The Pennsylvania State University

The Graduate School

Department of Plant Pathology and Environmental Microbiology

PERSISTENT VIRUSES IN PEPPERS:

THEIR EVOLUTION AND BENEFICIAL EFFECTS

A Dissertation in

Plant Pathology

by

Maliheh Safari

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

December 2017

The dissertation of Maliheh Safari was reviewed and approved* by the following:

Marilyn J. Roossinck
Professor of Plant Pathology
Dissertation Advisor
Chair of Committee

Cristina Rosa
Assistant Professor of Plant Pathology

David M. Geiser
Professor of Plant Pathology

Andrew G. Stephenson
Distinguished Professor of Biology, Associate Dean for Research

Craig E. Cameron
Professor of Biochemistry and Molecular Biology

Carolee Bull
Head of the Department of Plant Pathology and Environmental Microbiology

*Signatures are on file in the Graduate School

ABSTRACT

There are many non-pathogenic viruses that are maintained in a persistent lifestyle in plants. Plant persistent viruses are widespread, replicating in their hosts for many generations. The roles of plant persistent viruses have not been studied thoroughly, but their very long-term relationships with their hosts, and their high level of vertical transmission imply beneficial interactions. So far, *Endornaviridae* is the only family with a single-stranded RNA genome, containing one large open reading frame. *Bell pepper endornavirus* (BPEV), *Hot pepper endornavirus, Capsicum frutescens* endornavirus 1 (CFEV 1) have been identified from peppers. Peppers are native to Central and South America, and as domesticated plants human selection accelerated their evolution. Using a collection of 97 different peppers the evolution of pepper persistent viruses was studied. The evolution of endornaviruses in different peppers was investigated using two fragments from the viral helicase (Hel) and RNA dependent RNA polymerase (RdRp) domains. In addition, by using single nucleotide polymorphisms the pepper host populations and phylogenies were analyzed. The endornaviruses phylogeny was correlated with its *Capsicum* species host. In this study BPEV was limited to *C. annuum* species, and the RdRp and Hel phylogenies identified two clades that correlated with the host pungency. No *C. annuum* infected with CFEV 1 was found in this study, but the CFEV 1 RdRp fragment was recovered from *C.chinense*, *C. frutescens*, and *C.bacccutum* and *C. pubescens*.

Partitiviridae is the most common persistent virus family in wild plants. Jalapeño and Hungarian Wax peppers (*Capsicum annuum*) have been reported with *Pepper cryptic*

virus 1 (PCV 1) and *Pepper cryptic virus* 2 (PCV 2), respectively. Both viruses belong to the genus *Deltapartitivirus* from *Partitiviridae* family. The evolution of PCV 1 and PCV 2 was investigated using the RdRp and coat protein of both viruses. Both viruses were detected in cultivated and wild peppers. It is shown that these viruses have a remarkably slow evolution rate in comparison with acute RNA viruses. This might be correlated with the strong purifying selection related to the lifestyle of theses viruses, or the replication strategies in double stranded RNA viruses, that use a stamping machine mode of replication.

Finally, some potential beneficial effects of persistent virus were examined. In wild plants partitivirus infection decreased the likelihood of acute virus infection. I compared the effect of odor cues from PCV 1 infected (J+) and virus free (J-) Jalapeño pepper on the aphid *Myzus periscae*, a common vector of acute plant viruses. Pairwise preference experiments showed a stark contrast to insect-plant interactions in acute virus infections: virus infected plants deterred aphids. The acute plant virus *Cucumber mosaic virus* (CMV) manipulates its host's volatile emission to attract aphid vectors and facilitate its transmission. Volatiles of J+ and J- CMV infected plants were more attractive to aphids than J+ and J- mock inoculated plants. However, in pairwise preference between J+ CMV- and J- CMV-infected plants, aphids preferred the J- CMV volatile blend. Also, aphid fecundity on J+ and J- plants was measured as an indicator for the effect of PCV 1 on host quality for aphids. Aphid reproduction on J+ plants was more than two fold lower than J- plants. This study demonstrates a beneficial relationship between PCV 1 and Jalapeño plants by protecting the plants from the vector of acute viruses. In addition, the

effect of PCV 1 on the Jalapeño's developmental growth was tested by measuring the average time required for seed germination, emergence of first true leaf, first open flower and dried biomass of plants. PCV 1 showed no significant effect on the developmental growth of Jalapeño plants.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere appreciation to my advisor, Dr. Marilyn J. Roossinck, for her valuable guidance, great encouragement, unlimited support and providing me with an excellent atmosphere for performing research. Not only her mentorship has helped me to grow as a research scientist, but also her advice on my life has been priceless. I hope that I can apply what I have learned from her mentorship to help and inspire others in the future.

I would also like to thank my doctoral committee: Dr. Cristina Rosa, Dr. David M. Geiser, Dr. Andrew G. Stephenson, and Dr. Craig E. Cameron for their comments and suggestions throughout my project, and for generously spending their time helping me to improve this dissertation. Special thank to Dr. Rosa and Dr. Stephenson for providing some of aphid equipment.

In addition, I want to express my sincere gratitude to Dr. Juan Jose Lopez-Moya for all his invaluable help and support during his short visit in our lab. Thanks to his incredible willingness to help. I am grateful for his advice and guidance. I acknowledge Dr. Matthew J. Ferrari's guidance for statistical analysis, Dr. Rodrigo Valverde for supplying Jalapeño seeds and CFEV 1 sequence, and Dr. Gary Thompson for providing aphid colony. I am thankful to generous people who helped in pepper collection; Dr. Michael D. Orzolek, Dr. Rafael Bello-Bedoy and particularly Dr. Fernando García-Arenal who also sent me dried chiltepin samples. Many thanks to Dr. Robert Jarret for his guidance in pepper information, Dr. Hamid Ashrafi and Dr. Theresa Hill for their discussions about pepper markers and analysis.

I am also grateful to all of the staff and current and former students from the plant pathology department, and center of infectious disease dynamics for their friendship, caring and assistance during last five years.

Many thanks to all of my friends or better called my family at Roossinck lab: Dr. Mahtab Peyambari, Dr. Vaskar Thapa, Rimnoma Serge Ouedraogo, and Edelio Bazan who have always been there through all the ups and downs, happiness and sadness, and helping me in the lab. I would like to thank Dr. Luis M. Marquez, Dr. Justin Pita, and Dr. Xiaodong Bao who trained me in the Roossinck lab.

I wish to thank my friends and family for their unconditional love and support through all the stages of my personal and academic life. My dearest dad and mom, Hamid and Fatemeh, who believed in me and supported me all the way through, no matter how far I have been from them. My lovely sister, Mahsa, that her presence at Penn state has been heartwarming. And last but not least, I must express my profound gratitude to Amin, my husband who has been standing beside me throughout my career and writing this dissertation.

The work presented herein was funded by the Pennsylvania State University College of Agricultural Sciences, and the Huck Institute for Life Sciences, thank you vey much for your support.

Chapter 1

Evolution of Endornaviruses in *Capsicum* spp.

Abstract

There are many non-pathogenic viruses that are maintained in a persistent lifestyle in plants. Plant persistent viruses are widespread, replicating in their hosts for many generations. So far, *Endornaviridae* is the only family of plant persistent viruses with a single-stranded RNA genome, containing one large open reading frame. *Bell pepper endornavirus* (BPEV), *Hot pepper endornavirus, Capsicum frutescens* endornavirus 1 (CFEV 1) have been identified from peppers. Peppers are native to Central and South America, and as domesticated plants human selection accelerated their evolution. We investigated the evolution of these endornaviruses in different peppers including *Capsicum annuum*, *C. chacoense*, *C.chinense*, *C. frutescens*, *C.bacccutum* and *C. pubescens* using two fragments from the viral helicase (Hel) and RNA dependent RNA polymerase (RdRp) domains. In addition, by using single nucleotide polymorphisms we analyzed the pepper host populations and phylogenies. The endornaviruses phylogeny was correlated with its *Capsicum* species host. In this study BPEV was limited to *C. annuum* species, and the RdRp and Hel phylogenies identified two clades that correlated with the host pungency. No *C. annuum* infected with CFEV 1 was found in this study, but the CFEV 1 RdRp fragment was recovered from *C.chinense*, *C. frutescens*, and *C.bacccutum* and *C. pubescens*.

Introduction

There are many cytoplasmic RNA viruses reported from plants that persist for many generations through nearly 100% vertical transmission. These viruses are very poorly studied, even though they are found in many crops and are the most common viruses found in wild plants (Roossinck 2012, Roossinck 2014). Generally persistent plant viruses are reported with no effect on their host, although finding an isogenic virus-free plant to use for comparison is not always possible. However, in clover and in a model legume, *Lotus japonicus*, the coat protein gene of the partitivirus *White clover cryptic virus* 2 affects nodulation regulation (Nakatsukasa-Akune, Yamashita et al. 2005). In a recent study, we showed that a persistent virus protected its host from a vector of acute viruses (Safari et al 2017, under review).

