmitted at rate of 50 and 73.3%, respectively. Symptoms appeared 3-4 weeks post-inoculation (Fig. 2, Table 2). Mosaic and chlorotic vein banding were major characteristic symptoms of DMV compared to mosaic only for DCMV on *V. encelioides*. Delaying flowering

by removing shoot apex to encourage axillary shoot leaves development enhanced symptom appearance on these leaves followed by general chlorosis (Fig. 2)

Mechanically inoculated plants showed symptoms 4 weeks post-inoculation (Fig. 3) and transmission was tested by PCR using species-specific primers. Inoculated verbesina plants developed symptoms ranging from chlorotic spots to vein banding and leaf distortion. DMV was more efficiently mechanically transmitted with all the buffers tested compared to DCMV (Table 3). The rate of transmission ranged from 36 to 58.3% and nicotinic acid (1% w/v) was most effective followed by sodium diethyldithiocarbamate. Mechanical transmission of DCMV ranged between 36 to 20% and the buffer with sodium bisulfite was most efficient.

Most caulimoviruses have limited host range and their transmission in nature is facilitated by many aphid species in a noncirculative manner (Shepherd, 1979; Prione and Blanc, 1996). Aphid transmission of caulimoviruses requires helpers or aphid transmission factor (ATF) which is the product of ORFII (P2) (Lung et al., 1974; Armour et al., 1983; Woolston et al., 1983). The DNA-binding protein product of ORFIII (P3) represents the second factor of transmission by aphids (Leh et al., 1999) serving as a bridge between P2 and the coat protein. *Cauliflower mosaic virus* (CaMV), type species of Caulimovirus group, is transmitted by aphids in a non-circulative manner (Gray and Banerjee, 1999). The transmission is mediated by interaction of ORFII protein (P2) with a non-identified attachment site in the aphid stylet and with the ORFIII protein (P3) associated with the virus capsid (encoded by ORFIV) (Woolston et al., 1987; Leh et al., 1999; Blanc et al., 2001). This interaction involves the C-terminal domain of P2 and N-terminal domain of P3 while the C-terminal domain of P3 interacts with the coat protein; thus P3 serves as a bridge between P2 and virus particles (Leh et al., 2001). In the case of DMV and DCMV, both proteins (P2 and P3) showed sequence similarities with those *Mirabilis mosaic virus* and

Figwort mosaic virus (Pappu et al., 2008) and the presence of amino acid sequence motif, IXG, X being any amino acid, necessary for the interaction between the ATF and virus particles for aphid transmission (Schmidt et al., 1994). On the contrary, DMV-D10 lacks the ORFII product (P2) but it conserved the presence of P3 in its genome (Pahalawatta et al., 2008a). The results of this study on aphid transmission of DMV, DCMV, and DMV-D10 confirmed the importance of P2 in virus transmission. A study by Jacquot et al. (1998) showed the importance of P3 (ORFIII product) of CaMV for infectivity on turnip plants by using point-mutation experiments. On the other hand, the vegetative propagation of dahlias and the host plant-virus co-evolution might have lead to DMV-D10 integration as an EPRS in the plant genome and the loss of P2 and conservation of P3. It was demonstrated that EPRS are not neutral components of plant genomes but can potentially contribute to pathogenicity (Lockhart et al., 2000; Richert-Pöggeler et al., 2003; Geering et al., 2005) and virus evolution.

An alternative method is mechanical transmission of caulimoviruses that requires wounds made by abrasive or broken leaf hairs for virus particles to enter and initiate infection. Not all caulimoviruses are easily transmitted mechanically (Lockhart et al., 2000; Zhang et al., 2008, Mollov et al., 2009); it requires investigating the right buffer and an experimental host. Buffers used with nicotinic acid, sodium diethyldithiocarbamate and sulfite were efficient for DMV and DCMV mechanical transmission. Nicotinic acid has high antioxidant activity associated with its reaction with hydrogen peroxide and prevents peroxide oxidation of lipids in biological membranes (Gromovaya et al., 2002). Sodium diethyldithiocarbamate acts as a chelating agent with specificity to copper (Barnett and Fulton, 1971). Copper is important for polyphenoloxidase activity that may result in virus inactivation. Sulfite probably inhibits the phenol oxidase by combining with quinone (Pierpoint, 1966). Repeated efforts to mechanically transmit DMV-D10

were unsuccessful. There are a few examples of caulimoviruses that were not successfully transmitted by mechanical inoculation or by *Myzus persicae* to index hosts, such as *Tobacco vein clearing virus* (TVCV) which also was found to be integrated in the host genome (Lockhart et al., 2000), Lamium leaf distortion-associated virus (Zhang et al., 2008), and Rose yellow vein virus (Mollov et al., 2009). However, these results should be interpreted with caution. It is probably required to use different buffers for DMV-D10 or different aphid species since its ORFII product has no similarity with that of DMV and DCMV and is of smaller size. There could be some unknown reasons for loss of DMV-D10 transmission by mechanical inoculation and could be its integration in host genome has exerted a selection for virus structure that is probably less compatible with aphid and mechanical transmission. Loss of aphid transmissibility following prolonged vegetative propagation was documented for several viruses (Husted, 1995; Delmer et al., 1997; Domier et al., 2007) or it is not mechanically transmissible.

This study compared the biological properties of the three caulimoviruses associated with dahlia (DMV, DCMV, and DMV-D10). Differences in their transmission rates reflect the divergence of these viruses as distinct species which was in agreement with the divergence in their respective genomic sequences.

Table 2. Aphid transmission of the three caulimoviruses [*Dahlia mosaic virus* (DMV), Dahlia common mosaic virus (DCMV), and DMV-D10]

Virus	Source plant/ test plant	Infected/ tested	Transmission(%)
DMV	Dahlia/ Verbesina	9 ／ 18	50
DCMV	Dahlia/ Verbesina	11 ／ 15	73.3
DMV-D10	Dahlia/ Verbesina	0 ／ 19	0

A

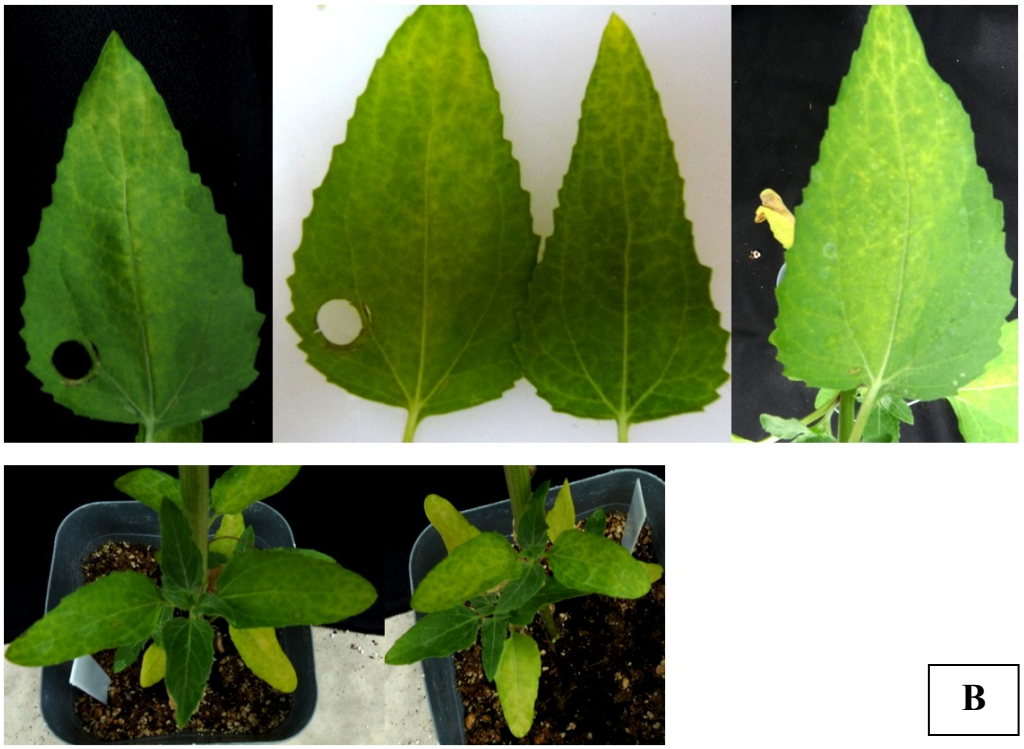


Fig. 2. Symptoms on verbesina following aphid transmission. A: *Dahlia mosaic virus* (DMV) symptoms on verbesina; B: Dahlia common mosaic virus (DCMV) symptoms on verbesina.

Table 3. Mechanical transmission of the three caulimoviruses [*Dahlia mosaic virus* (DMV), Dahlia common mosaic virus (DCMV), and DMV-D10]

Virus	Source plant/ test plant	Buffer	Infected/ tested	Transmission (%)
DMV	Dahlia / Verbesina	Water	9／25	36
	Dahlia / Verbesina	Phosphate buffers + 1% nicotinic acid	7／12	58.33
	Dahlia / Verbesina	Phosphate buffers + 5% sodium bisulfite	7／17	41.17
	Dahlia / Verbesina	Phosphate buffers + 0.03M sodium diethyldithiocarbamate	6／11	54.54
DCMV	Dahlia / Verbesina	Water	2／ 9	22.22
	Dahlia / Verbesina	Phosphate buffers + 1% nicotinic acid	1／ 5	20
	Dahlia / Verbesina	Phosphate buffers + 5% sodium bisulfite	4／ 11	36.36
	Dahlia / Verbesina	Phosphate buffers + 0.03M sodium diethyldithiocarbamate	2／ 8	25
DMV-D10	Dahlia / Verbesina	Water	0／ 10	0
	Dahlia / Verbesina	Phosphate buffers + 1% nicotinic acid	0／ 10	0
	Dahlia / Verbesina	Phosphate buffers + 5% sodium bisulfite	0／ 10	0
	Dahlia / Verbesina	Phosphate buffers + 0.03M sodium diethyldithiocarbamate	0／ 10	0

A

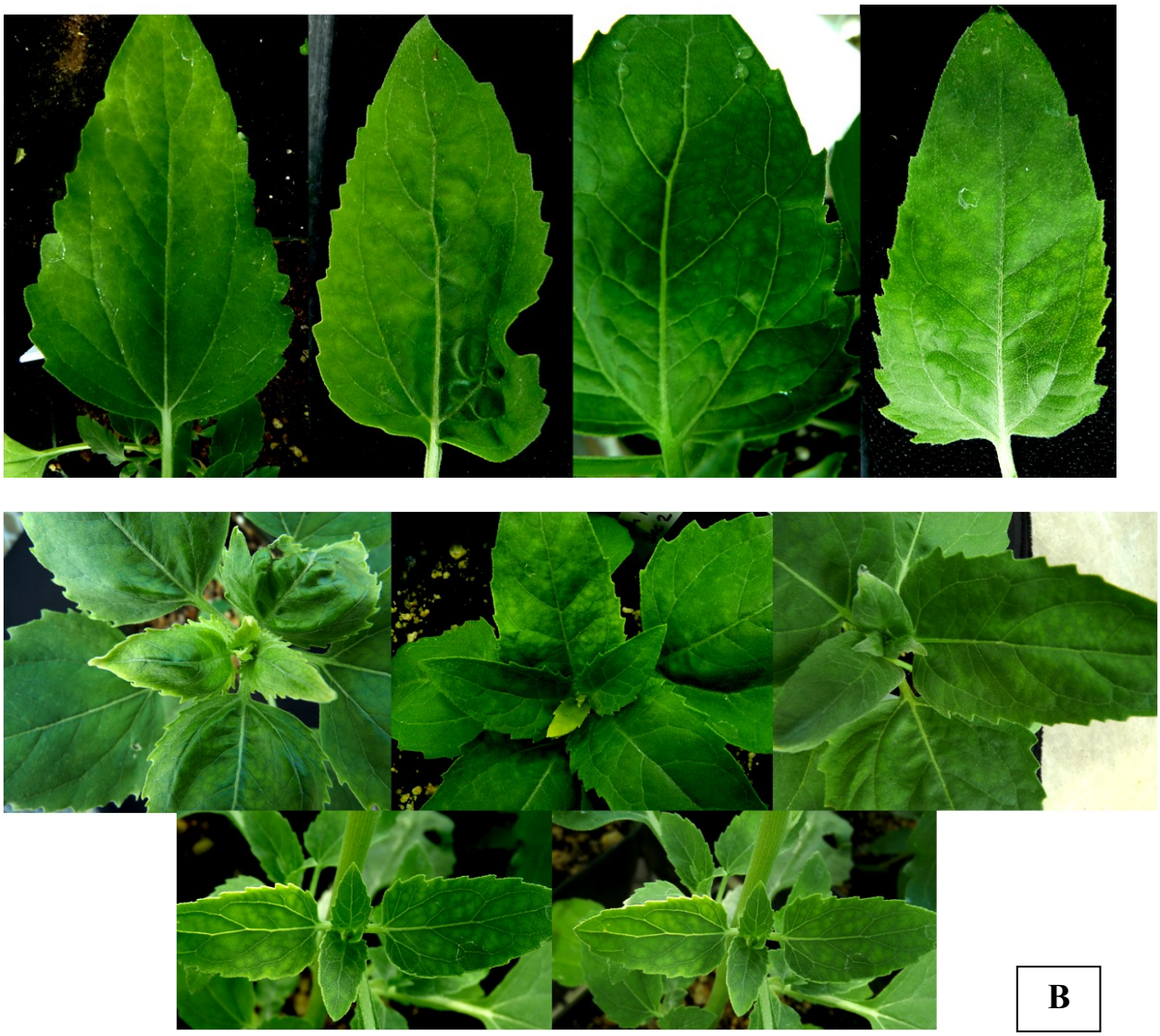


Fig. 3. Symptoms on verbesina following Mechanical inoculation. A: *Dahlia mosaic virus*-(DMV) symptoms on verbesina; B: Dahlia common mosaic virus (DCMV) symptoms on verbesina.

REFERENCES

Albouy, J. 1995. Dahlia, p.265-273. In: G. Lobenstein, R.H. Lawson, and A.A. Brunt (eds.). Virus and virus-like diseases of bulb and flower crops. John Wiley and Sons, Chichester, NY.

Armour, S.L., U. Melcher, T.P. Pirone, D.G. Lyttle, and R.C. Essenberg. 1983. Helper component for aphid transmission encoded by region II of cauliflower mosaic virus DNA. Virology 129:25–30

Barnett, O.W. and R.W. Fulton. 1971. Differential response of *Prunus* necrotic ringspot and tulare apple mosaic viruses to stabilizing agents. Virology 46: 613-619.

Blanc, S., E. Hébrard, M. Drucker, and R. Froissart. 2001. Molecular basis of vector transmission: caulimoviruses, p.143-166. In : K., Harris, O.P., Smith, and J.E. Duffus (eds.). Virus-Insect-Plant Interactions. San Diego, CA: Academic Press.

Brierley, P. and F.F. Smith. 1950. Some vectors, hosts and properties of dahlia mosaic viruses. Plant Disease Reporter 35: 405-407.

Brierly, P. 1951. Value of index for detecting dahlia viruses. Plant Disease. Reptr. 35: 405.

Brunt, A. A. 1971. Some hosts and properties of dahlia mosaic virus. Annals of Applied Biology 67: 357-368.

Demler, S.A., D.G. Rucker-Feeney, J.S. Skaf, and G.A. de Zoeten. 1997. Expression and suppression of circulative aphid transmission in pea enation mosaic virus. Journal of General Virology 78:511–523.

Domier, L.L., T.A. Steinlage, H.A. Hobbs, Y. Wang, G. Herrera-Rodriguez, J.S. Haudenshield, N.K. McCoppin, and G.L. Hartman.2007.Similarities in seed and aphid transmission among soybean mosaic virus isolates. Plant Disease 91:546–550.

Eid ,S., K.L. Druffel, D. Saar, and H.R. Pappu. 2009. Incidence of multiple and distinct species of caulimoviruses in dahlias (*D. variabilis*). HortScience 44:1498-1500.

Geering, A.D.W., N.E. Olszewski, G. Harper, B. E.L. Lockhart, R. Hull, and J.E. Thomas. 2005. Banana contains a diverse array of endogenous badnaviruses. Journal of General Virology 86: 511–520.