The viruses in the family *Endornavirirdae* are found in plants, fungi, *oomycetes*, and protists. The endornavirus genome has a single stranded RNA genome (9.8– 17.6 kbp) that contains a single open reading frame (ORF) of around 5,000 amino acids (Fukuhara and Moriyama 2008). The virus in not encapsidated, and for this reason the replicative intermediate double-stranded (ds) RNA is most often found in hosts, but the RNA dependent RNA polymerase (RdRp) indicates that the genome is single-stranded (Roossinck, Sabanadzovic et al. 2011). Additional identified domains in the large polyprotein are Methyltransferase (MTR), Helicase (Hel), Capsular polysaccharide synthase protein (CPS), and Glycosyltransferase (GT) domains. All members of this family have a highly conserved RdRp at the C terminal end of the ORF, that is closest to the RdRp of closteroviruses. The other domains are not conserved among all endornaviruses (Roossinck, Sabanadzovic et al. 2011). A Hel domain is reported from most of the endornaviruses, but the origins of the helicase domains are clearly different in different viruses (Roossinck, Sabanadzovic et al. 2011). Endornaviruses are found in many plants including rice, wild rice, kidney bean, broad bean, barley, cucurbits, Yerba mate, Malabar spinach, Grapevine,

Avocado, and peppers (Wakarchuk and Hamilton 1990, Zabalgogeazcoa and Gildow 1992, Moriyama, Nitta et al. 1995, Pfeiffer 1998, Coutts 2005, Valverde and Gutierrez 2008, Espach, Maree et al. 2012, Villanueva, Sabanadzovic et al. 2012, Debat, Grabiele et al. 2014, Okada, Moriyama et al. 2014). All the plants with endornaviruses are reported as healthy and normal, although in *Vicia faba* male sterility is associated with the presence of *Vicia faba endornavirus* (Pfeiffer 1998).

Oryza sativa endornavirus is found in all japonica cultivars, but not in indica cultivars (Horiuchi, Moriyama et al. 2003). *Oryza rufipogon* is the ancestor of domesticated japonica rice and has a related persistent virus, *Oryza rufipogon endornavirus*. Around 10,000 years ago, the hosts of these two viruses diverged during the cultivation of rice (Molina, Sikora et al. 2011). Therefore, assuming the virus was already present at the time of divergence, these viruses have diverged by only about 24% during 10,000 years. This is a surprisingly slow evolution rate for RNA viruses with previously reported rate of 10^{-3} to 10^{-4} substitution/site/year (Roossinck 1997, Roossinck and Ali 2007, Pagán and Holmes 2010).

A large dsRNA was first reported from tissue extracts of healthy bell peppers (*Capsicum annuum*) (Valverde, Nameth et al. 1990), and in 2007 the partial sequence analysis revealed that the dsRNA is the genome of an endornavirus, now named *Bell pepper endornavirus* (BPEV) (Valverde and Gutierrez 2007). The vertical transmission of BPEV is very high (close to 100%) when both parents are virus infected (Okada, Kiyota et al. 2011). The single ORF of BPEV starts at nucleotide 225 and ends at nucleotide 14,670, encoding a 4815 aa polyprotein with an estimated molecular mass of 545 kDa (Okada, Kiyota et al. 2011). BPEV contains four identified domains: MTR, Hel, GT, and RdRp (Figure 1-1) (Roossinck, Sabanadzovic et al. 2011). It is presumed that the large protein is proteolytically cleaved into functional doamins, although a protease has not been identified. Studying pepper viromes revealed co-infection with BPEV and

several acute viruses in Pusa Jwala (PJ) and Taiwan 2 (TW) pepper cultivars (Jo, Choi et al. 2017).

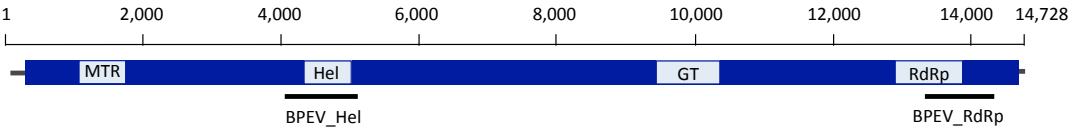

Figure **1-1.** Schematic representation of the genome organization of BPEV, and BPEV_Hel and BPEV_RdRp fragments positions. The analagous regions of Hel and RdRp were used for CFEV 1 phylogenies.

In addition to BPEV, *Hot pepper endornavirus* (HPEV) was reported from a chili pepper (*C. annuum*), with 72% identity to BPEV Yolo Wonder isolate, and containing all four identified domains in BEPV (Lim, Kim et al. 2015), and an additional endornavirus, *Capsicum frutescens* endornavirus 1, was isolated from the tabasco cultivar of *C. frutescens* (Dr. Rodrigo Valverde, personal communication).

Peppers are economically important crops of the *Solanaceae* family. A broad area along the Andes in South America including Peru, Ecuador and Colombia is thought to be the origin of *Capsicum* (García, Barfuss et al. 2016). The genus apparently migrated from South America northward around 6Ma resulting in the dispersal and speciation events. Bolivia and Peru are thought to be the centers of diversification and cultivation of peppers (García, Barfuss et al. 2016). The *Capsicum* genus consists of approximately 35 species; which *C. annuum* L., *C. baccatum* L., *C. chinense* Jacq., *C. frutescens* L., and *C. pubescens* (Ruiz & Pavon) are domesticated species that now grown all over the world (Garcia, Sterpetti et al. 2013, Qin, Yu et al. 2014). Domesticated species of peppers are derived from three distinct genetic lineages: one lineage includes the members of a species complex, *C. frutescens*, *C. annuum*, and *C. chinense*, that were derived independently from wild progenitors, whereas *C. pubescens* and *C. baccatum* are each independent lineages (Eshbaugh 1993, Pickersgill 1997, Hill, Ashrafi et al. 2013). The

most common cultivated species of *Capsicum* is *C. annuum*, which was domesticated in Mexico from chiltepin (*C. annuum* var. *glabriusculum*); chiltepin has small erect, pungent, and edible fruits, red in color when ripe (Tewksbury, Nabhan et al. 1999, Hernández-Verdugo, Guevara-González et al. 2001, González-Jara, Moreno-Letelier et al. 2011, Kraft, Brown et al. 2014). Cultivated peppers are divided into two main groups of sweet and pungent peppers. Bell peppers are all sweet *C. annuum* cultivars and are found in a variety of colors with blocky shaped fruits. Pungent peppers, often called chile peppers, refer to a large number of varieties with mild to extreme pungency and belong to *C. annuum*, *C. frutescens*, *C. chinense*, *C. pubescens*, and *C. baccatum* (Hulse-Kemp, Ashrafi et al. 2016).

Based on the whole genome sequencing data, domestication has led to an increase in the size of the pepper genome (Qin, Yu et al. 2014), but cultivation has resulted in a significant decline in genetic variation and the genetic structure of wild pepper populations and significant decreases in chiltepin populations have occurred over the last few thousand years (González-Jara, Moreno-Letelier et al. 2011). Until very recently the origins, relationships and domestication of peppers were unclear, but markers are now available that allow detailed phylogenetic trees of peppers to be constructed (Hill, Ashrafi et al. 2013, Kraft, Brown et al. 2014, Qin, Yu et al. 2014). For polymorphism detection and expression analysis in *Capsicum*, a Pepper GeneChipH array was developed; forty different *C. annuum* lines and cultivated *C. frutescens*, *C. chinense* and *C. pubescens* were tested resulting in detection of 33,401 single position polymorphism markers. The diversity in non-pungent peppers was reduced three-fold compared with pungent lines (Hill, Ashrafi et al. 2013). Endornaviruses are found in a number of both sweet and pungent pepper cultivars. In this study, we investigated the origin, natural history and evolution of two persistent viruses BPEV and CFEV 1 in different pepper cultivars. We used a SNPs-based method, Kompetitive Allele Specific PCR (KASP) assay (He, Holme et al. , Ashrafi, Hill et al. 2012, Hulse-Kemp, Ashrafi et al. 2016), to assess the evolutionary history of the hosts of these two

endornaviruses.

Materials and Methods

Pepper collection

We collected 96 different peppers including *C. annuum*, *C. chinense*, *C. frutescens*, *C. bacccutum, C. pubescens* and *C. chacoense* with different improvement levels (cultivars/landraces/wild materials). The seed sources and information are found in Table S1. Seeds were germinated in wet paper towels and then transferred to plastic pots containing Sun-Gro Horticulture soil and grown in an insect-free environmental room at 24°C and fluorescent light (16:8 light:dark photoperiod).

Screening *Capsicum* spp. for the presence of BPEV and CFEV 1

To screen plants for the presence of endornaviruses, dsRNAs were extracted from 5g of fresh weight leaves using the protocol described previously (Márquez, Redman et al. 2007).

For BPEV, based on the available sequences in GeneBank, primer pairs specific to part of the RdRp (BPEV_RdRp) and Hel domain (BPEV_Hel) were designed to amplify 1,240 bp and 1,162 bp fragments, respectively. Primers sequences and their positions are provided in Table 1-1 and Figure 1-1. Based on sequence information for CFEV 1 (kindly provided by Dr. Rodrigo Valverde, Louisiana State University), specific primers for part of the RdRp (CFEV1_RdRp) and Hel domain (CFEV1_Hel) were designed to amplify 1,101 bp and 1,016 bp, respectively (Table 1-1). For the RT reaction, we used about 2 µg of dsRNA mixed with 2 µM reverse primer, 0.5 mM of Tris-EDTA (pH 8.0) and nuclease-free water to a final volume of 12 µl, and boiled for 2

min. The mixture was incubated on ice for 2 min and 8 µl of Reverse Transcriptase (RT) mix [200 U of M-MuLV reverse transcriptase (New England Biolabs), 2 µl of 10X M-MuLV buffer (supplied by the manufacturer) and 10 mM dNTPs] was added and incubation continued at 42° C for 1 h. Then, cDNA was incubated with 10 µg of boiled RNase A (Sigma) for 15 min at room temperature and cleaned with E.Z.N.A Cycle Pure Kit (Omega Bio-tech) according to the manufactures instruction. About 0.5 µg of cleaned cDNA was used as a template for a 25 µl polymerase chain reaction (PCR) with Taq DNA Polymerase (ThermoFisher Scientific), buffer (30 mM MgCl2, Idaho Technologies), 2 mM dNTPs, and 0.2µM forward and reverse primers for each fragment. The PCR reactions were completed in capillary tubes with an Idaho Technologies Rapid Cycler for 40 cycles (94°C denaturation for 0 s, 48°C annealing for 0 s, and 72°C extension for 45 s). The RT-PCR products were separated on a 1.2% Agarose gel in 0.5X TBE, and appropriate bands were excised and purified using E.Z.N.A.® Gel Extraction Kit (Omega Bio-tech). Sequence analysis of the PCR products was done by the Genomic Core Facility of Pennsylvania State University, University Park, PA. The sequences have been deposited in GenBank under accession numbers listed in Table S2.

Nucleotide sequences were aligned with Clustal W using default settings in the program Geneious 10.0.9 (Drummond, Ashton et al. 2011). The alignment was visually corrected as necessary. Phylogenetic analysis was performed using MrBayes (Ronquist and Huelsenbeck 2003) implemented as a Geneious plug-in. The nucleotide sequences of 12 full length sequences of BPEV available in GenBank (listed in Table S3) were included in the phylogenetic analysis. *Hot pepper endornavirus* was used as an outgroup for the phylogenies of BPEV regions. For the phylogenies of all pepper endornaviruses *Phaseolus vulgaris endornavirus* 2 was used as an outgroup. For BPEV_RdRp and BPEV_Hel phylogenies HKY 98 + I substitution model, and for Endornaviridae _RdRp and _Hel phylogenies GTR + G + I were selected as the best fit models according to JModelTest (Darriba, Taboada et al. 2012); burn-in and chain length were 100,000

and 1,100,000, respectively.

Table **1-1**: List of the specific primers for BPEV and CFEV 1

Primers	Sequence	Position	Size (bp)[1]
BPEV_RdRp			1,240 bp
Forward	5' GAGTCTCTGGGAAGATACAG 3'	13,228-13,247	
Reverse	5' TACATCTGGACCCAGTGAGC 3'	14,352-14,371	
BPEV_Hel			1,162 bp
Forward	5' GCATGGGTAAAGGTGTTCGC 3'	3,934-3,953	
Reverse	5' GAAGTGTGTCGCTATGCTC 3'	5,077-5,095	
CFEV1_RdRp			1,101 bp
Forward	5' GAGCTTATGGGAAGACACTGAC 3'	-	
Reverse	5' GCCATCATGCAGCAGAACAC 3'	-	
CFEV1_Hel			1,016 bp
Forward	5' ATGGAGCCCAGGACCGAATA 3'	-	
Reverse	5' TTTGCCACCACCCAGTGTTTCC 3'	-	

[1] size of amplicon

Deep sequencing of the Marengo isolate of BPEV

The dsRNA extracted from Marengo bell pepper was treated with RQ1 DNase (Promega) and after cDNA preparation as described above using tagged primers for multiplexing, sequenced as part of a pool of 96 samples at the Oklahoma Medical Research Foundation, Clinical Genomic Center, using the Illumina HiSeq 3000 system. The raw data were trimmed using trimmomatic and quality was assessed by fastqc. Reads were assembled using Velvet (1.2.10) program, and mapped against the BPEV Yolo Wonder isolate (JN019858.1) as a reference sequence. The consensus sequence was used for the analysis.

Pepper phylogenies and population structure

To determine the phylogeny of peppers, DNA of plants listed in Table S4 was extracted using CTAB as described previously (Healy, Futrado et al. 2014) with some modifications. Plant tissue (0.5 g) was flash frozen in liquid nitrogen and ground to a fine powder, and mixed with 5 ml preheated extraction buffer (100 mM Tris-HCl pH 7.5, 25 mM EDTA, 1.5 M NaCl, 2% (w/v) CTAB, and 0.3% (v/v) β-mercaptoethanol). The sample was incubated at 65° C, mixed by inversion every 10 min for 30 min, followed by centrifugation for 5 min at 5000 × g. The supernatant was transferred to a fresh tube and mixed with an equal volume of phenol:chloroform (1:1, w:w). After mixing by inversion for 5 minutes, the sample was centrifuged for 10 min at 5000 × g and the upper phase was transferred to a new tube and treated with 5 µl RNAse A (Sigma-Aldrich; 10 mg/ml) for 15 min at 37 °C, followed by an additional extraction as above with phenol:chloroform. The final aqueous phase was precipitated with 0.3 M Sodium Acetate and two volumes absolute ethanol. The pellet was washed with 70% ethanol, dried, and resuspended in 100 µl of TE buffer. DNA samples were sent to TraitGenetics Company in Germany for KASP (Kompetitive Allele Specific PCR) assay with 96 SNPs. The KASP genotyping assay is a fluorescence-based reporting system for detection of the alleles (SNPs) at a specific locus within the genomic DNA (He, Holme et al. , Ashrafi, Hill et al. 2012). The SNP information is provided in Table S5.

SNPs within different pepper lines identified in the KASP analysis were used for a Bayesian cluster estimation of population structure using the STRUCTURE Software (version 2.3.4) (Pritchard, Stephens et al. 2000). STRUCTURE was run for K=1 to K=10 (K number of clusters), and for each K value ten replicates were performed with a burn-in of 35,000 followed by 30,000 Markov Chain Monte Carlo iterations after burn-in. The replicate with highest probability of K was selected. In addition, to reveal the phylogenetic origin of these pepper lines,

a phylogenetic tree was generated with the SNP data using MrBayes in Geneious. JModelTest suggested a Sym substitution model, but since it is not implemented in MrBayes several models were tested that all resulted the same tree and the GTR + I substitution model was selected as the best fit model; burn-in was 100,000 and total chain length was 1,100,000.

Results

Screening *Capsicum* spp. for the presence of endornaviruses

After dsRNA extraction from peppers, the presence of BPEV or CFEV 1 was confirmed by RT-PCR using specific primers for each virus. The RT-PCR products were sent for Sanger sequencing to obtain the consensus sequence for each virus. We detected BPEV in 10 out of 33 different *C. annuum* pepper cultivars including Marengo, Chocolate, Greek pepperonici, Hungarian YW, Peter, Joe's long Cayenne, Padron hot, Jimmy Nardello's, Neapolitan, and Feher ozon Paprika (Table 1-2, Table 1-3). BPEV has been detected mainly in the samples that were from the USA, except for Chocolate pepper from Guatemala (Figure 1-2).

Table **1-2:** Number of positive samples in each Capsicum species.

Capsicum Species	# of lines[1]	+ BPEV[2]	+ CFEV 1[3]
C. annuum	33	10	0
C. annuum var. *glabriusculum*	8	0	0
C. baccatum	16	0	13
C. frutescens	18	0	5
C. chinense	11	0	5
C. pubescens	8	0	1
C. chacoense	2	0	0

[1] Number of lines tested for BPEV and CFEV
[2] Number of lines positive for BPEV

[3] Number of lines positive for CFEV 1

We did not detect CFEV 1 in any *C. annuum* peppers, but 24 different peppers in diverse species of *Capsicum* were infected with this virus; 13 out of 16 *C. baccatum*, five out of 11 *C. chinense*, five out of 18 *C. frutescese*, and one out of 8 *C. baccatum* were positive for CFEV 1 (Table 1-2, Table 1-3). Origins of the seeds for the CPEV 1 positive plants are shown in Figure 1-2.