Gray, S.M., and N. Banerjee. 1999. Mechanisms of arthropod transmission of plant ad animal viruses. Microbiology and Molecular Biology Reviews 63: 128-148.

Gromovaya, V.F., G.S. Shapoval, and I.E. Mironyuk. 2002. Antiradical and antioxidant activity of biologically active carboxylic acids. Russian Journal of General Chemistry 72: 774-777.

Husted, K. 1995. Reason for non-aphid transmissibility in a strain of Kalanchoe¨ mosaic potyvirus. Virus Genes 1:59–61.

Jacquot, E., A. Geldreich, M. Keller, and P. Yot. 1998. Mapping regions of the cauliflower mosaic virus ORFIII product required for infectivity. Virology 242: 395-402.

Leh, V., E. Jacquot, A. Geldreich, M. Haas, S. Blanc, M. Keller, and P. Yot. 2001. Interaction between cauliflower mosaic virus ORFIII product and the coat protein is required for transmission of the virus by aphids. Journal of Virology 75: 100-106.

Leh, V., E. Jacquot, A. Geldreich, T. Hermann, D. Leclerc, M. Cerrutti, P. Yot, M. Keller, and S. Blanc. 1999. Aphid transmission of cauliflower mosaic virus requires the viral PIII protein. European Molecular Biology Organization Journal 18: 7077-7085.

Lockhart, B.E., J. Menke, G. Dahal, and N. E. Olszewski. 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. Journal General Virology 81: 1579-1585.

Lung, M.C.Y., and T.P. Pirone. 1974. Acquisition factor required for aphid transmission of purified cauliflower mosaic virus. Virology 60: 260-264.

Mollov, D, B. Lockhart, and D. Zlesak. 2009. Genome characterization and transmission of Rose yellow vein virus, a new caulimovirus occurring in garden rose. Phytopathology 99 : S87.

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

Pahalawatta, V., K.L. Druffel, S.D. Wyatt, K.C. Eastwell, and H.R. Pappu. 2008a. Genome structure and organization of a novel and distinct species of genus *Caulimovirus* (Family *Caulimoviridae*) associated with dahlia mosaic. Archives of Virology 153: 733-738.

Pahalawatta, V., K.L. Druffel, and H.R. Pappu. 2008b. A new and distinct species in the genus *Caulimovirus* exists as an endogenous plant pararetroviral sequence in its host, *Dahlia variabilis*. Virology 376: 253-257.

Pahalawatta, V., R. Miglino, K.L. Druffel, A. Jodlowska, A.R. Van Schadewijk, and H.R. Pappu. 2007. Incidence and relative prevalence of distinct caulimoviruses (Genus *Caulimovirus*, Family *Caulimoviridae*) associated with dahlia mosaic in *Dahlia variabilis.* Plant Disease 91:1194-1197.

Pappu, H. R., K. L. Druffel, R. Miglino, and A.R. Van Schadewijk. 2008. Nucleotide sequence and genome organization of a new and distinct caulimovirus associated with dahlia mosaic. Archives of Virology 153:2145–2148.

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

Pappu, H.R., and S.D. Wyatt. 2003. Dahlia mosaic virus. A description of symptoms on dahlia. APSnet Image of the Week. http://apsnet.org/online/archive/2003/IW000032.asp.

Pearson, M.N., D. Cohen, S.J. Cowell, D. Jones, A. Boulin, B.S.M. Lebas, J.B. Shiller, and G.R.G Clover. 2009. A survey of viruses of flower bulbs in New Zealand. Australian Plant Pathology 38: 305-309.

Pierpoint, W.S. 1966. The enzymic oxidation of chlorogenic acid and some reactions of the quinine produced. Biochemical Journal 98: 567-580.

Presting, G.G., O.P. Smith, and C.R. Brown. 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.

Pirone, T. P., and S. Blanc. 1996. Helper-dependent vector transmission of plant viruses. Annual Review of Phytopathology 34:227-247.

Richert-Poggeler, K. R., F. Noreen, T. Schwarzacher, G. Harper, and T. Hohn, T. 2003. Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia. European Molecular Biology Organization Journal 22: 4836-4845.

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

Schmidt, I., S. Blanc, P. Esperandieu, G. Kuhl, G. Devauchelle, C. Louis, and M. Cerutti. 1994. Ineteraction between the aphid transmission factor and virus particles is part of the molecular mechanism of cauliflower mosaic virus aphid transmission. Proceedings of the National Academy of Science USA 91: 8885-8889.

Shepherd, R.J. 1979. DNA plant viruses. Annual Review of Plant Physiology 30: 405-423.

Woolston, C.J., S.N. Covey, J.R. Penswick, and J.W. Davies. 1983. Aphid transmission and a polypeptide are specified by a defined region of the cauliflower mosaic virus genome. Gene 23: 15-23.

Woolston, C.J., L.G. Czaplewski, P.G. Markham, A.S. Goad, R. Hull, and J.W. Davies. 1987. Location and sequence of a region of cauliflower mosaic virus gene 2 responsible for aphid transmissibility. Virology 160: 264-251.

Zhang, L., B. Lockhart, G. Dahal, and N. Olszewski. 2008. Studies on biology and genomic characterization of a caulimo-like virus associated with a leaf distortion disease of Lamium maculatum. Archives of Virology. 153: 1181-1184.

CHAPTER FIVE

GENOMIC CHARACTERIZATION OF PARA-RETROVIRAL SEQUENCES IN WILD *DAHLIA* SPP IN NATURAL HABITATS

ABSTRACT

More than 35 species of *Dahlia* are endemic to the mountain ranges of central Mexico. A survey on the incidence of the three caulimoviruses (*Dahlia mosaic virus* DMV, D10-US, and Dahlia common mosaic virus DCMV) in wild dahlia species showed the presence of DMV-D10 sequences and no evidence of DMV or DCMV. To better understand the possible emergence, co-existence and co-evolution of para-retroviruses and their host plants, the genome structure and organization of DMV-D10 from three wild species *D. coccinea* (D10-DC), *D. sherffii* (D10-DS) and *D. tenuicaulis* (D10-DT) were determined and compared to that of D10-US from cultivated species *D. variabilis*. The complete ca. 7 kb dsDNA genome of D10-DC, D10-DS, and D10-DT has the structure and organization typical of a *Caulimovirus* species and shared 89.3 to 96.6% amino acid sequence identity among various ORFs when compared to those of D10-US from *D. variabilis.* The lack of an aphid transmission factor homolog (ATF) and the partial coat protein fused with the reverse transcriptase (RT) open reading frame are common among these DMV-D10 isolates from wild and cultivated *Dahlia* species.

INTRODUCTION

The genus *Dahlia* is within the Asteraceae family and it was first described by Cavanilles in 1791. Currently, 35 wild species are recognized plus the cultivated forms, known as either *D. pinnata* or *D. variabilis.* Dahlias are perennial plants used as garden or potted ornamentals and it is an economically important bulb crop in the US and several parts of the world. Most species of *Dahlia* are native to the mountain ranges in central Mexico (Sørensen, 1969; Saar, 1999). Mexico is considered to have the greatest diversity for *Dahlia*.

Dahlia production is threatened by many viral diseases due to the vegetative propagation method adopted for dahlias. Dahlia mosaic disease is one of the most important and widely prevalent diseases in cultivated dahlias (Brunt, 1971; Albouy, 1995). The most characteristic symptoms of dahlia mosaic include mosaic, vein clearing, vein-banding, leaf distortion, systemic chlorosis, flower breaking, and overall stunting of the plant. Symptoms vary depending on the cultivar, and cultivars with few or no symptoms are common. The disease is caused by *Dahlia mosaic caulimovirus* (DMV) which is a distinct species in the genus *Caulimovirus* based on the molecular characterization (Richins and Shepherd, 1983). Recently, it was shown that in addition to DMV, there are two more distinct caulimoviruses associated with dahlia. They are tentatively referred as Dahlia common mosaic virus (DCMV) (Pappu et al., 2008) and D10-US (referred to as DMV-D10 in Pahalawatta et al., 2008a). Both DCMV and D10-US were found to be widespread based on the surveys conducted in the US and Europe (Pappu et al., 2005; Pahalawatta et al., 2007a; Eid et al., 2009).

D10-US was recently shown to exist as an endogenous plant pararetrovirus (DvEPRS) (Pahalawatta et al., 2008b). Most of the research considering the prevalence of caulimoviruses was done in cultivated dahlia but little is known about their occurrence in natural populations of

wild dahlia spp. We investigated their potential presence in genus *Dahlia*'s in their natural habitats from west central Mexico. D10-US was found to be predominant (91%) of the tested samples (Eid et al., 2010) while DMV and DCMV were not found. Study of this naturally occurring D10-US in *D. coccinea* (one of the parents of *D. variabilis*), *D. sherffii*, and *D. tenuicaulis* could provide valuable insights into the molecular aspects of host-virus interactions and the evolutionary pathways of plant-associated para-retroviruses. Here we report the genome organization and sequence diversity of DMV-D10 form occurring D10-DC (from *D. coccinea*), D10-DS (from *D. sherffii*), and D10-DT (from *D. tenuicaulis*).

MATERIALS AND METHODS

Plant material. Dahlia leaf samples were collected, during summer 2006, from wild dahlia species *D. coccinea, D. sherffii,* and *D. tenuicaulis,* found in their natural wild habitats in the mountain ranges of central Mexico.

DNA extraction and polymerase chain reaction. Total gnomic DNA was extracted from infected dahlia leaves using the DNeasy plant mini kit (Qiagen Inc, Valencia, CA) according to manufacturer's instructions. The quality of the DNA was ensured by polymerase chain reaction (PCR) for the dahlia internal transcribed spacer (ITS) region using specific primers (Table 1). D10-US specific primers were used to obtain the whole genome (Table 1). The PCR reaction included 10 ng DNA, 4 µl 5X buffer (Promega, Madison, WI) buffer and a final concentration of 150 µM dNTP's, 2 mM $MgCl_2$ and 0.6 µM of each primer and 10U GoTaq® Flexi DNA Polymerase (Promega, Madison, WI) in a total reaction volume of 20 µl. PCR amplification was performed in a DNA thermal cycler (BioRad, Hercules, CA) programmed for 4 min at 94°C for

initial denaturation and 50 cycles each consisting of 30 sec at 94^{o}C, 25 sec at T_m-5^{o}C and 1 min extension per 1000 bp product at 72^{o}C, followed by a final extension for 10 min at 72^{o}C.

Cloning and sequencing. PCR products of expected size were cloned using pGEM-T (Promega, Madison, WI) cloning kit according to manufacturers' instructions. The ca. 7-7.2 kb of D10-DC, D10-DS, and D10-DT genomic DNA was cloned in 6 cloning steps (Fig. 1b, 2b, 3b). Recombinant plasmids were isolated (Sambrook et al., 1989) and were sequenced using M13 sense and antisense primers. 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.

Phylogenetic 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). Pair-wise and multiple alignments were done using CLUSTAL W (Thompson et al., 1994). 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-D10 open reading frame (ORF) with ORFs of other caulimoviruses (Table 2) was done using CLUSTALW version 1.83 (Thompson et al., 1994) and MEGA4 (Tamura et al., 2007).

Table 1. Primers used to cover DMV-D10 whole genome in wild *Dahlia* species

Fragments	Primers pair	D10-DC	D10-DS	D10-DT
		Sequence (5'-3')	Sequence (5'-3')	Sequence (5'-3')
A	Sense	TGG TAT CAG AGC CTT GTG C	CTT TTA TTG CTT TAA TAC TG	TGG TAT CAG AGC CTT GTG C
	Antisense	TTC TAA AGG TTG CTT GC	TTC TAA AGG TTG CTT GC	TTT CTG TGT TTC TAC TGG
B	Sense	GAG GTT ATA GCA GCA AA	GAA AAC CTA GAG GTT ATA GC	AAT TGA GTA TAA AGA AAA GGA GAG
	Antisense	GTG GGA ATA GGA AAT GG	GTG GGA ATA GGA AAT GG	AGG TCA TGA GTT GAA CGA
C	Sense	GAG GAT CTA TGG GAA AAC AC	GAG GAT CTA TGG GAA AAC AC	AGG TAA TTT TAT CTT GTG GA
	Antisense	ATT ATG GCT GTC TCC TTT GG	GGT TTC CTG AGC TCC GCA AG	CAA GAA AAG CTG GAG AC
D	Sense	TGG ATG AAT GCC TCA ATA GA	GAT CAA GGG ACA CAT TGT	TGG ATG AAT GCC TCA ATA GA
	Antisense	GTC TGA ATT CGT CTA AAG AG	TGT TTC TAG ACT TTC GTC G	AGA TTC AGT TTG TAA CCT CG
E	Sense	TCC GTT GAT GGC TGA CAG TT	G AAA AAA TTC TTT TAA TA	TCTCCATTTATCTTACTCCG
	Antisense	AGG TAG GGG TTT ACT CAC ACA TT	GTC TGA ATT CGT CTA AAG AG	TCT GGT AGA GTA GAG TTG C
F	Sense	CAC ATG GGT ATG TCG ACT AAA TAC G	CAA ATC CGT TGA TGG CT	AG ACA AAT GAA AAG AGA ACG
	Antisense	TAA AGT TTT TAA TTT TTC TTT TTC AA	TTTATATCTACTGACAATTCTTC	TTT ATA TCT ACT GAC AWT TCT T
Dahlia ITS[a]	Sense	AAT GCG ATA CTT GGT GTG AA		
	Antisense	TCG AAG CAT CAT AAC AAG AC		

[a] Internal transcribed spacer region of dahlias

Table 2. List of caulimoviruses used for sequence comparisons

Caulimoviruses	ORF I	ORF III	ORF IV	ORF V	ORF VI
Figwort mosaic virus (FMV)	NP_619544	NP_619546	NP_619547.1	NP_619548	NP_619549
Strawberry vein banding virus (SVBV)	NP_043929	NP_043931	NP_043932	NP_043933	NP_043934
Horseradish latent virus (HLV)	AAW56085	AAW56087	AAW56088	AAW56089	AAW56090
Miriabilis mosaic virus (MMV)	NP_659393	NP_659395	NP_659396	NP_659397	NP_659398
Blueberry red ringspot virus (BRRV)	NP_395472		NP_395468	NP_395469	NP_395470
Commelina yellow mottle virus (ComYMV)	NP_039818				
Cestrum yellow leaf curling virus (CmYLCV)	NP_861411	NP_861408		NP_861410	NP_861547
Carnation etched ring virus (CERV)	NP_612573	NP_612575	NP_612576	CAA28360	NP_612578
Peanut chlorotic streak virus (PSCV)	NP_042515		NP_042512	NP_042513	NP_042514
Soybean chlorotic mottle virus (SbCMV)	NP_044299	NP_068727	NP_068728	NP_068729	NP_068730
Cauliflower mosaic virus (CaMV)	Q02968	Q00967	Q00956	Q00962	Q00957
Dahlia mosaic virus (DMV)	AAP44107	AAP44109		AAP75615	AAP75616
Dahlia common mosaic virus (DCMV)	ABW81759	ABW81761		ABW81763	ABW81764
Dahlia mosaic virus D10 (DMV-D10)	ABW80579	ABW80580		ABW80581	ABW80582