Table **1-3**: Peppers with a detected endornavirus

Names in the tree[1]	Type	Species	PI[2]	Virus
BPEV_Marengo	Sweet	*C. annuum*	-	BPEV
BPEV_Jimmy Nardello's	Pungent	*C. annuum*	-	BPEV
BPEV_Feher ozon paprika	Sweet	*C. annuum*	-	BPEV
BPEV_Joe's long cayenne	Pungent	*C. annuum*	-	BPEV
BPEV_Greek pepperoncini	Pungent	*C. annuum*	-	BPEV
BPEV_ Hungarian YW	Pungent	*C. annuum*	-	BPEV
BPEV_Padron Hot	Pungent	*C. annuum*	-	BPEV
BPEV_Neapolitan	Pungent	*C. annuum*	-	BPEV
BPEV_Peter	Pungent	*C. annuum*	-	BPEV
BPEV_Chocolate	Sweet	*C. annuum*	PI 666471	BPEV
CFEV1_Habanero Red	Pungent	*C. chinense*	-	CFEV1
CFEV1_Chile blanco	Pungent	*C. chinense*	PI 574545	CFEV1
CFEV1_30040	Pungent	*C. chinense*	PI 159236	CFEV1
CFEV1_Royal gold	Pungent	*C. chinense*	PI 315023	CFEV1
CFEV1_Lemon drop	Pungent	*C. chinense*	PI 315024	CFEV1
CFEV1_Aribibi gusano	Pungent	*C. frutescens*	PI 573337	CFEV1
CFEV1_Greenleaf tabascoAL	Pungent	*C. frutescens*	PI 634826	CFEV1
CFEV1_Tabasco L-167	Pungent	*C. frutescens*	PI 640909	CFEV1
CFEV1_Habanero	Pungent	*C. chinense*	-	CFEV1
CFEV1_Ecu	Pungent	*C. pubescens*	PI 585262	CFEV1
CFEV1_Lemon drop hot	Pungent	*C. baccatum*	-	CFEV1
CFEV1_WTS-14	Pungent	*C. baccatum* var. *pendulum*	PI 595905	CFEV1
CFEV1_WWMC126	Pungent	*C. baccatum* var. *pendulum*	PI 632927	CFEV1
CFEV1_WW141	Pungent	*C. baccatum* var. *pendulum*	PI 633756	CFEV1
CFEV1_MC145	Pungent	*C. baccatum* var. *pendulum*	PI 633757	CFEV1
CFEV1_MC147	Pungent	*C. baccatum* var. *pendulum*	PI 633758	CFEV1
CFEV1_Malagueta	Pungent	*C. baccatum* var. *pendulum*	PI 260543	CFEV1
CFEV1_Valentine	Pungent	*C. baccatum* var. *pendulum*	PI 260549	CFEV1
CFEV1_Omnicolor	Pungent	*C. baccatum* var. *pendulum*	PI 260590	CFEV1
CFEV1_82	Pungent	*C. baccatum* var. *pendulum*	PI 337522	CFEV1
CFEV1_BGH 4215	Pungent	*C. baccatum* var. *pendulum*	PI 441589	CFEV1
CFEV1_470	Pungent	*C. baccatum* var. *baccatum*	PI 215699	CFEV1
CFEV1_No 1553	Pungent	*C. baccatum* var. *baccatum*	PI 238061	CFEV1
HPEV_Ember	Pungent	*C. annuum*	PI 273426	HPEV

[1] Virus name_cultivar name or plant designation
[2] PI, Plant Identification number

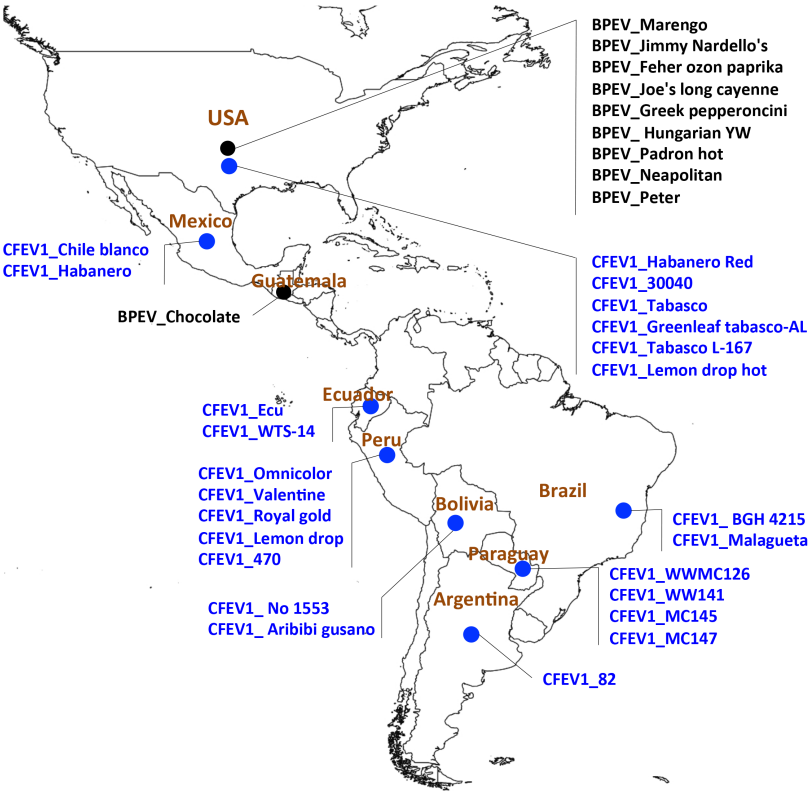

Figure **1-2**: The distribution map of detected endornaviruses: *Bell pepper endornavirus* (BPEV) and *Capsicum frutescens* endornavirus 1 (CFEV 1). Location indicates the country that the plants were obtained from.

Using CFEV1_RdRp primers an 875 bp fragment was amplified from dsRNA extracted from the Ember cultivar (*C. annuum*), which had 94% nt identity to the Cs isolate of HPEV reported from South Korea (Lim, Kim et al. 2015). We did not detect BPEV or CFEV 1 in eight accessions of chiltepin (*C. annuum* var. *glabriusculum*) or two accessions of *C. chacoense* in our pepper collection.

Deep sequencing of Marengo isolate of BPEV

Using Illumina sequencing we obtained a contig of 14,290 bp from Marengo peppers, corresponding to nt 223 to 14,512 of BPEV that differed from the strain amplified using BPEV

specific primers (strain Marengo1). The RdRp and Hel fragments of this isolate (Marengo2) had 90% and 88% identity to BPEV Marengo1, respectively. The Marengo2 strain had 99% identity to the Maor isolate (KP455654.1), while both RdRp and Hel fragments of BPEV Marengo1 showed highest identity (99% and 97%, respectively) to the Yolo Wonder isolate (JN019858.1).

Virus phylogenies

Pairwise identity was 88-100% among BPEV_RdRp nt sequences, and 87-100% among BPEV_Hel sequences. Phylogenetic analysis using the nt sequence of both fragments resulted in two main clades (Figure 1-3). In the BPEV_RdRp phylogeny clade 1, viruses isolated from different pepper varieties (Marengo1, Greek pepperonici, Hungarian YW, Peter, Joe's long cayenne, Padron hot, Jimmy Nardello's, and Neapolitan) are grouped together with Yolo Wonder, PJ, and TW isolates (accession numbers in Table S3). All isolates in clade 1 are from pungent peppers, except Marengo1 and Yolo Wonder. BPEV isolates from Chocolate, Feher ozon paprika, and Marengo2 strain are grouped in clade 2, which included all sweet pepper isolates (lj, Phenol, Santa Fe, Kyosuzu, IS, Maor, and Haeley) except Chocolate and Santa Fe isolates which are pungent peppers (Muñoz-Baena, Marín-Montoya et al. 2017). In BPEV_Hel phylogeny the same pattern of two clades was observed; the clade 1 contained isolates from pungent peppers, and clade 2 had the sweet pepper isolates (except Chocolate and Santa Fe). However, the BPEV_Hel fragment from Yolo Wonder grouped with the other sweet pepper isolates (Figure 1-3B, Clade 2), which indicates recombination in this virus as was previously suggested in another study (Jo, Choi et al. 2016).