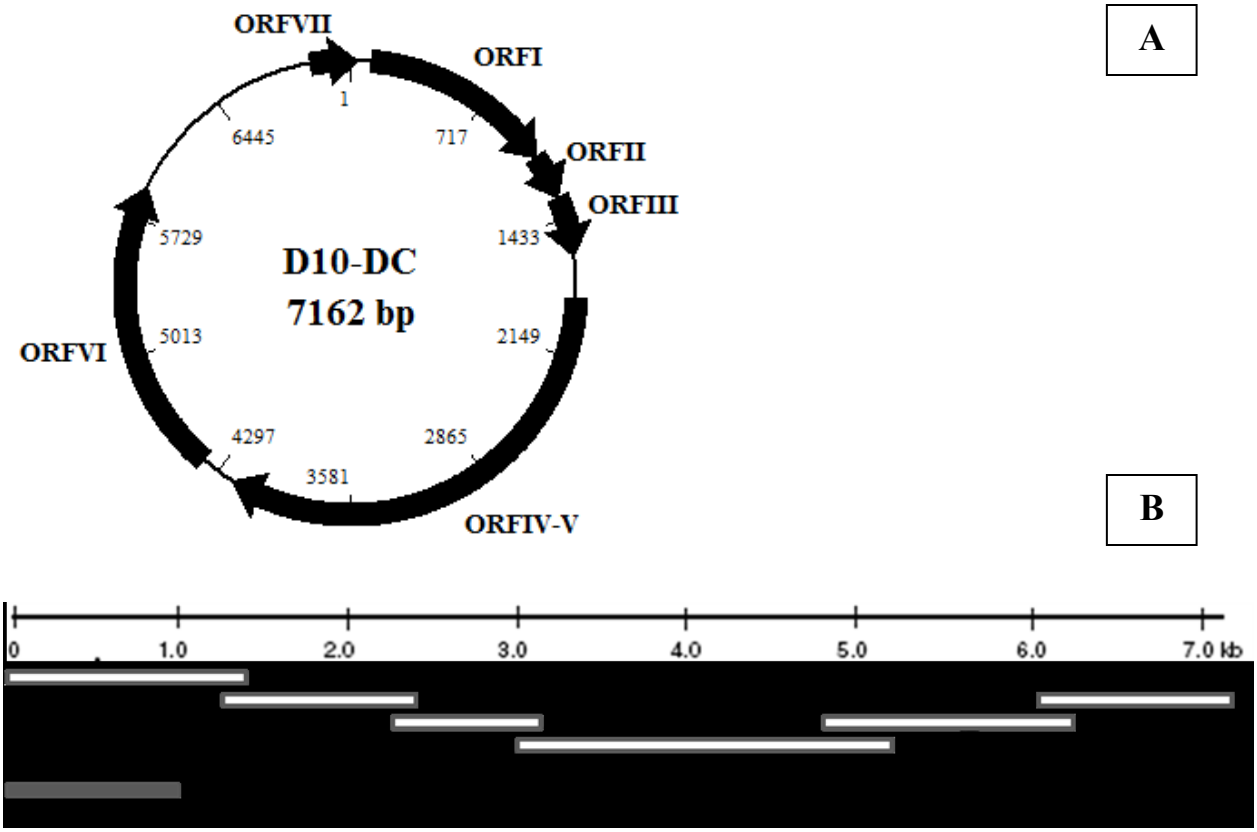


Fig. 1. A. Dahlia mosaic virus-D10 genome organization from *Dahlia. coccinea* (D10-DC). The closed circle represents the viral genome. Lines with arrows represent the various open reading frames. ORFs I, III, V, and VI code for putative movement protein, DNA binding protein, reverse transcriptase and inclusion body protein, respectively. **B.** A linear representation of the locations of the clones used in the sequencing of the entire genome. Letters A-F represent clones. The size of each clone is indicated in base pairs (bp). The hatched box represents a single clone

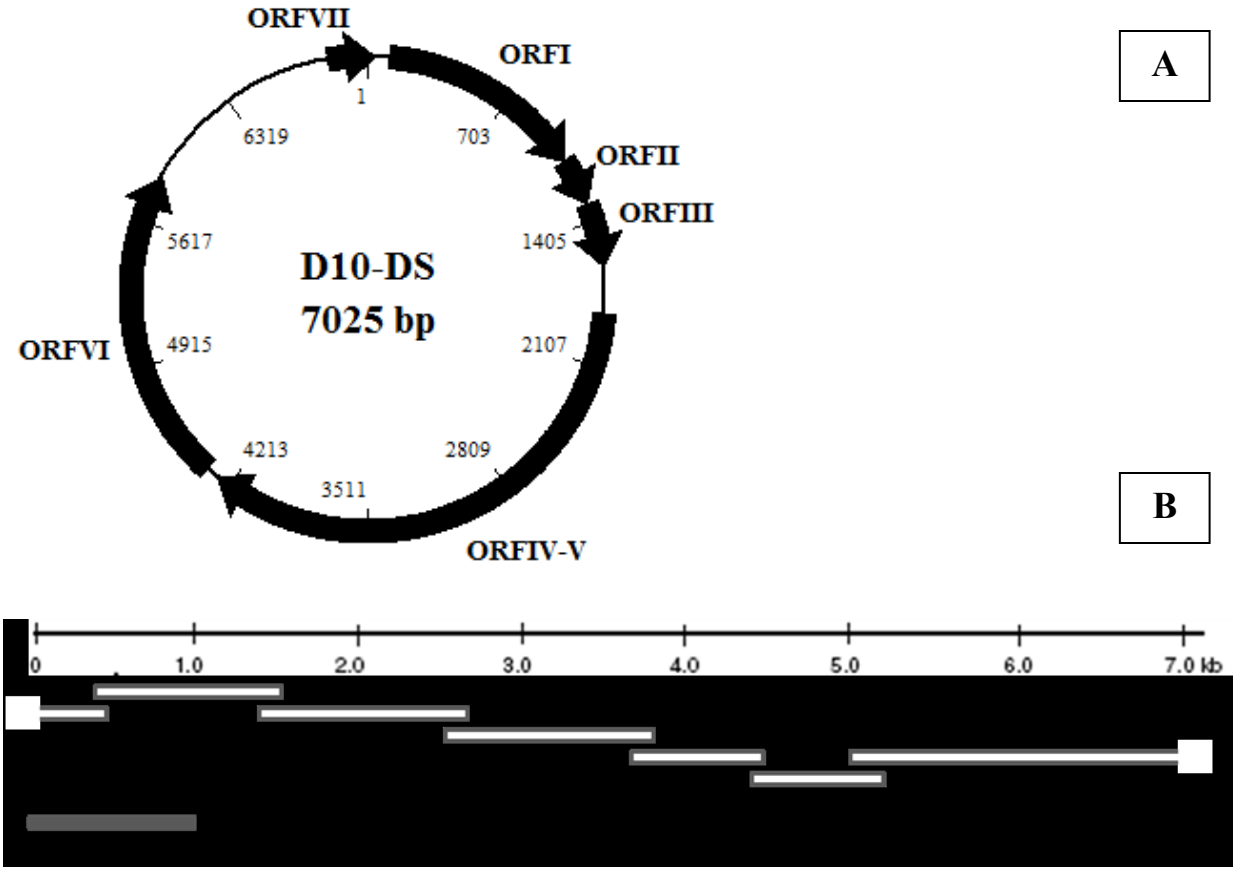


Fig. 2. A. Dahlia mosaic virus-D10 genome organization from *Dahlia sherffii* (D10-DS). The closed circle represents the viral genome. Lines with arrows represent the various open reading frames. ORFs I, III, V, and VI code for putative movement protein, DNA binding protein, reverse transcriptase and inclusion body protein, respectively. **B.** A linear representation of the locations of the clones used in the sequencing of the entire genome. Letters A-F represent clones. The size of each clone is indicated in base pairs (bp). The hatched box represents a single clone

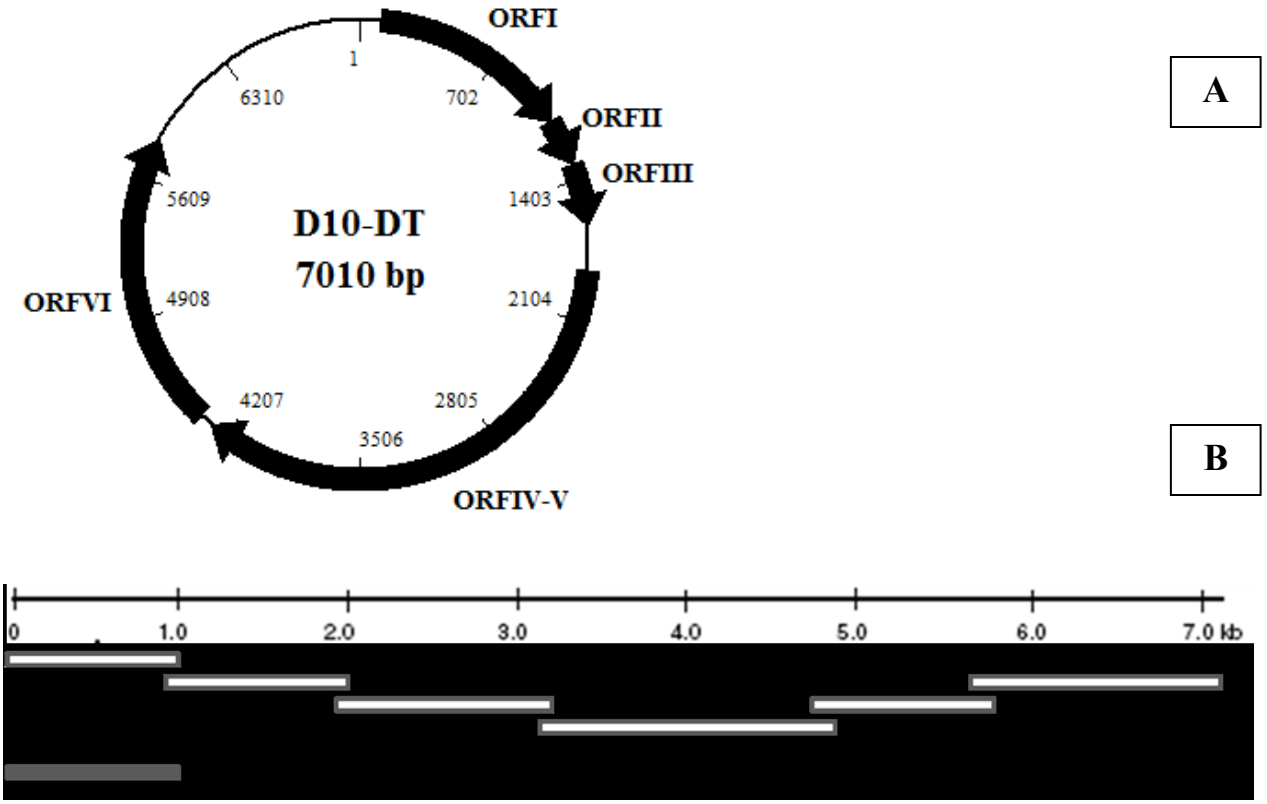


Fig. 3. A. Dahlia mosaic virus-D10 genome organization from *Dahlia tenuicaulis* (D10-DT). The closed circle represents the viral genome. Lines with arrows represent the various open reading frames. ORFs I, III, V, and VI code for putative movement protein, DNA binding protein, reverse transcriptase and inclusion body protein, respectively. **B.** A linear representation of the locations of the clones used in the sequencing of the entire genome. Letters A-F represent clones. The size of each clone is indicated in base pairs (bp). The hatched box represents a single clone

RESULTS AND DISCUSSION

Sequence analysis. The viral genome of DMV-D10 occurring naturally in wild *D. coccinea* (D10-DC), *D. sherffii* (D10-DS) and *D. tenuicaulis* comprises 7162, 7025 and 7010 nucleotides respectively (Fig. 1a, 2a, 3a). G+C composition of D10-DC, D10-DS and D10-DT is 33.14, 33.49, and 32.82 % respectively. This G+C percentage is higher than that of D10-US (30.5% G+C) from cultivated dahlias (Pahalawatta et al., 2008a) but slightly lower than that of CaMV, DCMV, CERV, and SoyCMV genome of 40, 37.4, 36.4, and 34.0 % G+C respectively (Pappu et al., 2008; Hasegawa et al., 1989; Hull et al., 1986). The numbering begins at 5' end of the (-) strand primer binding site (5'- TGGTATCAGAGC -3') which shows complementarity to the first twelve 3'-terminal nucleotides of $tRNA^{Met}$ from plant cytoplasm (Canaday et al., 1980; Ghosh et al., 1982). Computer analysis of the (+) strand predicted five complete putative open reading frames (ORFs). The organization of these ORFs was typical to the members of the genus *Caulimovirus*. They are tightly clustered and separated by one intergenic region (Hull, 2002) (Fig. 1a, 2a, 3a).

Coding regions. ORFI of D10-DC, D10-DS, and D10-DT is 990-993 bp long encoding the movement protein (MP). The domains GNLAYGKFMFAVY in D10-DC and GNLAYGKFMFTVY in D10-DS, and D10-DT were identified which were also conserved in other caulimoviruses. This domain is important for cell-to-cell movement (Glasheen et al., 2002; Hasegawa et al., 1989). Sequence comparison of D10-DC, D10-DS, and D10-DT with D10-US showed sequence divergence of 3.9 to 8.3 % and 10.7 to 4.6 % at nucleotide and amino acid levels, respectively (Table 3a). By comparison with DMV and DCMV, D10-DC, D10-DS, and D10-DT showed sequence identity of 45.3 to 46.6% (Table 3a). Phylognetic analysis showed

clustering of the MP of D10-US, D10-NZ, D10-DC, D10-DS, and D10-DT together and they were closest to that of *Figwort mosaic virus* (FMV) (Fig. 4a).

ORF II encodes for aphid transmission factor (ATF) in most caulimoviruses (Pirone and Blanc, 1996). Its homologous ORF was not identified in either D10-DC or D10-DS or D10-DT as well in D10-US (Pahalawatta et al., 2008a). A smaller ORF was found and the sequence identity with that of D10-US ranged between 54.5 to 81.8%. The IXG motif, common to caulimoviruses ATF, was absent.

ORF III codes for DNA-binding protein (DNAb). The N terminal of ORFIII protein was shown to be required for aphid transmission of *Cauliflower mosaic virus* (CaMV) (Leh et al., 1999) and it acts as a second "helper" factor for CaMV transmission by aphids. The D10 sequences from wild and cultivated dahlias are clustered together by phylogenetic analysis (Fig. 4b) and very close to DCMV. The sequence identity of D10-DC, D10-DS, and D10-DT with D10-US ranged 91.1 to 96% while it diverged of 69.9 to 71.2 % with DCMV and DMV (Table 3b). Jacquot et al. (1998) have shown that ORFIII product is essential for infectivity within the plant host.

ORF IV in D10-DC, D10-DS, and D10-DT found fused in-frame with ORF V due to absence of stop codon and consisted of 100 aa contrary to 480 aa of known caulimoviruses containing the HX4C motif in D10-DC and D10-DT but not in D10-DS (Pahalawatta et al., 2008a). The percentage of amino acid sequence identity of D10-DC, D10-DS, and D10-DT with DMV and DCMV was low (4.4 to 6.1 %).

ORF V of D10-DC, D10-DS, and D10-DT contained conserved amino acid sequences similar to those of D10-US and those reported in other caulimoviruses: YVDTGASLC (aa 194-202/ 196-204/198-206) amino acid sequence for putative protease domain (Glasheen et al., 2002;

Hasegawa et al., 1989; Torruella et al., 1989). The RT domain for caulimoviruses was present in D10-DC, D10-DS and D10-DT as YVDGIIVF (D10-DC, aa 535-542) and YVDDIIVF (D10-DS and D10-DT, aa 537-544/539-546) (Glasheen et al., 2002; Hasegawa et al., 1989). The sequences for RNase H motifs reported in other caulimoviruses and D10-US (Stavolone et al., 2003; Pahalawatta et al., 2008a) also found: EAVCRYTSGSFK (D10-DC, aa 703-714), EEVCRYTSGSFK (D10-DS, and D10-DT, aa 705-716/707-718), IIETDASNDFWG (D10-DC, and D10-DS, aa 681-692) / IIETDASNDYWG (D10-DT, aa 689-692), and EHLEGSKNVLADCL (D10-DC, and D10-DS, aa 787- 800/683-694) / EHLEGLKNVLADCL (D10-DT, aa 791-804). D10-DC, D10-DS, D10-DT, D10-US and D10-NZ were tightly clustered together and formed one big clade with FMV, MMV, DMV, and DCMV by phylogenetic analysis (Fig. 4c). The amino acid sequences showed 34.7 to 42.8 % divergence with DCMV and DMV, and 6.1 to 7.7 % with D10-US (Table 3c).

The product of ORFVI in caulimoviruses is a multifunctional protein called trans-activator or TAV. It regulates the translation of the viral polycistronic mRNA (Bonneville et al., 1989; Daubert et al., 1983). It is involved in inclusion body formation, and virus assembly. It influences both disease severity and determination of host range (Hohn and Fütterer, 1997; Kobayashi et al., 1998; Schoelz and Shepherd, 1988). It was found that transgenic plants expressing CaMV- ORFVI contained inclusion bodies and symptoms were sometimes observed (Covey et al., 2000), and the authors suggested that ORFVI product has elicitor-like functions triggering, directly or indirectly, a host response and leading to symptom expression. The ORFVI of D10-DC, D10-DS, and D10-DT conserves the key domains (letters in capital) of CaMV TAV (GLTKYIY; aa 303-309/302-308/289-295) (Richins et al., 1987), similar to the amino acid sequence motif that is well conserved in D10-US (Pahalawatta et al., 2008a) and in other

caulimoviruses (Glasheen et al., 2002; Hasegawa et al., 1989). Phylogenetic relationships among ORFVI protein of caulimoviruses showed clustering of D10-DC, D10-DS, and D10-DT with D10-US and D10-NZ and forming all one cluster with FMV (Fig. 4d). The amino acid sequence comparison with D10-US showed 93 to 96.6% identity and did diverge of 61.4 to 62.8 % with DCMV and DMV (Table 3d).

ORFVII of D10-DC and D10-DS encodes 82 and 73aa protein respectively but no functional proteins have been found *In planta* (Wurch et al., 1990). ORFVII was not found in D10-DT. This ORF in D10-DC and D10-DS is slightly bigger than that of D10-US and with sequence identity of 79.6 to 82.3% at the nucleotides level.

Non-coding regions. The D10-DC, D10-DS, and D10-DT genomes each contain one large intergenic region between ORFs VI and VII/ I, similar to that of D10-US (Fig. 1a, 2a, 3a). This region contained a putative 35S promoter homolog. The TATA box (TATATAA) is followed by a poly-adenylation signal (AATAAA) (Fig. 5). The transcriptional activity of 35S promoter in caulimoviruses is a result of synergistic and combinatorial effects of enhancer elements (Benfey and Chua, 1990; Benfey et al., 1989, 1990 a,b). The known activating sequence factor -1 is found in D10-DC, D10-DS, and D10-DT putative 35S promoter upstream of TATA box (TGACG) as a single copy and not in duplicate as in FMV, CaMV and MMV (Odell et al., 1985; Maiti et al., 1997; Dey and Maiti, 1999) (Fig. 5). This motif is essential for DNA-protein interaction (Lam et al., 1989). The 50 nucleotide-distance separating this motif from TATA box is similar to that found in CaMV promoter (34 nt) and much less than in the CoYMV (176 nt) (Medberry et al., 1992). However, the CAAT box downstream of the TATA box was found in D10-DC, D10-DS, and D10-DT, in addition to CAAT box-like sequences upstream to TATA box (CGACT and

CCACT, and CGAAT) as in FMV (CC/TACT of 49bp distance from TATA box) (Maiti et al., 1997) and CaMV(CCAAT) (Ow et al., 1987) (Fig. 5).

Analysis of the D10-DC, D10-DS, and D10-DT sequences, and comparison with the other available sequences of DMV-D10, demonstrated that single virus properties can differ significantly among members of the same genus. There was little or no sequence variability among clones for a given region of DMV-D10 from each of the wild species. The overall genome organization of D10-DC, D10-DS, and D10-DT is similar to that of D10-US with some differences. The absence of ATF IXG motif in ORFII of both wild and cultivated D10-US indicates that it might be inactive for aphid transmission or it has a different role. D10-US was found to be seed-transmitted and can be detected in pollen (Pahalwatta et al., 2007b) while further investigations showed that it is integrated in plant genome as an endogenous pararetrovirus (Pahalawatta et al., 2008b). The possibility that D10-DC, D10-DS, D10-DT are seed-transmitted or integrated need to be studied. We can not rule out the possibility of presence of a helper virus that could code for CP and ATF and contribute to spread of D10-US in cultivated dahlias (Pahalwatta et al., 2007a) and in wild species. The presence of cis-elements (binding activators) enhancing transcription efficiency of D10-DC, D10-DS, and D10-DT remain to be seen and their role in the promoter activity of theses para-retroviruses in comparison with CaMV 35S promoter. Molecular characterization of D10-DC, D10-DS, and D10-DT is an important step in understanding the evolution of D10-US.

Table 3. Sequence identity matrix (%) of the *Dahlia mosaic virus*-D10 from *Dahlia coccinea* (D10-DC), *Dahlia sherffii* (D10-DS), and *Dahlia tenuicaulis* (D10-DT) ORFs with other caulimoviruses. The figures above the diagonal line represent the nucleotide identity and below the diagonal line are the amino acid identities. Virus acronyms were explained in Table 2.

A: ORFI

aa \ nt	D10-US	D10-NZ	D10-DC	D10-DS	D10-DT	DMV	DCMV
D10-US	**100**	97.1	92.6	91.7	96.1	56.5	56.9
D10-NZ	96.6	**100**	91.2	91.8	94.5	56.4	56.9
D10-DC	89.3	87.2	**100**	91.1	93.5	56.5	56.2
D10-DS	91.2	90.9	88.1	**100**	92.7	56.3	55.9
D10-DT	95.4	93.3	90.9	90.9	**100**	55.9	55.9
DMV	45.6	45	44.4	45.7	45.9	**100**	79
DCMV	46.5	46.8	45.3	46.6	46.8	79	**100**

B: ORF III

aa \ nt	D10-US	D10-NZ	D10-DC	D10-DS	D10-DT	DMV	DCMV
D10-US	**100**	96.1	94.8	95.4	96.1	43.9	43.2
D10-NZ	95	**100**	96.1	95.4	96.1	45.2	44
D10-DC	91.1	96	**100**	95.1	94.8	44.1	43.2
D10-DS	92.1	95	91.1	**100**	95.1	43.6	42.7
D10-DT	96	95	91.1	92.1	**100**	44.1	43.7
DMV	29.6	29.6	28.8	29.6	30.4	**100**	59.7
DCMV	33.3	31.7	30.1	30.1	31.7	42.2	**100**

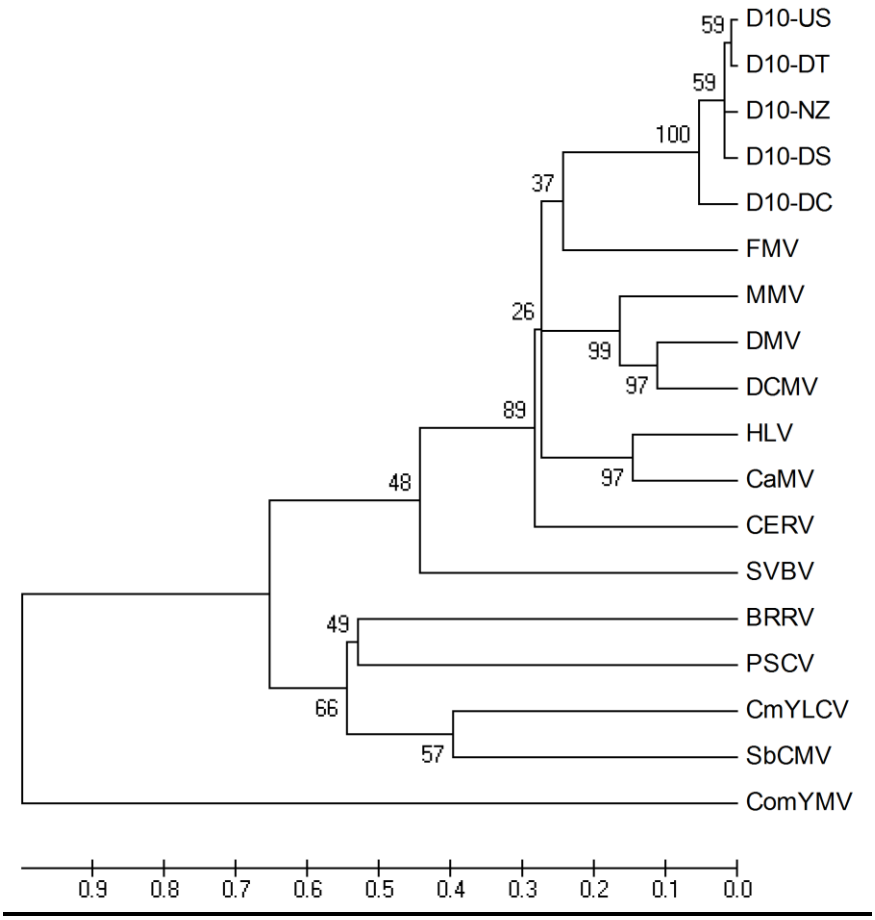
C: ORFV

aa \ nt	D10-US	D10-NZ	D10-DC	D10-DS	D10-DT	DMV	DCMV
D10-US	**100**	88	92	92.3	88.7	59.3	63.7
D10-NZ	89.5	**100**	86.5	87.3	86.1	58.6	63.9
D10-DC	92.9	86.9	**100**	90.5	90	58.9	63.9
D10-DS	93.9	87.6	93.2	**100**	88.9	59.6	64.2
D10-DT	92.3	86.9	92.4	91.7	**100**	59.1	63.7
DMV	57.9	56.3	57.5	58.2	57.2	**100**	67.9
DCMV	66.4	64.5	65.6	65.9	65.3	68.7	**100**

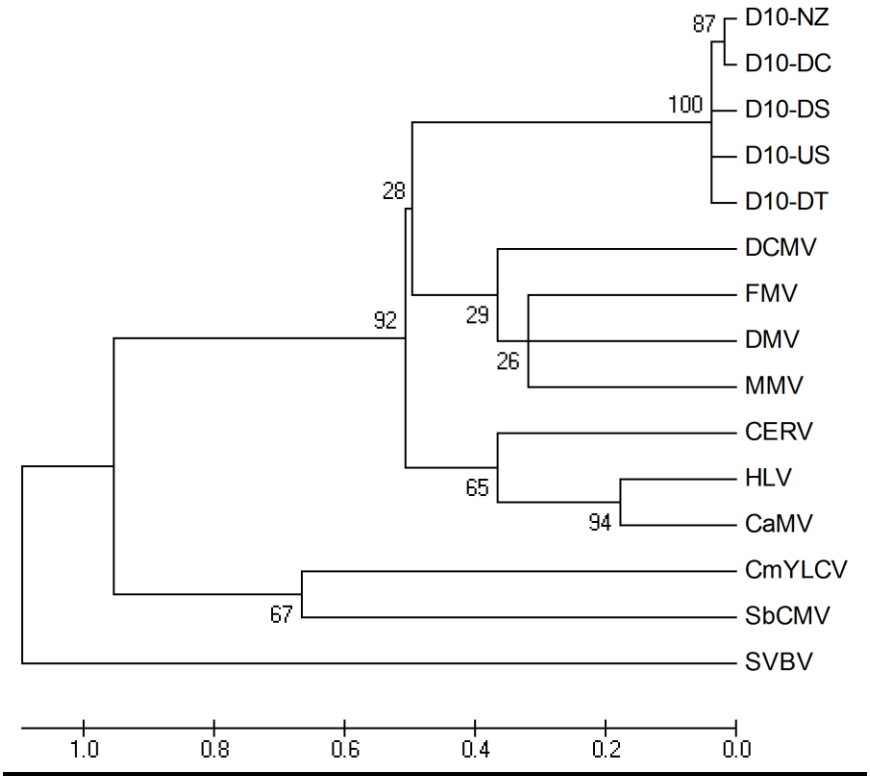
D: ORFVI

aa \ nt	D10-US	D10-NZ	D10-DC	D10-DS	D10-DT	DMV	DCMV
D10-US	**100**	95.8	96.6	93.6	93.1	51.8	50.6
D10-NZ	95.6	**100**	95.2	93.3	91.1	51.4	50.2
D10-DC	96.6	94.8	**100**	94.2	92	51.9	51.1
D10-DS	93.2	92.6	92	**100**	90.5	51.4	50.2
D10-DT	93	90.4	91.6	89.8	**100**	52.6	50.9
DMV	38.6	38.2	38.6	39	38.7	**100**	77.4
DCMV	37	37	37.2	37.2	37.3	75.4	**100**

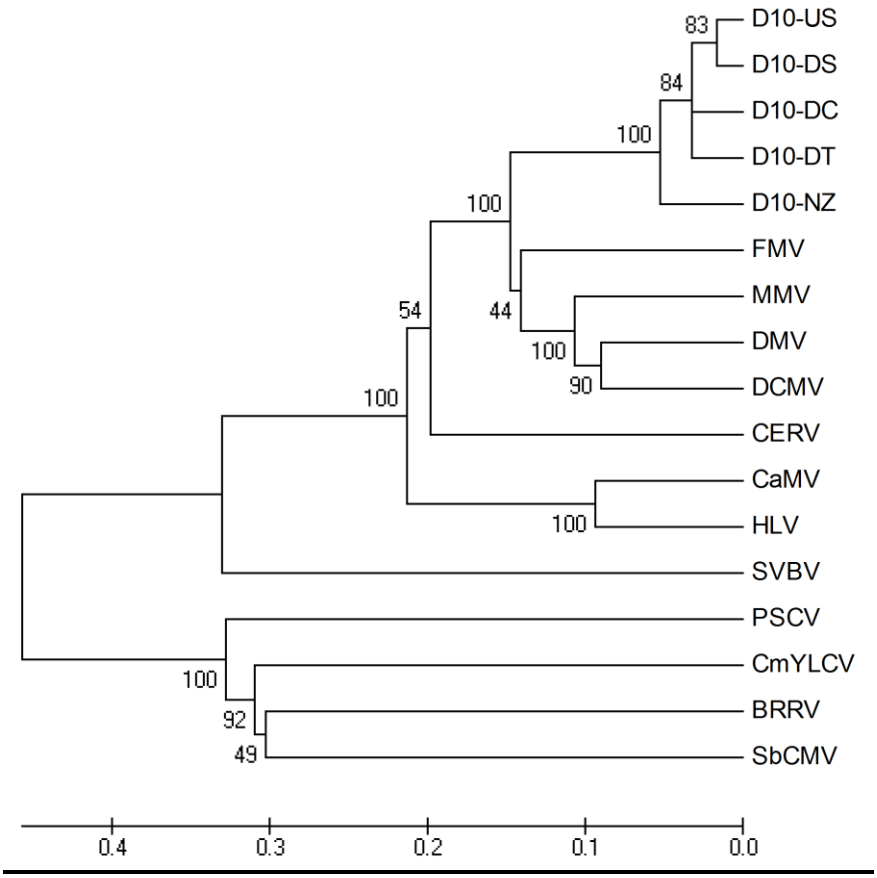
4A: ORFI


D10-US
D10-DT
D10-NZ
D10-DS
D10-DC
FMV
MMV
DMV
DCMV
HLV
CaMV
CERV
SVBV
BRRV
PSCV
CmYLCV
SbCMV
ComYMV
59
59
100
37
26
99
97
89
97
48
49
66
57
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

4B: ORFIII


87
D10-NZ
D10-DC
100
D10-DS
D10-US
28
D10-DT
DCMV
FMV
92
29
DMV
26
MMV
CERV
65
HLV
94
CaMV
CmYLCV
67
SbCMV
SVBV
1.0
0.8
0.6
0.4
0.2
0.0

4C: ORFV


D10-US
D10-DS
D10-DC
D10-DT
D10-NZ
FMV
MMV
DMV
DCMV
CERV
CaMV
HLV
SVBV
PSCV
CmYLCV
BRRV
SbCMV
83
84
100
100
54
44
100
90
100
100
100
92
49
0.4
0.3
0.2
0.1
0.0