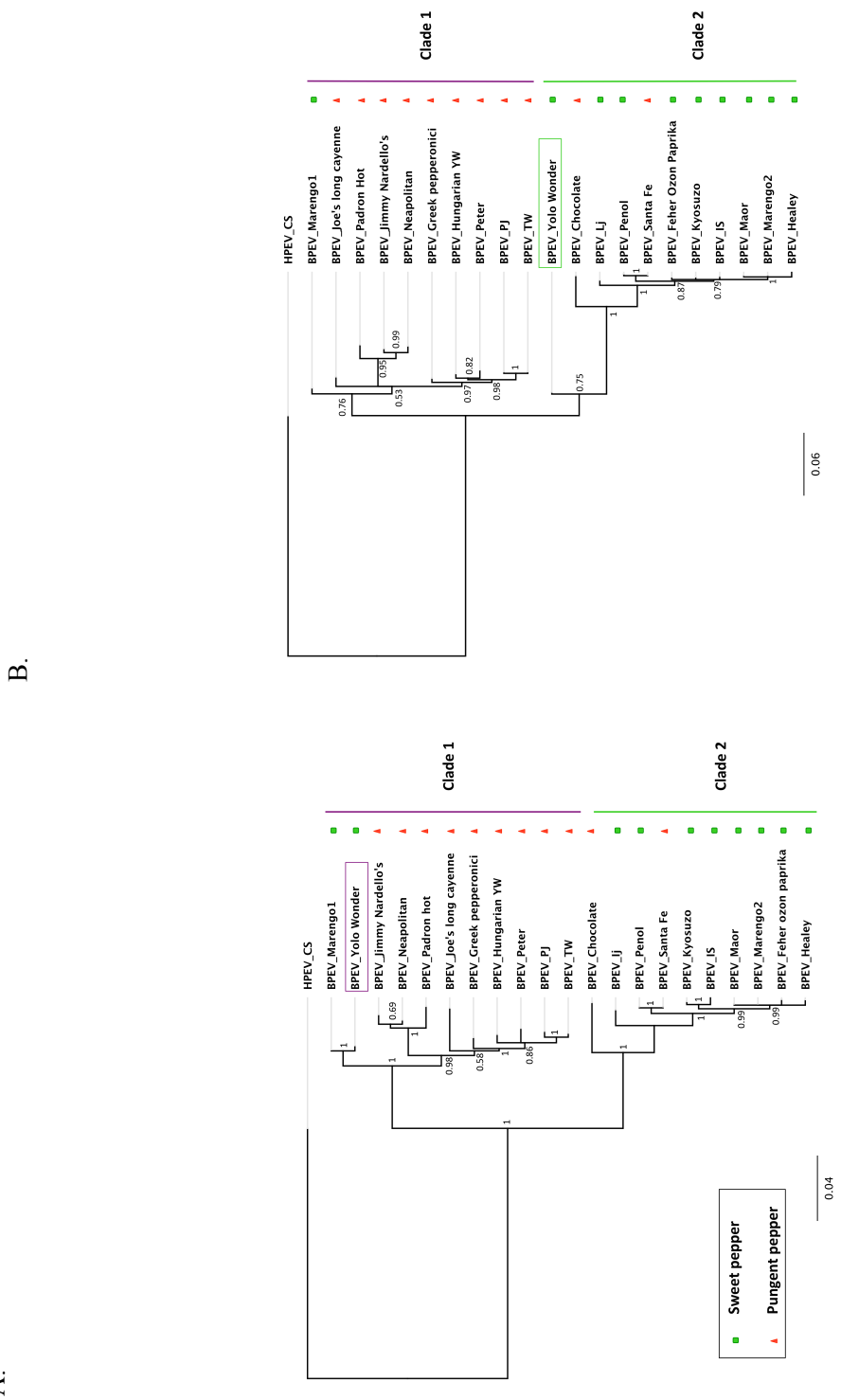

Figure **1-3**: Bayesian analysis of the relationships among aligned nt sequences of BPEV isolated from different peppers as described in Table 2 and S4. Trees were generated by MrBayes in Geneious 10.0.9 using the best fit model (HKY + I). Posterior probilities are shown on nodes. HPEV was used as an outgroup. A. BPEV_RdRp; B. BPEV_Hel.

A Bayesian phylogenetic tree was constructed using nt sequences of pepper endornavirus RdRp regions (BPEV, CFEV 1, and HPEV isolates), using *Phaseolus vulgaris endornavirus* 2 as an outgroup (Figure 1-4). This tree contains two main clades: two HPEV isolates (Cs and Ember) were grouped together with BPEV isolates in clade A, all of which are isolated from *C. annuum* species. The second clade contained the CFEV 1 isolates. CFEV 1 from Aribibi gusano pepper (*C. frutescens*) was distinct from the rest of isolates and with a posterior probability of 1 may be ancestral to the rest of the CFEV 1 isolates. CFEV 1 isolated from two Habaneros, Royal gold, and 30040 peppers (all from *C. chinense*), were grouped with three Tabasco peppers (*C. frutescens*) and placed in clade CF. The BCP clade included isolates mostly from *C. baccatum,* with the exceptions of Ecu (from *C. pubescens*), and Chile blanco and Lemon drop (*C. chinense*) (Figure 1-4). The CFEV 1 isolates from Aribibi gusano and Habanero Red showed 83-85% identity to other CFEV 1 isolates, while there was 94-100% nt identity among isolates in the CF clade, and 90-100% identity among isolates in the BCP clade.

Using specific primers the CFEV1_Hel fragment was amplified only from Tabasco (Tabasco L-64 and Greenleaf Tabasco; both *C. frutescens*) and 30040 (*C. chinense*) peppers. Also, in the Endornaviruses Hel phylogeny all *C. annum* isolates (BPEV and HPEV isolates) grouped together and CFEV 1 isolates from *C. frutescens* and *C. chinense* grouped in a second clade (Figure 1-5).

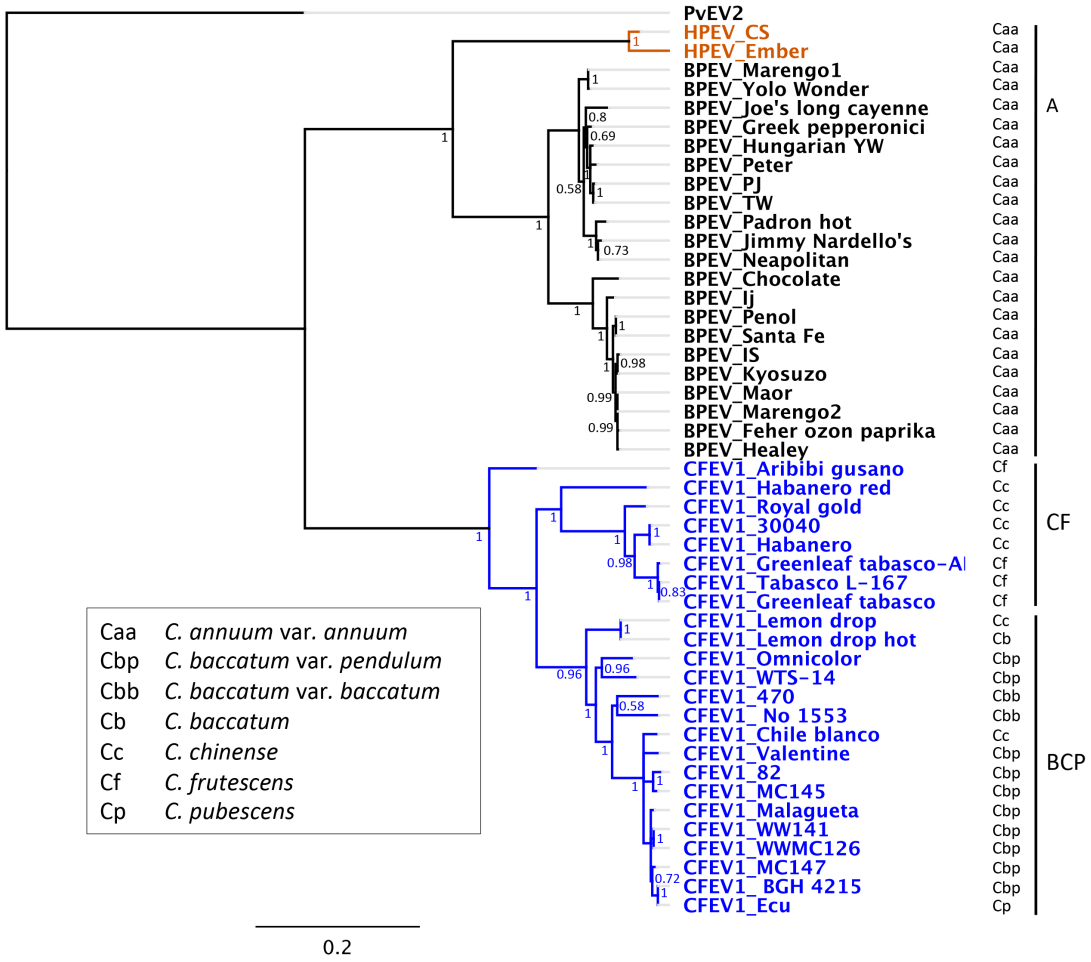

Figure **1-4**: Bayesian analysis of the relationships among aligned nt sequences of the RdRp fragment of pepper Endornaviruses. PvEV 2, *Phaseolus vulgaris endornavirus* 2, used as an outgroup. Trees were generated by MrBayes in Geneious 10.0.9 using GTR + I +G. Posterior probilities are shown on nodes.

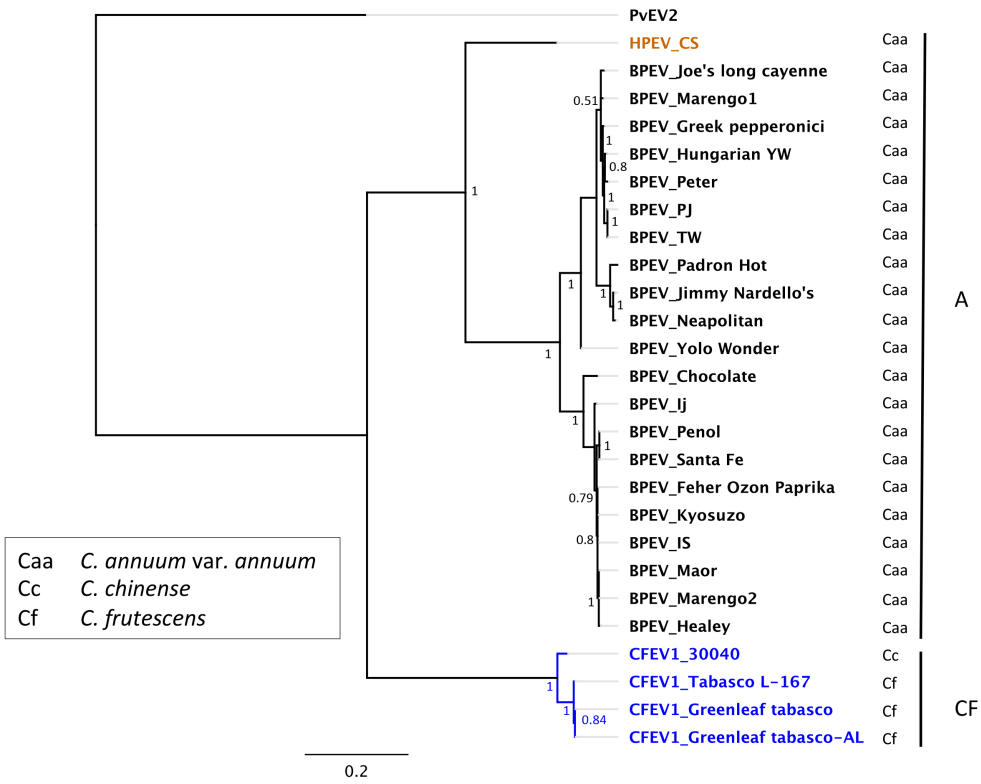

Figure **1-5**: Bayesian analysis of the relationships among aligned nt sequences of the pepper Endornaviruses_Hel fragment. Used PvEV2 as the outgroup. Tree is generated by MrBayes in Geneious 10.0.9 using GTR + I +G model.