4D: ORFVI

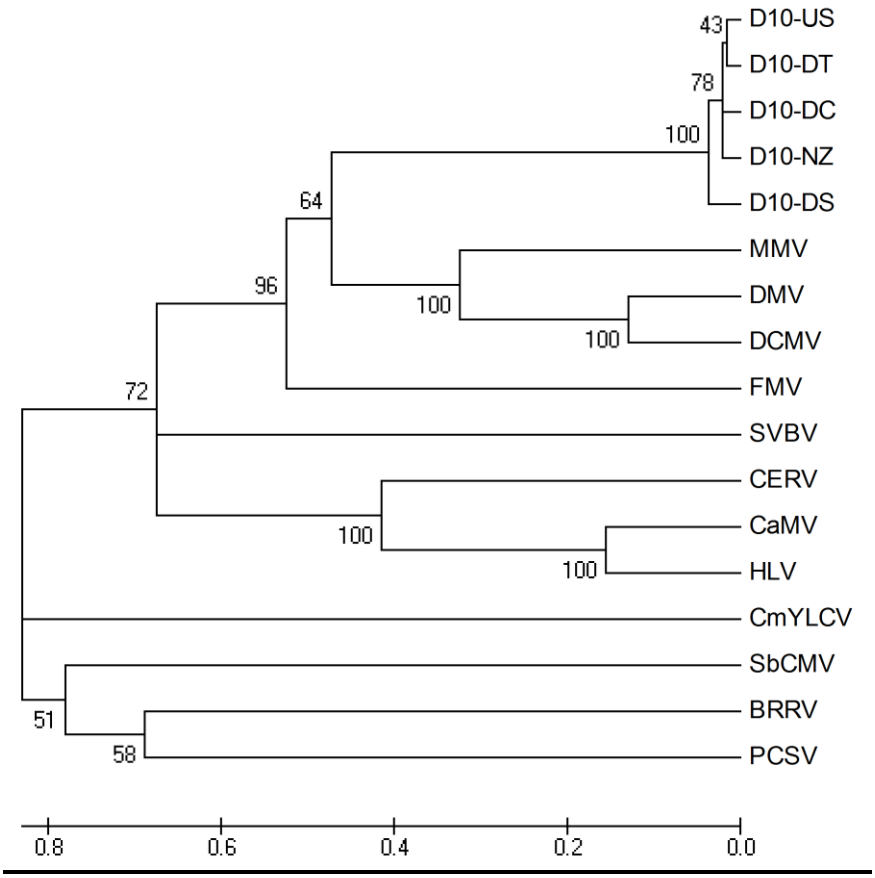


Fig. 4. Phylograms drawn from Clustal W alignments and MEGA4 of the different open reading frames (ORF) of selected members of the family Caulimoviridae compared with those from *Dahlia. coccinea* (D10-DC), *Dahlia. Sherffii* (D10-DS), and *Dahlia. Tenuicaulis* (D10-DT) (refer Table 2). Bootstrap values are indicated at branching points in the phylogram as percentage of 1000 iterations

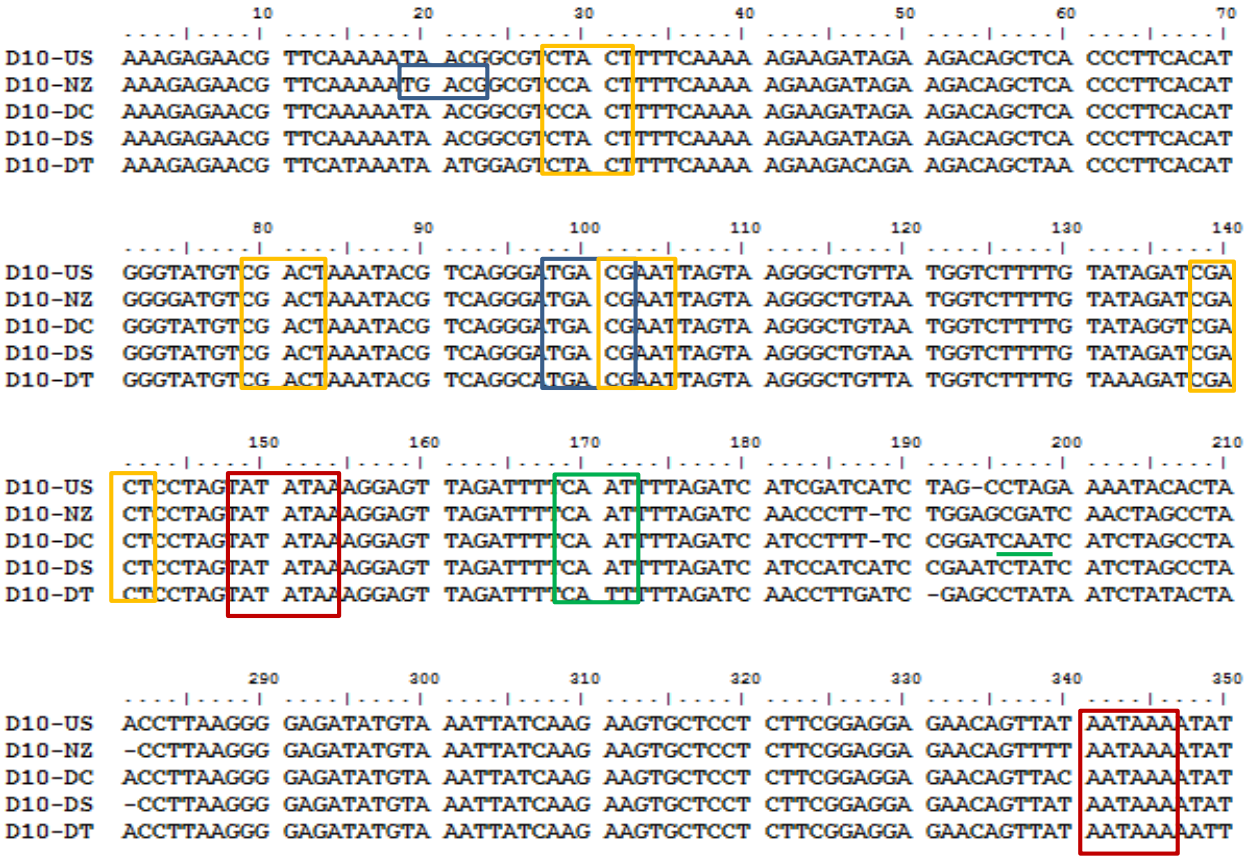


Fig. 5. The DNA sequence of the 35S putative promoter of D10-US (United States), D10-NZ (New Zealand) (*Dahlia variabilis*), D10-DC (*Dahlia coccinea*), D10-DS (*Dahlia sherffii*), and D10-DT (*Dahlia tenuicaulis*) including the TATA box (TATATAA), the poly-adenylation signal (AATAAA), and all repeat sequence domains (TGACG, CC/TACT**,** CGAAT, CAAT) shown in box

REFERENCES

Albouy, J. 1995. Dahlia, p.265-273. In: G. Lobenstein, R.H. Lawson, and A.A. Brunt (eds.). Virus and virus-like diseases of bulb and flower crops. John Wiley and Sons, Chichester, NY.

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

Benfey, P.N. and N.H. Chua. 1990. The cauliflower mosaic virus 35S promoter: Combinatorial regulation of transcription in plants. Science 250: 959-966.

Benfey, P.N., L. Ren, and N.H. Chua. 1989. The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue specific expression patterns. European Molecular Biology Organization Journal 8: 2195-2202.

Benfey, P.N., L. Ren, and N.H. Chua. 1990a. Combinatorial and synergistic properties of CaMV 35S enhancer subdomains. European Molecular Biology Organization Journal 9: 1685-1696.

Benfey, P.N., L. Ren, and N.H. Chua. 1990b. Tissue-specific expression from CaMV 35S enhancer subdomains in early stages of plant development. European Molecular Biology Organization Journal 9: 1677-1684.

Benson, D.A., I. Karsch-Mizrach, D.J. Lipman, J. Ostell, and D.L. Wheeler. 2005. GenBank. Nucleic Acids Research 33: 34-38.

Bonneville, J.M., H. Sanfacon, J. Futterer, and T. Hohn. 1989. Posttranscriptional trans-activation in cauliflower mosaic virus. Cell 6: 1135-1143.

Brunt, A.A. 1971. Some hosts and properties of dahlia mosaic virus. Annals of Applied Biology 67: 357-368.

Canady, J., P. Guillemaut, and J.H. Weil. 1980. The nucleotide sequences of the initiator transfer RNAs from bean cytoplasm and chloroplasts. Nucleic Acids Research 8: 999-1008.

Cavanilles, A.J. 1791. Icones et Descriptiones Plantarum 1**:** 57.

Covey, S.N., D.G. McCallum, D.S. Turner, N.S. Al-Kaff, P. Dale, E. Cecchini, and J.J. Milner. 2000. Pararetrovirus-crucifer interactions: attack and defence or modus Vivendi?. Molecular Plant Pathology 1: 77-86.

Daubert, S., R.J. Shepherd, and R.C. Gardner. 1983. Insertional mutagenesis of the cauliflower mosaic virus genome. Gene 25: 201-208.

Dey, N., and I.B. Maiti. 1999. Structure and promoter/leade deletion analysis of Mirabilis mosaic virus (MMV) full length transcript promoter in transgenic plants. Plant Molecular Biology 40: 771-782.

Eid, S., K.L. Druffel, D. Saar, and H.R. Pappu. 2009. Incidence of multiple and distinct species of caulimoviruses in dahlias (*Dahlia variabilis*). HortScience 44:1498-1500.

Eid ,S., K.L. Druffel, D. Saar, and H.R. Pappu. 2010. Plant Para-retroviral sequences in wild *Dahlia* species in their natural habitats in Mexican mountain ranges. Plant Pathology (in press).

Ghosh, H.P., K. Ghosh, M. Simsek, and U.T. RajBhandary. 1982. Nucleotide sequence of wheat germ cytoplasmic initiator methionine transfer ribonucleic acid. Nucleic Acids Research 10: 3241-3247.

Glasheen, B.M., I.J. Polashock, D.M. Lawrence, J.M. Gillet, D.C. Ramsdell, N. Vorsa, and B.I. Hillman. 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. Archives of Virology 147: 2169-2186.

Hasegawa, A., J. Verver, A. Shimada, M. Saito, R. Goldbach, A. Van Kammen, K. Miki, Kameya-Iwaki, and T. Hibi. 1989. The complete sequence of soybean chlorotic mottle virus DNA and the identification of a novel promoter. Nucleic Acids Research 17: 9993-10013.

Hohn, T., and J. Fütterer. 1997. The proteins and functions of plant pararetroviruses: Knowns and unknowns. Critical Reviews in Plant Sciences 16: 133-161.

Hull, R. 2002. Mathew's Plant Virology. 4th ed. New York, Academic Press, pp 1001.

Hull, R., J. Sadler, and M. Longstaff. 1986. The sequence of carnation etched ring virus DNA: comparison with cauliflower mosaic virus and retroviruses. European Molecular Biology Organization Journal 5: 3083-3090.

Jacquot, E., A. Geldreich, M. Keller, and P. Yot. 1998. Mapping regions of the cauliflower mosaic virus ORFIII product required for infectivity. Virology 242: 395-402

Kobayashi, K., S. Tsuge, H. Nakayashiki, K. Mise, and I. Furusawa. 1998. Requirments of cauliflower mosaic virus open reading frame VI product for viral gene expression and multiplication in turnip protoplasts. Microbiology and Immunology 42: 377-386.

Lam, E., P.N. Benfey, P.M. Gilmartin, R.X. Fang, and N.H. Chua. 1989. Site-specific mutations alter *in vitro* factor binding and change promoter expression pattern in transgenic plants. Proceedings of the National Academy of Science USA 86: 7890-7894.

Leh, V., E. Jacquot, A. Geldreich, H. Thomas, D. Leclerc, M. Cerutti, P. Yot, M. Keller, and S. Blanc. 1999. Aphid transmission of cauliflower mosaic virus requires the viral PIII protein. European Molecular Biology Organization Journal 18: 7077-7085.

Maiti, I.B., S. Gowda, J. Kierman, S.K. Ghosh, and R.J. Shepherd. 1997. Promoter/leader analysis and plant expression vectors with the Figwort mosaic virus (FMV) full – length transcript (FLt) promoter containing single and double enhancer domains. Transgene Research 6: 143-156

Medberry, S.L., B.E.L. Lockhart, and N.E. Olszewski. 1992. The Commelina yellow mottle virus promoter is a strong promoter in vascular and reproductive tissues. The Plant Cell 4: 185-192.

Odell, J.T., F. Nagy, and N.H. Chua. 1985. Identification of DNA sequences required for activity of the *Cauliflower mosaic virus* 35S promoter. Nature 313: 810-812

Ow, D.W., J.D. Jacobs, and S.H. Howell. 1987. Functional regions of the cauliflower mosaic virus 35S promoter determined by the use of the firefly luciferase gene as a reporter of promoter activity. Proceedings of the National Academy of Science USA 84: 4870-4874.

Pahalawatta, V., K.B. Druffel, and H.R. Pappu. 2007a. Incidence and relative prevalence of distinct caulimoviruses (Genus *Caulimovirus*, Family *Caulimoviridae*) associated with dahlia mosaic in *Dahlia variabilis*. Plant Disease 91: 1194-1197.

Pahalawatta, V., K. Druffel, and H.R. Pappu. 2007b. Seed transmission of dahlia mosaic virus in *Dahlia pinnata*. Plant Disease 91: 88-91.

Pahalawatta, V., K.L. Druffel, S.D. Wyatt, K.C. Eastwell, and H.R. Pappu. 2008a. Genome structure and organization of a novel and distinct species of genus *Caulimovirus* (Family *Caulimoviridae*) associated with dahlia mosaic. Archives of Virology 153:733-738.

Pahalawatta, V., K.L. Druffel, and H.R. Pappu. 2008b. A new and distinct species in the genus *Caulimovirus* exists as an endogenous plant pararetroviral sequence in its host, *Dahlia variabilis*. Virology 376: 253-257.

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

Pappu, H. R., K. L. Druffel, R. Miglino, and A. R. Van Schadewijk. 2008. Nucleotide sequence and genome organization of a new and distinct caulimovirus associated with dahlia mosaic. Archives of Virology 153:2145–2148.

Pirone, T. P., and S. Blanc. 1996. Helper-dependent vector transmission of plant viruses. Annual Review Phytopathology 34:227-247.

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

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

Saar, D.E. 1999. A phylogenetic analysis of the genus *Dahlia* (Asteraceae): an interdisciplinary study. Dissertation, Northern Illinois University.

Sambrook, J., E.F. Fritsch, and T. Maniatis.1989. Molecular cloning: A laboratory manual 2nd ed. New York: Cold Spring Harbor Laboratory, 1659 p.

Schoelz, J.E., and R.J. Shepherd. 1988. Host range control of cauliflower mosaic virus. Virology 162: 30-37.

Sørensen, P.D. 1969. Revision of the genus *Dahlia* (Compositae, Heliantheae-Coreopsidinae). Rhodora 71: 309-365, 367-416.

Stavolone, L., A. Ragozzino, and T. Hohn. 2003. Characterization of cestrum yellow leaf curling virus: a new member of the family Caulimoviridae. Journal of General Virology 84: 3459-3464.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596-1599.

Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acid Research 22: 4673-4680.

Torruella, M., K. Gordon, and T. Hohn. 1989. Cauliflower mosaic virus produces an aspartic proteinase to cleave its polyproteins. European Molecular Biology Organization Journal 8: 2819-2825.

Wurch, T., D. Kirchherr, J. Mesnard, and G. Lebeurier. 1990. The cauliflower mosaic virus open reading frame VII product can be expressed in *Saccharomyces cerevisiae* but is not detected in infected plants. Journal of Virology 64: 2594-2598.

CHAPTER SIX

MOLECULAR ANALYSIS OF THE INTEGRATION EVENTS AND EXPRESSION OF AN ENDOGENOUS PARA-RETROVIRUS

ABSTRACT

Para-retroviruses are distinguished from retroviruses by the lack of an integrase domain and integration is not necessary for virus replication. However, pararetroviral sequences were recently discovered in several plants species and designated as endogenous pararetrovirus sequences (EPRS). Three distinct caulimoviruses were found to be associated with dahlia mosaic disease where two are typical caulimoviruses (*Dahlia mosaic virus* -DMV, and Dahlia common mosaic virus-DCMV) and one is an EPRS (DMV-D10). Genomic Southern hybridization showed that DMV-D10 exists as an EPRS. To determine the presence of DMV-D10 transcripts, RT-PCR was done on DNase-treated total RNA from DMV-D10-infected dahlia plants. Results showed expression of open reading frames I, V, and VI regardless of symptom presence on the dahlia samples tested suggesting that DMV-D10 (EPRS) could be pathogenic and is sufficient for disease induction. Therefore, investigating the timeline of integration of DMV-D10 in relation to *Dahlia* speciation and the factors favoring integration and induction of transcription will be important to understand the mechanism of integration of these EPRS in the plant genome and management of the disease.