Pepper phylogenies and population structure

To examine the population structure in the *Capsicum* spp. using SNPs data from the KASP analysis Bayesian clustering was implemented in STRUCTURE software. The distribution of the 38 *Capsicum* spp. lines into three distinct clusters is shown in Figure 1-6 (red=cluster 1; green=cluster 2; blue=cluster 3). In cluster 1, Marengo, Jimmy Nardello's, Padron hot, Ember, Greek pepperoncini, Hungarian YW, Long red cayenne, Feher ozon paprika, Joe's long Cayenne, Chiltepin 102, and Diente de perro are clustered together. Chiltepin54, Serrano hot, Jalapeño, Chiltepin53, Peter, Chocolate, and Jalapeño criollo are grouped together in cluster 2. Cluster 3

includes Chiltepin30, Chile blanco, WWMC 126, Ecu, Lemon drop hot, Aribibi gusano, 30040, Omnicolor, Habanero red, Valentine, Royal gold, WW141, Malagueta, No1553, 470, BGH4215, Lemon drop, Tabasco L167, Greenleaf tabasco Al, and MC145.

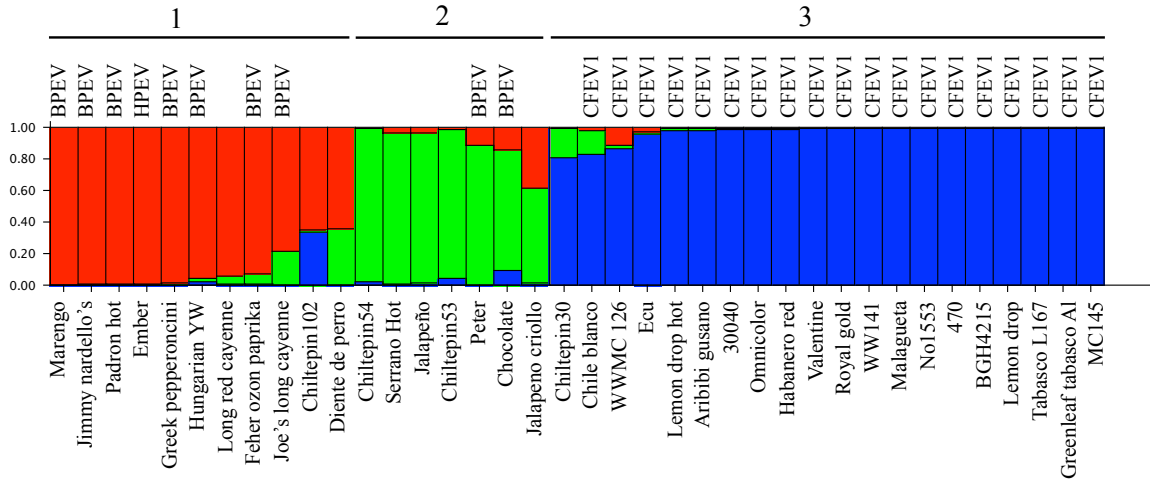

Figure **1-6**: Population strucure analysis of SNPs data using STRUCTURE software (2.3.4). The distribution of the 38 *Capsicum* spp. lines into three distinct clusters is shown in red=cluster 1, green=cluster 2, blue=cluster 3.

Population structure analysis revealed admixture in Joe's long cayenne and Diente de perro from cluster 1 with peppers in cluster 2. Peter, Chocolate, and Jalapeño criollo peppers have admixtures with genotypes in cluster 1. Peppers in cluster 1 and 2 belong to *C. annuum*, expect Diente de perro, which has been identified as *C. frutescens* based on its morphological characteristic, but this result indicates that it belongs to *C. annuum*. Pepper lines in cluster 3 corresponded to peppers positive for CFEV 1 except Chiltepin30. Chiltepin30 and Chile blanco showed admixed genome (<20%) with cluster 2. Based on our SNPs data few variations were detected in SNPs data from non-*annuum* species and they were grouped together in cluster 3.

Ecu pepper (*C. pubescens*) was used as an outgroup for the Baysian phylogenetic tree for the 38 pepper samples. This tree is in general agreement with the previous population clustering,

and the endornaviruses phylogenies (Figure 1-7). The clade highlighted in Red includes hosts of clade 1 from the BPEV phylogeny, including Feher ozon paprika. However, Peter and Chocolate peppers both positive for BPEV are in the green highlighted part of the pepper phylogenies. All pepper lines in the red and green highlighted groups belonged to *C. annuum* and are clustered with Chile blanco (*C. chinense*), and Chiltepin30; Chiltepin102 is the probable ancestor. The final grouping includes the various non-*annuum* species of peppers, where all of the CFEV 1 positive plants are placed, and in which there is less than 2% variation in the SNP data.

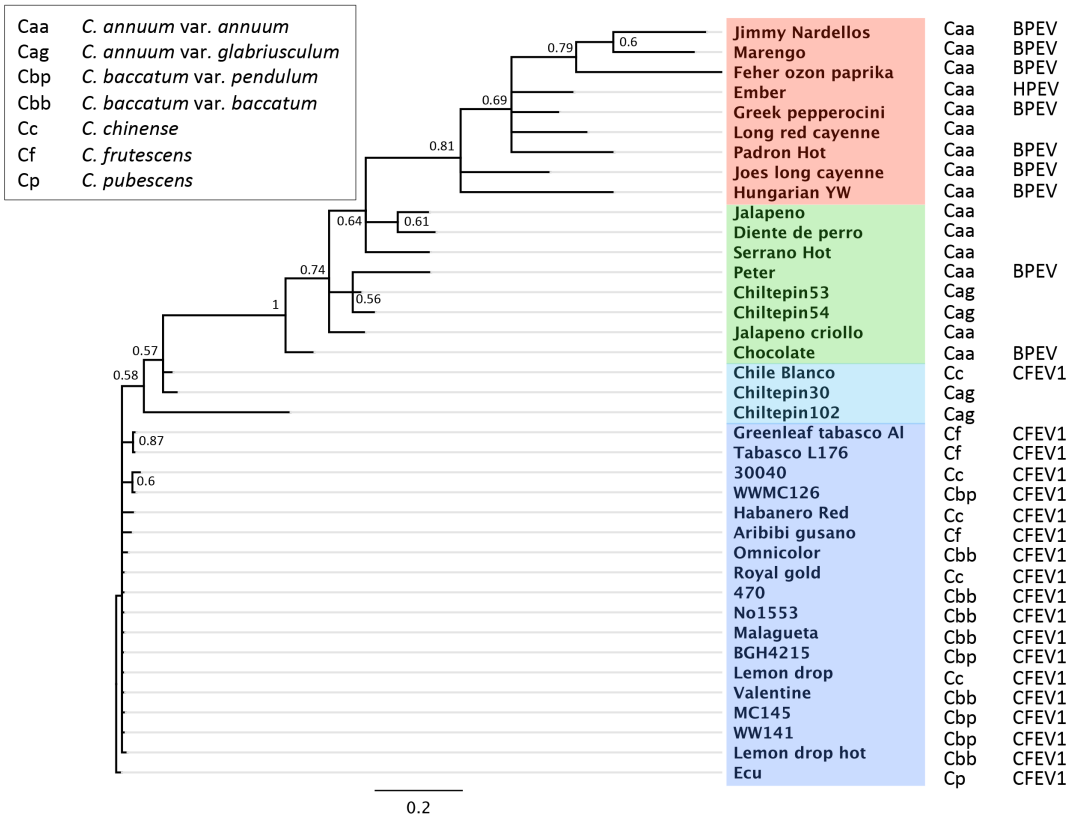

Figure **1-7**: Bayesian analysis of the relationships among SNPs in different pepper lines. Ecu was used as an outgroup. Tree was generated by MrBayes in Geneious 10.0.9 using GTR+ I. Posterior probilities are shown on nodes.

Discussion

In this study we detected BPEV only in *C. annuum*, and none of the other *Capsicum* species. Previously, BPEV was reported from several domesticated *Capsicum* species (*C. annuum*, *C. frutescens, C. chinense,* and *C. baccatum*) by using degenerate primers to amplify a 381 nt fragment of the RdRp (Okada, Kiyota et al. 2011). However, with a more comprehensive analysis of the RdRp we found that all the non-*annuum* isolates were CFEV 1. The limitation of BPEV to *C. annuum* is surprising because this virus can be transmitted by cross pollination (Valverde and Gutierrez 2007), and hybrids of C. *annuum* with *C. baccatum* and *C.chinense* are fertile.