INTRODUCTION

During the past few years, numerous reports on the detection of pararetrovirus-related sequences (EPRS) in the genomes of several plant species have been published (Harper et al., 1999; 2002; Jakowitsch et al., 1999; Ndowora et al., 1999; Budiman et al., 2000; Lockhart et al., 2000; Mao et al., 2000; Mette et al., 2002; Richert-Pöggeler et al., 2003; Kunii et al., 2004; Geering et. El., 2005a, b; Hansen et al., 2005; Staginnus and Richert-Pöggeler, 2006; Staginnus et al., 2007; Pahalawatta et al., 2008a). All these pararetrovirus-related sequences belong to genera *Petuvirus* (PVCV), *Cavemovirus* (TVCV), *Tungrovirus* (RTBV), *Badnavirus* (BSV) and one case in *Caulimovirus* (DMV-D10) but no integrated sequences of *Soymovirus* have been found in the plant genome yet. Due to this increasing number of EPRS, Staginnus et al. (2009) suggested a uniform nomenclature distinguishing between endogenous sequences that can be affiliated to an exogenous virus and those where no exogenous virus is so far known. The integration mechanism is believed to be through illegitimate recombination where the single-stranded overhanging sequences resulting from reverse transcription might serve as start point for integration as shown in the case of PVCV (Richert-Pöggeler et al., 2003), *Nicotiana sylvestris* EPRV (Jakowitsch et al., 1999), and RTBV (Kunii et al., 2004) or other unknown alternative mode (s) of integration. Few of these EPRS are active and were involved in spontaneous viral infection as in the case of petunia (PVCV), tobacco (TVCV), and banana (BSV) (Richert-Pöggeler and Shepherd, 1997; Richert-Pöggeler et al., 2003; Lockhart et al., 2000; Ndowora et al., 1999; Haper et al., 1999; Dallot et al., 2001). The activation found to occur following stress, wounding from excessive trimming, tissue culture, or interspecific hybridization. Despite the ability to give rise to exogenous viruses and cause disease, no vector

for horizontal transmission has been identified yet for PVCV and TVCV, and survival of these viruses rely on their vertical transmission.

This study focused on DMV-D10 which is the only reported EPRS in genus *Caulimovirus* (Pahalawatta et al., 2008a). DMV-D10 was reported to be seed transmitted, can be detected in pollen, which was expected for an EPRS (Pahalawatta et al., 2007a; 2008a). DMV-D10 was found to be the most prevalent in cultivated dahlias followed by Dahlia common mosaic virus (DCMV; Pappu et al., 2008) and *Dahlia mosaic virus* (DMV) (Pahalawatta et al., 2007b; Eid et al., 2009). The genome structure and organization of DMV-D10 is typical for a caulimovirus with few differences (Pahalawatta et al., 2008b). DMV and DCMV are considered to be exogenous viruses and are not integrated in the plant genome. The aim of this study is to investigate the extent of integration of DMV-D10 in *D. variabilis* and to determine if it is expressed and exists as active EPRS.

MATERIALS AND METHODS

Plant materials. Dahlia plants of various cultivars grown from tubers were brought from the Tacoma Dahlia Trial Garden in the summer of 2008. Plants were grown in the greenhouse (16 hours day light and 22-25°C temperature). Plants were tested for the presence of all three caulimoviruses associated with dahlia mosaic disease (DMV, DCMV, and DMV-D10) using virus-specific primers.

DNA extraction. Genomic DNA was isolated from dahlia leaves following the improved CTAB DNA isolation method by Ghosh et al. (2009) with minor modifications. 100 to 500mg of dahlia leaf tissue was ground to fine powder in liquid nitrogen, added to 10 to15ml of pre-warmed

CTAB extraction buffer (100 mM Tris–HCl pH 8, 10 mM EDTA pH 8, 1.4 M NaCl, 2% CTAB and 0.2% β-marcaptoethanol) and incubated at 65C for 30 min. DNA was precipitated by two-thirds chilled *iso*-propanol after removing the organic contaminants by 0.6 volume of chloroform:iso-amylalcohol (24:1). The DNA pellet was re-suspended in 1M NaCl (0.6ml – 1ml) and treated with RNase A (100mg/ml; Qiagen Inc, Valencia, CA) for 30 min at 37°C. An equal volume of phenol: chloroform was added to remove contaminants and centrifuged for 5min at 9,300xg followed by equal volume of chloroform:iso-amylalcohol (24:1) and centrifugation for 5min at 13,400xg. DNA from supernatant (aqueous phase) was precipitated by adding double volume absolute ethanol and the pellet was washed by 70% ethanol before re-suspension in 100µl autoclaved double-distilled water.

Sap extract and PCR. Sap extracts were prepared as described by Harper et al. (2003) from dahlias infected with DMV, DCMV or DMV-D10 with minor modifications. The infected dahlia leaves were ground in chilled mortar and pestle in 1X phosphates-buffered saline (136mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4 per liter, pH 7.4). A 10mM β mercaptoethanol final concentration was added fresh to the buffer before extraction. The extract was clarified by centrifugation for 5min at 6000x*g* and 1:100 dilution was made with same extraction buffer and kept at 4°C overnight and used the following day in PCR. Each 20 µL PCR reaction contained 1 µL of diluted sap extract, 1X Colorless GoTaq® Flexi Buffer and 2mM $MgCl_2$, 150 µM dNTP mix, and 0.6 pmole each of sense and antisense primers (Table 1), 9.8 µL sterile H_2O, and 1U GoTaq® DNA Polymerase (Promega, Madison, WI). PCR profile used was 94°C for 4min, followed by 32 cycles of 94°C x 0.5min, 48°C x 0.5min, 72°C x 1min, and a final extension at 72°C for 7min.

Table 1. Primers used to detect the three distinct caulimoviruses in dahlia sap extracts

	Primers	Sequence (5'-3')	Expected size(bp)
DMV-ORF IV	F	ACTTCCAAAGAGATCAATAC	1029
	R	TTATTCTGTTCCTGATGATT	
DCMV-ORF IV	F	CATATGGCCACCCAAATGACC	973
	R	TTTTTCCATTCTTCCCATGC	
D10-ORF I	F	ATGGATCGTAAAGATT	420
	R	CCTTCAGAAGGATTTTTACTGC	
D10-ORF VI	F	CCAGGTGCAAGTCCGGAATTA	513
	R	TAGGGCCAACTTCTGTCTCAT	
Dahlia-ITS[a]	ITS5m	GGAAGGAGAAGTCGTAACAAGG	750
	ITS4	TCCTCCGCTTATTGATATGC	

[a] Internal transcribed spacer region of dahlias (Saar et al., 2003)

Genomic Southern hybridization

Probe preparation. Total plant DNA from DMV-D10 infected dahlia plants was used as the template to amplify fragments from ORF I (ORF I start: 5'- ATG GAT CGT AA AGA T T-3'; ORF I end: 5'- CTG TTT TTC TGT GTT TCT ACT GG-3'), ORFV (RtD10-1218F: 5'- TTA ACA GAA CAA GAG TGC GA-3'; D10/KY ORF5R: 5'- TTT TGG AAG ATA AAG TTT TGG A-3'), and ORF VI (ORF VI start: 5'- ATG GAA GAA ATT AAG GCG T -3'; ORF VI end: TTG TCT TCA TCC ATA AAG CAG -3'). The PCR consisted of GoTaq® DNA Polymerase (Promega, Madison, WI) and the corresponding buffer and salt concentration as described previously. The PCR reaction conditions were 94°C x 4min, (94°C x 0.5min,

annealing temp (ORF1-50°C; ORF5-55°C; ORF6-60°C) x 0.33min, 72°C x 1min (ORF1) / x 1.33min (ORF6) / x 2min (ORF5)) x 50 cycles and a final 7 min extension at 72°C. The PCR products were cloned and sequenced. PCR amplification from clones was done with correspondent primers (same conditions) and then gel purified using QIAquick Gel Extraction Kit (Qiagen Inc, Valencia, CA). Labeling was performed with the DIG High Prime DNA Labeling and Detection Starter Kit II (Roche Applied Science, Indianapolis, IN) starting with 1µg PCR product and incubated for 20 hours at 37°C. The reaction was stopped by adding 2µl EDTA 0.2M (pH 8.0). To check the efficiency of labeling, serial dilutions were done as described in the instruction manual according to the expected yield (2300ng / 22µl). 1 µl of each dilution from labeled probe and DIG-labeled DNA (provided by the kit) was applied to Hybond N+ membrane (GE Healthcare BioSciences Corp, Piscataway, NJ) and detected by anti-Digoxigenin-AP conjugate and CSPD ready-to-use substrate by X-ray film exposure (Fig. 1).

Genomic DNA digestion and electrophoresis. Genomic DNA from four dahlia plants positive only to DMV-D10 was extracted using the above described CTAB method. The four dahlia plants were selected from dahlias grown from tubers brought from Tacoma Dahlia Trial Garden, WA, and they were numbered 5, 26, 49 and 63. As a first step, 80-90µg of the genomic DNA (measured by Nanodrop ND-8000) was digested overnight at 37°C with three different restriction enzymes (Invitrogen™, Carlsbad, CA), EcoRI (G/AATTC), Hind III (A/AGCTT), and BamHI (G/GATCC). The digestion of DNA was done in three 1.5ml microfuge tubes per restriction enzyme and in a total volume of 150µl per tube. The digested DNA was pooled from the three tubes and precipitated by adding two-thirds volume chilled *iso*-propanol and one tenth volume sodium acetate (3M, pH 5.3) for one hour at room temperature, then centrifuged at

maximum speed of 16,000xg for 15 min. The DNA pellet was washed in 70% ethanol and re-suspended in 20µl autoclaved water. Five microliters of DNA molecular weight marker II, Digoxigenin-labeled (Roche Applied Science, Indianapolis, IN), and digested DNA (20µl) were loaded on 1% agarose gel (25cm x 15cm) and electrophoresis was performed using 0.5 X TAE buffer (Tris-Acetate-EDTA) at 50V for 2 hrs then 30V for 20-22 hours. The gel was stained with ethidium bromide and photographed.

Gel treatments and transfer to nylon membrane. All steps were done at room temperature on a shaker starting with submerging the gel in 0.25M HCl for 10min (depurination) followed by rinse with ddH_2O and denaturation (0.5 N NaOH, 1.5 M NaCl) for 2x15min. The denaturation was followed by submerging the gel in neutralization buffer (0.5M tris-HCl pH 7.5, 3M NaCl) for 2x15min. The DNA was transformed from the gel onto Hybond N+ membrane (GE Healthcare BioSciences Corp, Piscataway, NJ) in 10X SSC (20XSSC: 3M NaCl, 0.3M Sodium Citrate, pH 7) by upward capillary transfer overnight as described by Sambrook et al. (1989). Fixation of DNA to the membrane was done by UV-crosslinking (Optimal crosslink 120 mJ/cm^2, UV crosslinker, Spectroline) using the wet membrane without prior washing, the membrane then was rinsed briefly (1min) in double distilled water and allowed to air dry.

Prehybridization and hybridization. The membrane was subjected for prehybridization at 42°C for three hours in 25ml DIG Easy Hyb working solution. Hybridization was done overnight at 42°C in 10ml pre-warmed DIG Easy Hyb working solution plus 25ng/ml of corresponding labeled probe (denatured in boiling water). The probe was filtered through a 0.45µm syringe

filter (PALL Life Sciences) performed after addition of the probe to the pre-warmed DIG Easy Hyb working solution. Prehybridization and hybridization were done in glass cylinders (Fisher Scientific) using hybridization oven (Fisher Scientific -Isotemp, Rotisserie motor). Stringency washes (2XSSC +0.1%SDS for 2x5min at room temperature and pre-warmed 0.5XSSC +0.1%SDS for 2x15min at 68°C) were done in the hybridization glass cylinders. Detection (blocking, antibody, and washing) was done using closed plastic trays and according to the manufacturer instructions. X-ray film (BioMax XAR Kodak Film) exposure was done for four hours, six hours and 24 hours to achieve the desired signal strength, after addition of diluted CSPD ready-to-use in detection buffer (1:3) and incubation for 8min at 37°C to enhance the luminescent reaction.

Reverse transcription - PCR. Total RNA extraction was done using RNeasy Plant Mini Kit (Qiagen Inc, Valencia, CA) from 31 dahlia plants found positive only to DMV-D10 by PCR. Dahlia plants were grown from tubers brought from Tacoma Dahlia Trial Garden, WA, in a greenhouse (16 hours day light and 22-25°C temperature) and subjected to regular trimming every three weeks. 14 samples out of 31 showed symptoms (Fig. 2) and the rest of the samples were asymptomatic. The QuantiTect Reverse Transcription Kit (Qiagen Inc, Valencia, CA) was used to generate cDNA from 1µg total RNA treated with gDNA wipe out buffer (provided in the kit) to remove contaminating genomic DNA (absence of DNA was determined in every sample by conducting PCR using the RNA obtained after the treatment with gDNA wipe out buffer as template with virus specific primers). cDNA were obtained by using DMV-D10 specific antisense primers for ORFI, ORFVI (listed in Table 1), and ORFV (EGL 2120-133C: 5'- TCT GTC AAT CCA TTG GG-3') according to manufacturer instructions. The PCR reaction (20 µl)

contained 1X Colorless GoTaq® Flexi Buffer and 2mM $MgCl_2$, 150 µM dNTP mix, and 0.6 pmole each of sense and antisense primers for ORF I, ORF VI (listed in Table 1) and ORF V (D10 1727-602C: 5'-GAA CAG GAA TCT TTG ACT AA-3'; EGL 2120-133C: 5'- TCT GTC AAT CCA TTG GG-3'), 1U GoTaq® DNA Polymerase (Promega, Madison, WI) and the volume was adjusted to 20 µl by adding nuclease free water and 4 µl of cDNA template. PCR parameters were as follows: one cycle at 94°C for 2 min followed by 50 cycles at 94°C for 30s, 50°C (ORF I and ORF V) or 60°C (ORF VI) for 20s, and 72 °C for 32s, , with a final incubation at 72°C for 10 min. PCR products (5 µL) were analyzed by agarose gel electrophoresis (1.2%) in 0.5 X TAE (Tris-acetate-EDTA) buffer. At least one amplicon for each of ORF I, ORF V, and ORF VI product was cloned into pGEM-T (Promega, Madison, WI). Recombinant plasmids were sequenced at the Washington State University sequencing core facility. Sequences were compared with the genomic sequences of DMV-D10 to confirm the identity of the amplicons.

Enzyme-Linked Immunosorbent Assay (ELISA). All 31 dahlia plants were tested for the presence of *Cucumber mosaic virus* (CMV) and *Tomato spotted wilt virus* (TSWV) using ELISA kits from Agdia, Inc (Elkhast, IN) according to manufacturer's instructions.

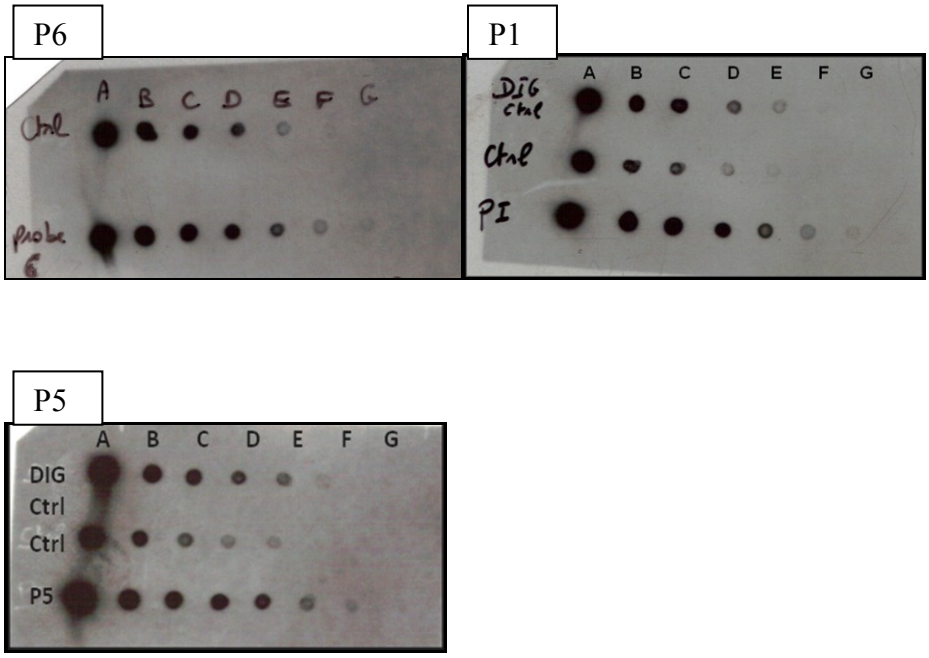


Fig. 1. Determination of labeling efficiency A to H dilutions series of the DIG-labeled control DNA provided by the DIG High Prime DNA Labeling and Detection Starter Kit II. Probes used were P6 (ORF VI), P1 (ORF I), and P5 (ORF V) (A -1ng/µl, B -10pg/ µl, C -3.3pg/ µl, D -1pg/ µl, E -0.33pg/ µl, F -0.1 pg/ µl, G -0.03pg/ µl, H -0.01pg/ µl).