BPEV has been reported from the USA, Colombia, Canada, China, Japan, Taiwan, India, South Korea, and Israel (Valverde and Gutierrez 2007, Okada, Kiyota et al. 2011, Lim, Kim et al. 2015, Jo, Choi et al. 2016, Muñoz-Baena, Marín-Montoya et al. 2017). In this study, all cultivars were from the Americas, and BPEV was found in the USA samples, and a Chocolate pepper from Guatemala. Hence, BPEV has been distributed all over the world by pepper dispersal. In structure analysis and pepper phylogenies (Figures 1-6 and 1-7), Chocolate pepper was nested within chiltepins, the ancestor of *C. annuum* peppers. While BPEV has not been identified in any chiltepin in this study or a previous study (Okada, Kiyota et al. 2011)), Chocolate pepper may be the origin of introduction of this virus to *C. annuum*.

The admixture observed within *C. annum* samples (Figure 1-6) revealed the exchange of genetic material between cultivars by breeding and natural cross-pollination. Among four different chiltepins (*C. annuum var. glabriusculum*) used in this analysis, Chiltepin102 included in cluster 1 revealed an admixture (~25%) with cluster 3, Chiltepin53 and Chiltepin54 clustered with peppers in cluster 2, and Chiltepin30 in cluster 3 with an admixed genome with cluster 1. This result is consistent with a previous study that used microsatellite markers for pepper

(Nicolaï, Cantet et al. 2013).

Phylogenetic analysis of nucleotide sequences of BPEV RdRp and Hel fragments revealed two main clades in both domains corresponding to the pepper pungency (Figure 1-3); clade 1 included BPEV isolates mainly from pungent pepper, and clade 2 from sweet peppers. Two strains of BPEV in Marengo were separated in both BPEV phylogenies. Strain Marengo1 was nested between isolates from pungent peppers, and the Marengo2 strain was placed within sweet pepper isolates. The presence of different strains could result from cross-pollination between peppers, since persistent viruses are transmitted by pollen and ovum (Roossinck 2010). SNP profiles in the transcriptome data from a Maor pepper revealed more than one BPEV variant in this cultivar as well (Jo, Choi et al. 2016).

BPEV_Hel from Yolo Wonder showed evidence of recombination in this virus. A schematic diagram of BPEV_Hel alignment is shown in Figure 1-8A, with the color code based on the SNPs similarity between different isolates. In Yolo Wonder sequences were similar to clade 1 isolates until position 875 (H1, shown in purple), while after position 875 (H2, shown in green) sequences were similar to isolates in clade 2. In the phylogenetic trees of the H1 and H2 fragments (Figure 1-8B) the Yolo Wonder isolate is grouped within clade 1 in H1, with pungent pepper isolates, and in clade 2 in the H2 phylogeny with sweet pepper isolates. In the H1 fragment pairwise identity was 88-100%, while identity was 80-100% in H2 (Figure 1-8C); so there is more variation is in the second part (H2) of the BPEV_Hel fragment. Except for three positions all the substitutions are synonymous.

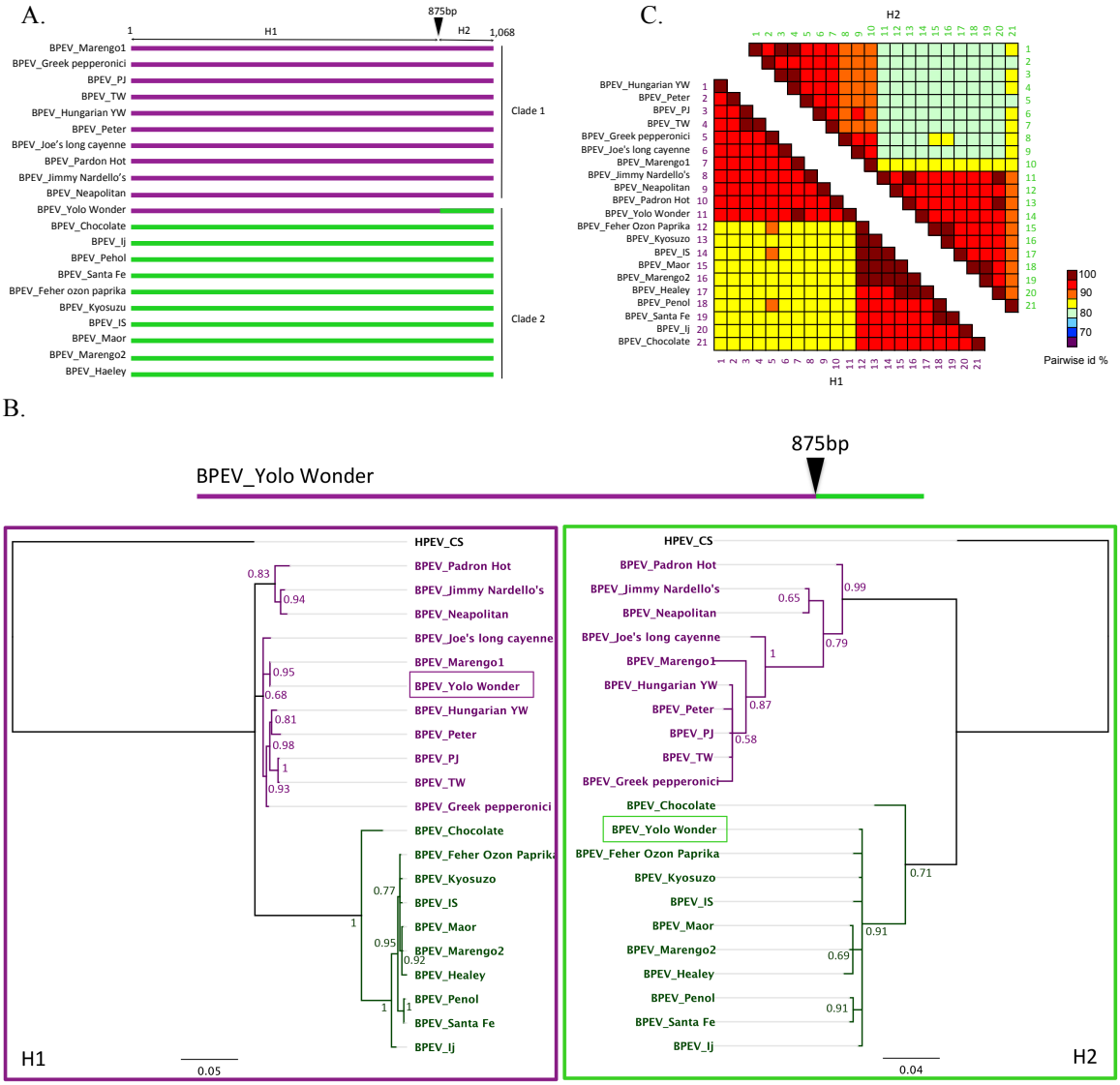

Figure **1-8**: Recombination in BPEV_Hel fragment. A. Schematic nt sequence alignment display of BPEV_Hel fragment, color is based on the similarity of the sequences. Purple corresponds to the first 875 nt of the BPEV_Hel, and green represents the rest of the sequence (~200bp). B. Bayesian analysis of the relationships among aligned nt sequences of the BPEV_Hel isolated from different peppers for the H1 and H2 alignments. Trees were generated by MrBayes in Geneious 10.0.9 using the best fit model (HKY + I). Posterior probilities are shown on the nodes. HPEV was used as an outgroup. C. Pairwise identity plot of BPEV-Hel for both H1 and H2 fragments.

The phylogenetic tree of pepper endornaviruses was correlated with pepper species with few exceptions. Two HPEVs (Cs and Ember isolates) and all BPEV isolates were included in

clade A, and they are all from *C. annuum* species. The CFEV 1 isolates had two main clades: isolates in the CF clade were from *C. chinense* and *C. frutescens*, and the BCP clade contained isolates mainly from *C. baccatum* peppers. Species classifications have been based on morphological traits (flowers, leaves, and fruits), which were defined by different people. So, in addition to genome introgression by interspecific crossing between compatibilities species, there is also a possibility of previous misclassification of species (Shirasawa, Ishii et al. 2013).

In Figure 1-9, the BPEV_RdRp (A) and *C. annuum* peppers (including Chile blanco) (B) phylogenies are shown by connecting viruses to their host; in many samples congruence is observed, and the admixture is likely related to cross-pollination between different peppers that can lead to viral transmission.