Fig. 2. Symptoms (mosaic and yellowing) on dahlia plants used as a source of RNA for the RT-PCR experiment.

RESULTS AND DISCUSSION

Sap extract PCR gave stronger bands for DMV and DCMV compared to ITS band contrary to DMV-D10 bands which were less intense for both primer pairs. This method is used to distinguish between episomal (DMV and DCMV) and DvEPRS (DMV-D10). DMV and DCMV were used as controls for detection of virus and *Dahlia* ITS region for detection of genomic host DNA (Fig. 3). Incubation at 4°C overnight was essential to dissociate the inclusion bodies and to make the viral sequences accessible to Taq DNA polymerase. A similar approach was used previously for PVCV (Harper et al., 2003) to prove the presence of integrated versus episomal viral sequences using plant sap extract.

To better understand the integration event of DMV-D10 into *D. variabilis* (cultivated dahlias) genome, genomic Southern hybridizations were conducted using probes from ORF I, V, and VI of DMV-D10. EcoRI, HindIII, BamHI size-separated genomic DNA digests were hybridized with probe 1 (ORF I), probe 5 (ORF V), and probe 6 (ORF VI). The relatively high quantity of genomic DNA (80-90µg) used for digestion and Southern hybridization was estimated by preliminary experiments using αdCTP-^{32}P labeled probe (ORFVI). While most Southern hybridizations utilize ^{32}P –labeled probes, digoxigenin (DIG) labeling was used for the detection of infectious *Petunia vein clearing virus* and endogenous provirus in petunia (Richert-Pöggeler et al., 2003). Hybridization gave many bands over a general background smear (Fig. 5 a, b, c). BamHI digest showed one or two strong bands depending on cultivars with a smear at about 23 kb (Fig. 5 a, b, c) contrary to the expected band at about 7 kb if DMV-D10 exists as an episomal virus (Table 2; Fig. 4) irrespective of the probe used. EcoRI digest showed bands with a size range of 2 kb to 9 kb and varied among different cultivars and probes used and contrary to the expected sizes if DMV-D10 were to exist as an episomal virus (Table 2, Fig. 4, 5 a, b, c). Similar observation was made for HindIII digests: detection of strong signals at high molecular weight DNA (much larger than the unit length of the virus gnome) in some cultivars and faint signals in other cultivars, confirming of DMV-D10 integration into dahlia plant genome in a different locations between different cultivars (Table 2, Fig. 4, 5 a, b, c). The variation between cultivars as frequency of integration and intensity of bands is expected as the genetic background of these plants is unknown (inter-specific crosses between unknown parents). The hybridization patterns with the three probes were dissimilar suggesting that these probes were not non-specifically hybridizing with digested DNA. The Southern hybridization of undigested DNA showed a single band at high molecular weight with all the probes used (Fig. 6).

To answer the question whether DMV-D10 is an active integrant, the presence of DMV-D10 transcripts was investigated by RT-PCR. Primers specific to ORF I (movement protein), ORF V (reverse transcriptase), and ORF VI (inclusion body/transactivator) gave amplicons of xpected size suggesting that the 35S and 19S transcripts could be present in the plants tested (Table 3). Transgenic plants expressing CaMV-ORFVI contained inclusion bodies and sometimes expression of symptoms (Covey et al., 2000), and it was suggested that ORF VI product has elicitor-like functions triggering directly or indirectly a host response and leading to symptom expression. The incidence of *Cucumber mosaic virus* (CMV) and *Tomato Spotted wilt virus* (TSWV) was very low in the 31 dahlia plants tested (Table 4). Presence of symptoms on dahlia plants infected only with DMV-D10 and amplification of various ORFs by RT-PCR from symptomatic plants suggests that DMV-D10 is pathogenic and its presence is sufficient for disease induction.

The finding of EPRS in many plant species including *Nicotiana* species (Jakowitsch et al., 1999; Lockhart et al., 2000; Gregor et al., 2004) , potato (Hansen et al., 2005), banana (Harper et al., 1999; Ndowora et al., 1999; Geering et al., 2005a,b), petunia (Richert-Poggeler et al., 2003), rice (Kunii et al., 2004) and tomato (Staginnus et al., 2007) reflects the great genome plasticity in the plant kingdom. However, all previously mentioned EPRS belong to different genera (*Cavemovirus, Petuvirus, Tungrovirus,* and *Badnavirus*) in the family *Caulimoviridae* and DMV-D10 is the only EPRS reported belonging to the genus *Caulimovirus* (Pahalawatta et al., 2008a). At present, PCR is the most reliable method for detecting caulimoviruses associated with dahlia mosaic disease (DMV, DCMV, and DMV-D10) and presence of symptoms can not be used as criterion for virus diagnosis (Pahalawatta et al., 2007b; Eid et al., 2009) since asymptomatic/latent infections are common. It takes a long time to hybridize dahlia flowers and

get them accepted by the American Dahlia Society. Therefore, it is desirable before going further in vegetative propagation of dahlia materials or developing virus management tactics to distinguish early whether viral or host plant sequences are being detected.

Southern hybridization of total genomic DNA and DNA libraries approaches led to detection EPRS sequences in potato (Hansen et al., 2005), tomato (Staginnus et al., 2007), and *Musa* species (Geering et al., 2005 a,b). The future steps will be to investigate the timeline of integration of DMV-D10 in relation to *Dahlia* speciation. No vector for DMV-D10 horizontal transmission has yet been identified as in the case of PVCV and TVCV (EPRS) whereas survival is guaranteed by vertical transmission (Lockhart et al., 2000; Richert-Pöggeler and Shepherd 1997). The aphid transmission factor is lacking in PVCV as well as in DMV-D10 (episomal and EPRS forms) (Richert-Pöggeler and Shepherd 1997). It was reported that some EPRS could give rise to episomal particles and cause symptoms under certain conditions. Spontaneous viral infections were noticed in petunia, tobacco, and banana by PVCV (Richert-Pöggeler and Shepherd 1997; Richert-Pöggeler et al., 2003), TVCV (Lockhart et al., 2000), and BSV (Ndowora et al., 1999; Harper et al., 1999; Dallot et al., 2001) respectively.

The finding of this study included strong instance of a first caulimovirus existing as EPRS and this implies serious consideration for dahlia trade because it is difficult if not impossible to eliminate the EPRS from propagative stock. The factors favoring integration of DMV-D10 into *Dahlia* genome and potential hot spots for integration in the plant genome are unknown and remain to be explored. In addition, the finding that DMV-D10 is an active integrant will necessitate studying the factors involved in the expression of EPRS resulting in pathogenicity. Information on the dynamics of transcription from DMV-D10 could provide

insights into the infection process of this EPRS and the extent of contribution to virus evolution and pathogenicity.

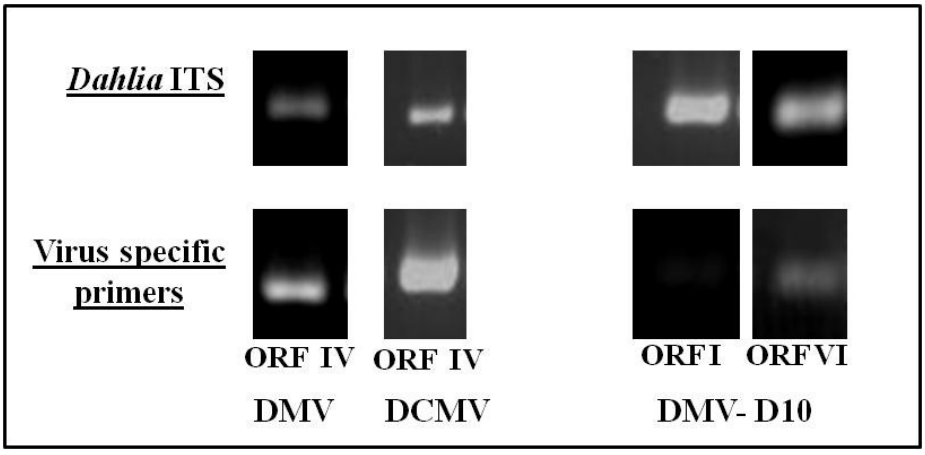


Fig. 3. Polymerase chain reaction from sap extract for *Dahlia mosaic virus* (DMV), Dahlia common mosaic virus (DCMV) and Dahlia mosaic virus- D10 (DMV-D10) using virus-specific primers, and primer specific to the internal transcribed spacer region of *Dahlia* (ITS).

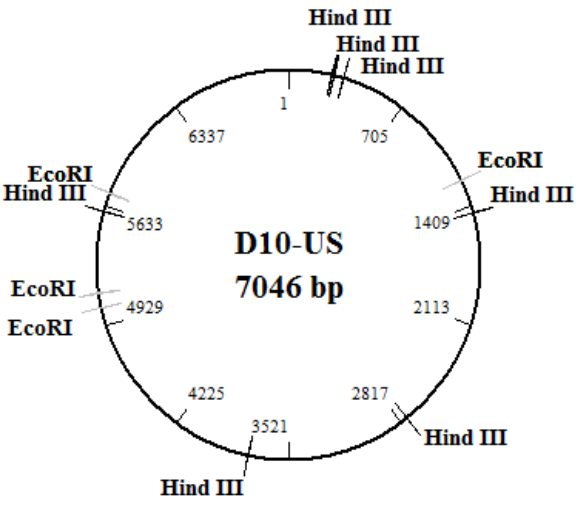
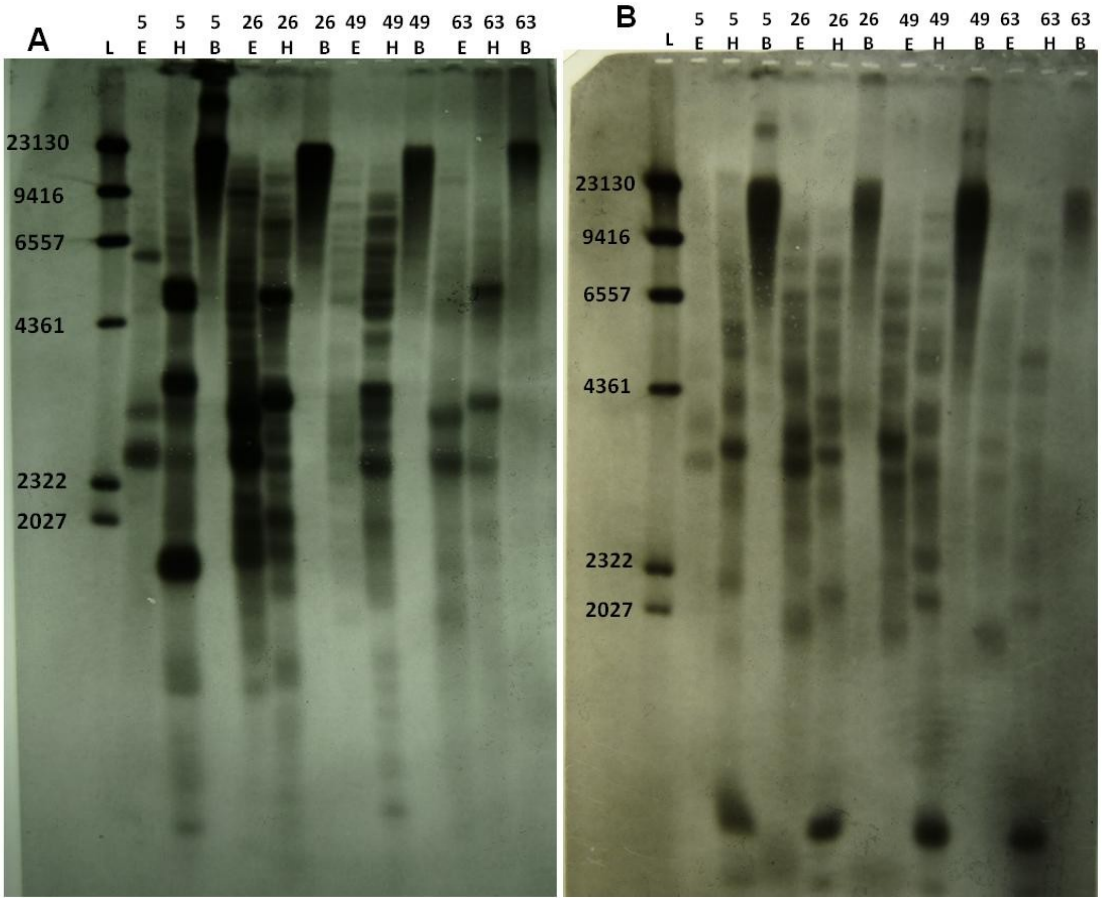


Fig. 4. Restriction enzyme map of Dahlia mosaic virus-D10 genome (referred as D10-US) showing EcoRI and HindIII digestion sites and positions.

Table 2. Restriction digestion of Dahlia mosaic virus-D10 with EcoRI, HindIII, and BamHI with expected sizes (bp) if DMV-D10 were to exist as an episomal virus.

Bands of expected sizes (bp) with EcoRI, HindIII, and BamHI restriction enzymes			
	EcoRI	HindIII	BamHI
If DMV-D10 episomal	3783 - 2590 - 589 – 84	1834 - 1700 - 1325 - 1125 - 992 - 58 – 12	7046
Probe 1 (ORF1)	2590	1700 – 1125 – 58 – 12	7046
Probe 5 (ORF5)	3783	1834 – 1325 – 992	7046
Probe 6 (ORF6)	2590 – 589 – 84	1834 – 1700	7046



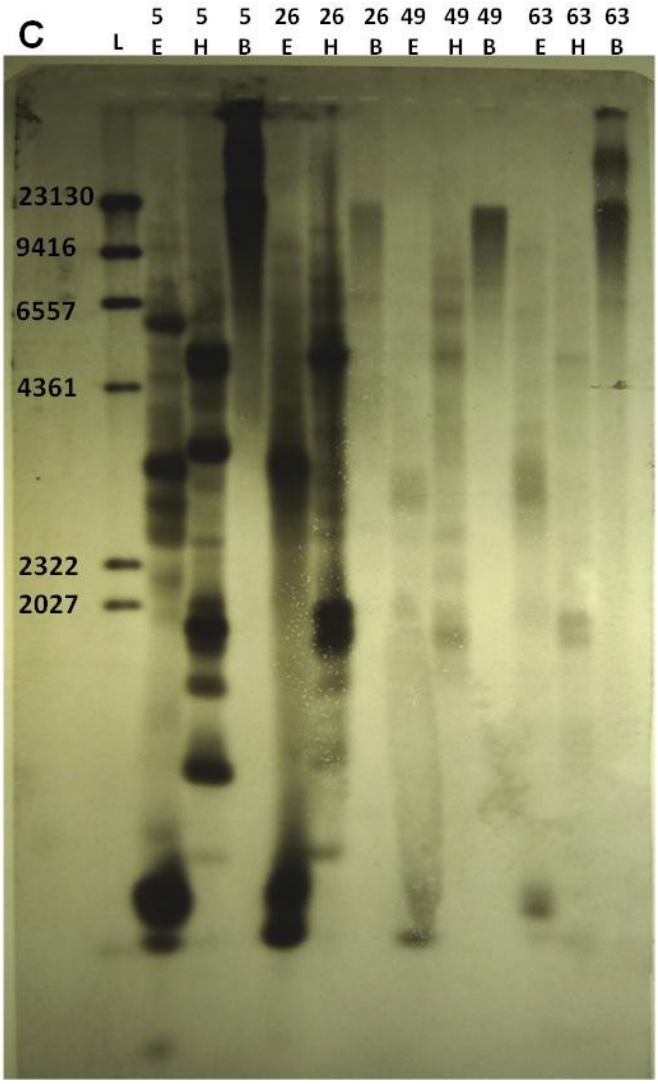


Fig. 5. Southern hybridization of genomic DNA from DMV-D10-infected dahlia plants digested with EcoRI (E), HindIII (H), and BamHI (B). 80-90µg of DNA was loaded for samples from different *D. variabilis* cultivars (5, 26, 49 and 63) using probes from ORF I (A), ORF V (B), ORF VI (C). Marker DNA (L- bp) consists of HindIII digest of digoxigenin-labeled λ DNA (Roche).

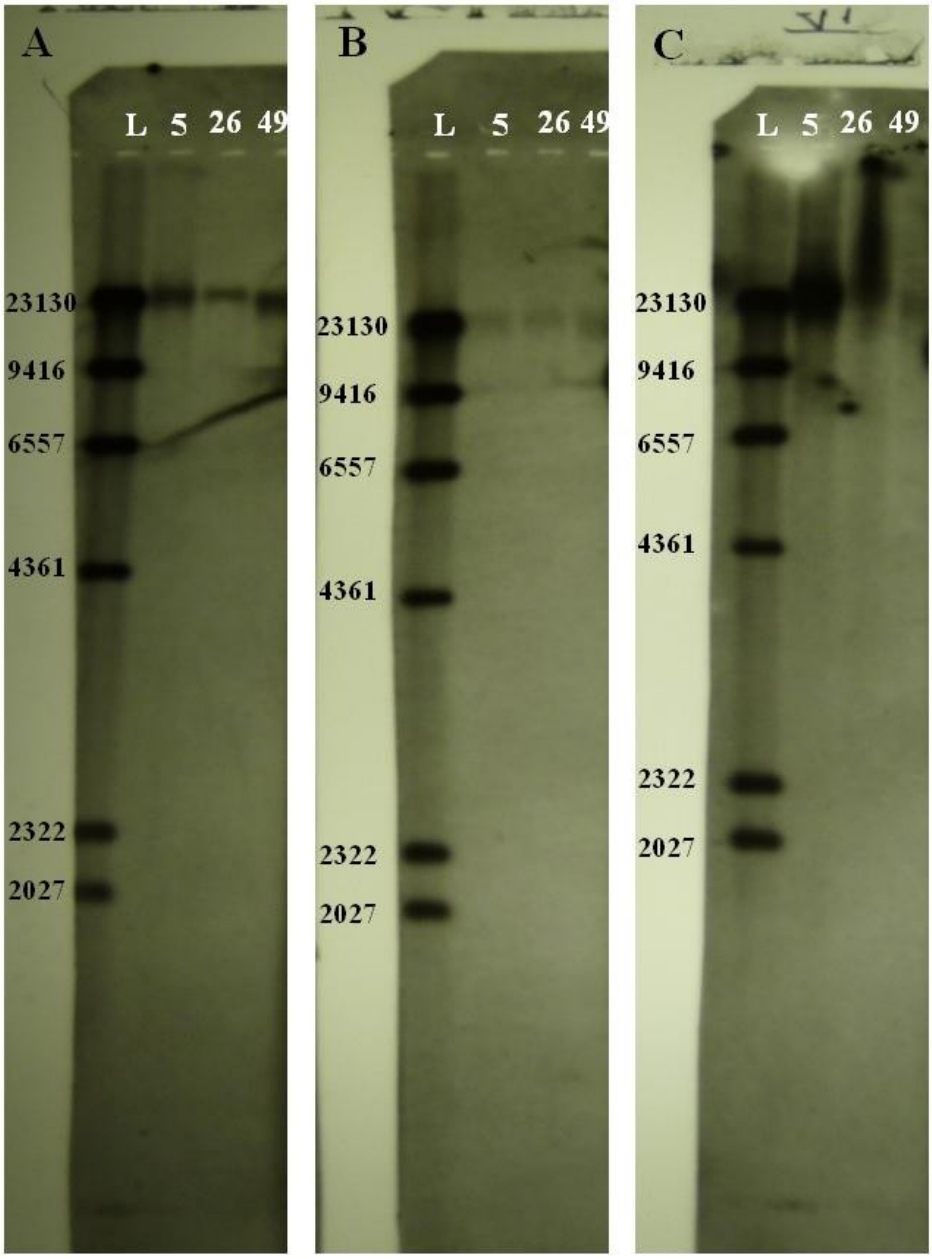


Fig. 6. Southern hybridization of undigested genomic DNA from DMV-D10-infected dahlia plants. 8-10µg of DNA was loaded for samples from different *D. variabilis* cultivars (5, 26, 49) using probes from ORF I (A), ORF V (B), ORF VI (C). Marker DNA (L- bp) consists of HindIII digest of digoxigenin-labeled λ DNA (Roche).

Table 3. Amplification of various DMV-D10 genes in RT-PCR using total RNA from DMV-D10-infected dahlia plants.

DMV-D10 Symptomatic samples			DMV-D10 Asymptomatic samples		
ORF I	ORF V	ORF VI	ORF I	ORF V	ORF VI
14/14*	9/14	11/14	13/17	11/17	13/17

*RT-PCR positive/total plants tested

Table 4. Detection of *Tomato spotted wilt virus* and *Cucumber mosaic virus* in symptomatic and asymptomatic dahlia plants

Tomato spotted wilt virus (TSWV)		*Cucumber mosaic virus* (CMV)	
Symptomatic	Asymptomatic	Symptomatic	Asymptomatic
5/14*	2/17	2/14	3/17

*ELISA positive/total plants tested

RFEERENCES

Budiman, M.A., L. Mao, T.C. Wood and R.A. Wing. 2000. A deep-coverage tomato BAC library and prospects toward development of an STC framework for genome sequencing. Genome Research 10: 129-136.

Covey, S.N., D.G. McCallum, D.S. Turner, N.S. Al-Kaff, P. Dale, E. Cecchini and J.J. Milner. 2000. Pararetrovirus-crucifer interactions: attack and defence or modus Vivendi?. Molecular Plant Pathology 1: 77-86.

Dallot, S., P. Acuna, C. Rivera, P. Ramirez, F. Cote, B.E.L. Lockhart and M.L. Caruana. 2001. Evidence that the proliferation stage of micropropagation procedure is determinant in the experience of *Bana streak virus* integrated into the genome of the FHIA 21 hybrid (*Musa* AAAB). Archives of Virology 146: 2179-2190.

Eid ,S., K.L. Druffel, D. Saar, and H.R. Pappu. 2009. Incidence of multiple and distinct species of caulimoviruses in dahlias (*D. variabilis*). HortScience 44:1498-1500.

Geering, A.D., M.M. Pooggin, N.E. Olszewski, B.E. Lockhart and J.E. Thomas. 2005a. Characterisation of Banana streak Mysore virus and evidence that its DNA is integrated in the B genome of cultivated Musa. Archives of Virology 150:787–796.

Geering, A.D., N.E. Olszewski, G. Harper, B.E. Lockhart, R. Hull and J.E. Thomas. 2005b. Banana contains a diverse array of endogenous badnaviruses. Journal of General Virology 86:511–520.

Ghosh, R., S. Paul, S.K. Ghosh, and A. Roy. 2009. An improved method of DNA isolation suitable for PCR-based detection of begomoviruses from jute and other mucilaginous plants. Journal of Virological Methods 159: 34-39.

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

Hansen, C.N, G. Harper and J.S. Heslop-Harrison. 2005. Characterization of pararetrovirus-like sequences in the genome of potato (*Solanum tuberosum*). Cytogenetic and Genome Research 110: 559-565.

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

Harper, G., R. Hull, B. Lockhart and N. Olszewski. 2002. Viral sequences integrated into plant genomes. Annual Review of Phytopathology 40: 119-136.

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

Jakowitsch, J., M.F. Mette, J. van der Winden, M.A. Matzke and A.J.M. Matzke. 1999. Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. Proceedings of the National Academy of Science USA 96:13241-13246.

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

Lockhart, B. E., J. Menke, G. Dahal and N.E. Olszewski. 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. Journal of General Virology 81:1579-1585.

Mao, L., T.C. Wood, Y. Yu, M.A. Budiman, J. Tomkins, S. Woo, M. Sasinowski, G. Presting, D. Frisch, S. Goff, R.A. Dean and R.A Wing. 2000. Rice transposable elements: a survey of 73000 sequence-tagged-connectors. Genome Research 10: 982-990.

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

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

Pahalawatta, V., K. Druffel and H.R. Pappu. 2008a. A new and distinct species in the genus *Caulimovirus* exists as an endogenous plant pararetroviral sequence in its host, *Dahlia variabilis*. Virology 376: 253-257.

Pahalawatta, V., K. L. Druffel, S. D. Wyatt, K. C. Eastwell, and H.R. Pappu. 2008b. Genome structure and organization of a novel and distinct species of genus *Caulimovirus* (Family *Caulimoviridae*) associated with dahlia mosaic. Archives of Virology 153:733-738.

Pahalawatta, V., K. Druffel and H.R. Pappu. 2007a. Seed transmission of *Dahlia mosaic virus* in *Dahlia pinnata*. Plant Disease 91: 88-91.

Pahalawatta, V., R. Miglino, K. L. Druffel, A. Jodlowska, A. R. Van Schadewijk, and H. R. Pappu. 2007b. Incidence and relative prevalence of distinct caulimoviruses (Genus *Caulimovirus*, Family *Caulimoviridae*) associated with dahlia mosaic in *Dahlia variabilis.* Plant Disease 91:1194-1197.

Pappu, H. R., K. L. Druffel, R. Miglino, and A. R. Van Schadewijk. 2008. Nucleotide sequence and genome organization of a new and distinct caulimovirus associated with dahlia mosaic. Archives of Virology 153:2145–2148.

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

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

Saar, D.E., N.O. Polans and P.D. Sørensen. 2003. A phylogenetic analysis of the genus Dahlia (Asteraceae) based on internal and external transcribed spacer regions of nuclear ribosomal DNA. Systematic Botany 28:627–639.

Sambrook, J., et al. (1989). *Molecular Cloning: A laboratory manual. 2nd ed*, Cold Spring Harbor Laboratory Press, 1659 p.

Staginnus, C., M.L. Iskara-Caruana, B. Lockhart, T. Hohn and K.R. Richert-Pöggeler. 2009. Suggestions for a nomenclature of endogenous pararetroviral sequences in plants. Archives of Virology 154: 1189-1193.

Staginnus, C., W. Gregor, M.F. Mette, C.H. Teo, E.G. Borroto-Fernandez, M. Laimer da Camara Machado, M.A. Marzke and T. Schwarzacher. 2007. Endogenous pararetroviral sequences in tomato (*Solanum lycopersicum*) and related species. BioMed Central Plant Biology 7:24.

Staginnus, C. and K.R. Richert-Pöggeler. 2006. Endogenous pararetroviruses: two-faced travelers in the plant genome. Trends in Plant Science 11: 485-491.

CHAPTER SEVEN

GENERAL CONCLUSIONS

Dahlia production is facing many challenges ranging from fungal diseases to spider mite infestation, viral diseases including several RNA DNA viruses. *Dahlia mosaic virus* (DMV) causing the dahlia mosaic disease is a serious disease affecting dahlias. In addition to DMV, two putative new caulimoviruses, tentatively designated as DMV-D10 (existing as endogenous plant pararetroviral sequence - DvEPRS) and Dahlia common mosaic virus (DCMV) were found to be associated with dahlia mosaic disease. Forty years have passed since the last and only report on biological properties of DMV by Alan Brunt (1970). *Verbesina encelioides* was found to be the best plant for virus source (DMV) and *Myzus persicae* (Sulz.) to be the principal vector for aphid transmission of DMV (Brunt, 1970).

The diversity of caulimoviruses extant in dahlia directed the necessity of conducting biological studies on DCMV and DMV-D10 in comparison with DMV. *V. encelioides* was found to be a good experimental host for DCMV and DMV for both aphid and mechanical transmission. DMV and DCMV were aphid transmitted at the rate of 50 and 73%, respectively. The most effective buffers for the highest rate of mechanical transmission of DMV and DCMV were phosphate buffers with nicotinic acid (1%) or sodium bisulfite (5%) which gave transmission rates of 58% and 36%, respectively. DMV-D10 was not transmissible by either aphids or mechanically. The reasons for loss of DMV-D10 transmission by mechanical transmission are unknown and it could be that its integration in its host genome has exerted a selection for virus structure making it less compatible with aphid and mechanical transmission. Availability of an

experimental host (*V. encelioides*) for DMV and DCMV could facilitate further studies on these viruses.

The incidence of the three caulimoviruses in dahlias showed the prevalence of DMV-D10 (94%) followed by DCMV (48.5%) and DMV (23.2%). Mixed infections were common and viruses were detected irrespective of symptom expression at the time of sampling. The high incidence of DMV-D10 could be due to its integration into the host genome contrary to DMV and DCMV which seem to exist as episomal molecules typical like other known species in the genus *Caulimovirus*. The results of this survey on cultivated dahlias highlight the importance of screening dahlia propagative material for these three viruses before going further in vegetative propagation. Presence of symptoms should not be the only criterion for virus diagnosis.

The incidence of these viruses was investigated in four wild *Dahlia* species collected from the Sierra Madre Occidental region of western Mexico. From total of 56 samples were tested and DMV-D10 was found in 51 samples and no evidence of the presence of DMV and DCMV. The Presence of DMV-D10 in wild species in their natural habitats in the center of diversity suggests that DMV-D10 might have been vertically transmitted to modern-day cultivated species. This was based on the assumption that DMV-D10 exists as an EPRS in wild species as well. However, this remains to be confirmed. Further studies on correlating the incidence of viral sequences with the species and population distribution of wild dahlias would provide new avenues of research into the evolutionary pathways of plant-associated pararetroviruses.

As a first step to understanding the evolution of this EPRS in relation to its host, molecular characterization of DMV-D10 in wild *Dahlia* spp. (*D. coccinea, D. sherffii, D.*

tenuicaulis) were carried out. The unique differences (absence of ATF, truncated CP gene fused with the RT gene) were conserved among the DMV-D10 isolates from wild *Dahlia* species. Amino acids sequence divergence at level of ORFI, III, and VI ranged between 89.3 to 96.6 % while phylograms of DMV-D10 from wild species (D10-DC, D10-DS, D10-DT) compared with sequences of other selected members of the family *Caulimoviridae* grouped them together in one cluster and in one big clade including *Figwort mosaic virus, Miriabilis mosaic virus, Dahlia mosaic virus* and Dahlia common mosaic virus. The 35S putative promoter sequence showed little divergence at the level of enhancer elements. It remains to be seen if DMV-D10 exists as an EPRS and whether the sequence divergence among the DMV-D10 in wild species is congruent with the evolution of *Dahlia* spp. *Dahlia* and DMV-D10 offer a model system to study the co-evolution of plant species and endogenous pararetroviruses.

The integration extent and frequency of DMV-D10 in cultivated dahlias was investigated by genomic Southern hybridization using various genes as probes (ORFI, V, and VI). Results showed that DMV-D10 exists as an EPRS and the extent of integration varies from cultivar to cultivar. The finding of EPRS in dahlia has implications for trade since it is difficult to eliminate the EPRS to produce EPRS-free stock. It is desirable to know the 'hot spots' in the plant genome favoring integration. It was reported for some endogenous para-retroviral sequences (i.e. BSV and PVCV) that they exist as functional integrants and can give rise to episomal particles and cause symptoms under certain conditions. Such information would be useful to produce DMV-D10-free stock.

The Genome expression studies of DMV-D10 by RT-PCR showed that the 35S and 19S transcripts could be present in DMV-D10-infected plants. Detection of transcripts by RT-PCR suggests that DMV-D10 as an EPRS is pathogenic and is sufficient for disease induction.

Studying the dynamics of transcription from DMV-D10 could provide insights into the infection process of this EPRS.