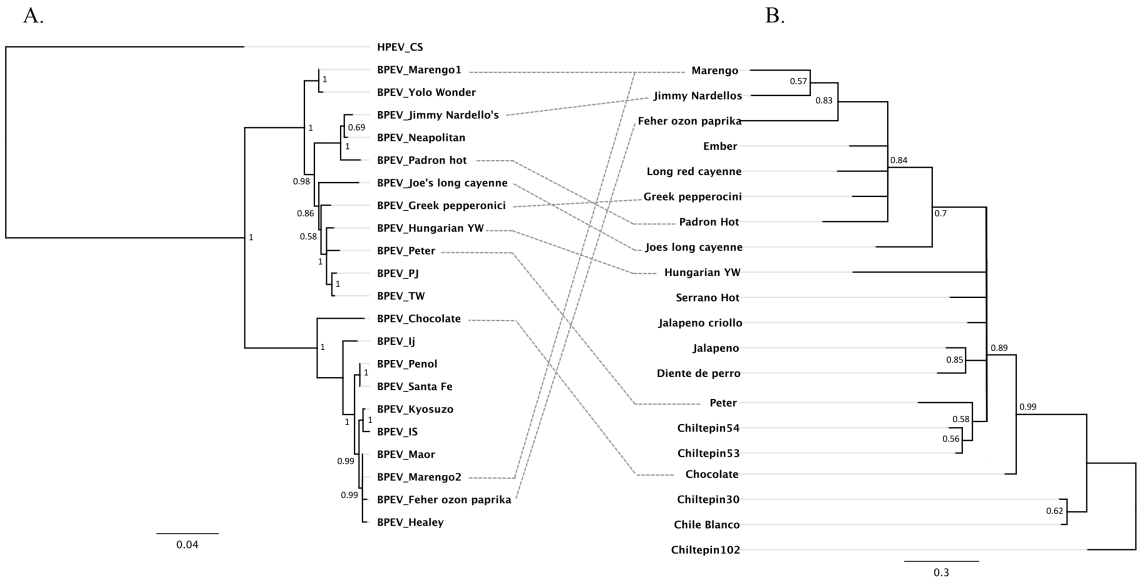

Figure **1-9**: Showing the congruence between the BPEV_RdRp (A) and pepper (B) phylogeny. In pepper phylogeny *C. annuum* samples and Chile banco were included, and Chiltepin102 was used as an outgroup. The tree was generated by MrBayes in Geneious 10.0.9 using GTR+ I. Posterior probilities are shown on nodes.

CFEV 1 was detected in four species of *Capsicum*, but in none of our *C. annuum* peppers. Ecu was the only *C. pubsecens* positive for this virus; however, it could have been due to our small collection (eight samples) from this species. While our SNP markers did not resolve non-*annuum* phylogenies to verify the species identification for this pepper, the possibility of misclassification of this pepper is low since morphological traits of *C. pubescens* and *C. baccatum* are quite distinct (Jarret 2008). Hybrids of *C. pubsecens* with the other four *Capsicum* spp. are sterile, so, there is no chance for the virus to be transmitted by cross-pollination. Hence this is evidence for the presence of an endornavirus in the ancestor of all of these *Capsicum* species.

Most of the screened *C. baccatum* peppers were positive for CFEV 1. CFEV 1 isolates from 470 and No 1553 peppers (*C. baccatum* var. *baccatum*), the wild form of this species (Nimmakayala, Abburi et al. 2016), are grouped together in bcp clade, separate from the other isolates. The high incidence of CFEV 1 in *C. baccatum*, and its detection in the wild subspecies implies that *C. baccatum* is the original host of CFEV 1. During *Capsicum* speciation, breeding, and dispersal this virus could have been transmitted to *C. frutescens* and *C. chinense*. The interspecific compatibility between these three species also supports this hypothesis (Walsh and Hoot 2001). The locations of *C. baccatum* peppers infected with CFEV 1 were mainly from South America (Peru, Bolivia, Ecuador, Paraguay, Brazil and Argentina), which is congruent with origin and speciation of *Capsicum* spp. During pepper speciation the ancestor of CFEV 1 may have evolved as a new endonavirus, BPEV, in *C. annuum* peppers.

Different approaches such as EST-SSR, isozyme, plastid DNA, and SNP analysis confirmed that *C. chinense* and *C. frutescens* were closely related to *C. annuum* with *C. baccatum* more distant (Walsh and Hoot 2001, Jarret 2008, Shirasawa, Ishii et al. 2013). It is interesting that we did not detect any BPEV in non-*annuum* capsicum, while CFEV 1 was found in *C. chinense* and *C. frutescens* and *C. baccatum*. It can be concluded that the ancestor of these Capsicums was

infected by an endornavirus, and during speciation and domestication of peppers, the virus was transmitted and diverged in different peppers.

Cited Literature

Ashrafi, H., et al. (2012). "*De novo* assembly of the pepper transcriptome (*Capsicum annuum*): a benchmark for *in silico* discovery of SNPs, SSRs and candidate genes." BMC Genomics **13**: 571.

Coutts, R. H. A. (2005). "First report of an *Endornavirus* in the *Cucurbitaceae*." Virus Genes **31**(3): 361-362.

Darriba, D., et al. (2012). "jModelTest 2: more models, new heuristics and parallel computing." Nat. Methods **9**(8): 772.

Debat, H. J., et al. (2014). "The complete genome of a putative endornavirus identified in yerba mate (*Ilex paraguariensis* St. Hil.)." VIrus Genes **49**: 348-350.

Drummond, A. J., et al. (2011). Geneious.

Eshbaugh, W. H. (1993). Peppers: History and Exploitation of a Serendipitous New Crop Discovery. New Crops. J. Janick and J. E. SImon. New York, Wiley**:** 132-139.

Espach, Y., et al. (2012). "Complete genome of a novel endornavirus assembled from next generation sequence data." J. Virol. **86**(23): 13142.

Fukuhara, T. and H. Moriyama (2008). Endornaviruses. Encyclopedia of Virology. B. W. J. Mahy and M. H. V. vanRegenmortel. Oxford, Elsevier. **2:** 109-116.

García, C. C., et al. (2016). "Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum*, Solanaceae)." Annals of Botany **118**: 35-51.

Garcia, C. C., et al. (2013). Wild Capsicums: identification and *In situ* analysis of Brazilian species. Breakthroughs in the Genetics and Breeding of Capsicum and Eggplant, Torino, Italy, Comitato per l'organizzazione degli evenit (COE) DISAFA.

González-Jara, P., et al. (2011). "Impact of human management on the genetic variation of wild pepper, *Capsicum annuum* var. *glabriusculum*." PLoS ONE **6**(12): e28715.

He, C., et al. SNP Genotyping: The KASP Assay. Crop Breeding: Methods and Protocols. D. Fleury and R. Whitford. New York, Springer. **1145:** 75-86.

Healy, A., et al. (2014). "Protocol: a simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species." Plant Methods **10**: 21.

Hernández-Verdugo, S., et al. (2001). "Screening wild plants of *Capsicum annuum* for resistance to pepper huasteco virus (PHV): presence of viral DNA and differentiation among populations." Euphytica **122**: 31-36.

Hill, T. A., et al. (2013). "Characterization of *Capsicum annuum* genetic diversity and population structure based on parallel polymorphism discovery with a 30K unigene pepper genechip." PLoS ONE **8**(2): e56200.

Horiuchi, H., et al. (2003). "Inheritance of *Oryza* sativa endornavirus in F1 and F2 hybrids betwween *japonica* and *indica* rice." Genes Genet Syst **78**: 229-234.

Hulse-Kemp, A. M., et al. (2016). "A HapMap leads to a *Capsicum annuum* SNP infinium array: a new tool for pepper breeding." Hort. Res. **3**: 10634.

Jarret, R. L. (2008). "DNA barcoding in a crop genebank: the *Capsicum annuum* species complex." Open Biol. J. **1**: 35-42.

Jo, Y., et al. (2017). "The pepper virome: natural co-infection of diverse viruses and their quasispecies." BMC Genomics **18**: 453.

Jo, Y., et al. (2016). "*In silico* identification of *Bell pepper endornavirus* from pepper transcriptomes and their phylogenetic and recombination analyses." Gene **575**: 712-717.

Kraft, K. H., et al. (2014). "Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico." Proc. Natl. Acad. Sci. USA **111**(17): 6165-6170.

Lim, S., et al. (2015). "Complete genom sequence of a novel endornavirus isolated from hot pepper." Arch. Virol. **160**: 3153-3156.

Márquez, L. M., et al. (2007). "A virus in a fungus in a plant --three way symbiosis required for thermal tolerance." Science **315**: 513-515.

Molina, J., et al. (2011). "Molecular evidence for a single evolutionary origin of domesticated rice." Proc. Natl. Acad. Sci. USA **108**(20): 8351-8356.

Moriyama, H., et al. (1995). "Double-stranded RNA in rice: a novel RNA replicon in plants." Mol Gen Genet **248**: 364-369.

Muñoz-Baena, L., et al. (2017). "Genome sequencing of two *Bell pepper endornavirus* (BPEV) variants infecting *Capsicum annuum* in Colombia." Agronomía Colombiana **35**(1): 44-52.

Nakatsukasa-Akune, M., et al. (2005). "Suppression of root nodule formation by artificial expression of the *TrEnodDR1* (coat protein of *White clover cryptic virus 2*) gene in *Lotus japonicus*." Mol. Plant-Microbe Interact. **18**(10): 1069-1080.

Nicolaï, M., et al. (2013). "Genotyping a large collection of pepper (*Capsicum* spp.) with SSR loci brings new evidence for the wild origin of cultivated *C. annuum* and the structuring of genetic diversity by human selection of cultivar types." Genet. Resour. Crop Evol. **60**: 2375-2390.

Nimmakayala, P., et al. (2016). "Genome-wide divergence and linkage disequilibrium analyses for *Capsicum baccatum* revealed by genome-anchored single nucleotide polymorphisms." Frontiers in Plant Science **7**: 1646.

Okada, R., et al. (2011). "Bell pepper endornavirus: molecular and biological properties, and occurrence in the genus *Capsicum*." J. Gen. Virol. **92**: 2664-2673.

Okada, R., et al. (2014). "A new endornavirus species infecting Malabar spinach (*Basella alba* L.)." Arch. Virol. **159**: 807-809.

Pagán, I. and E. C. Holmes (2010). "Long-term evolution of the *Luteoviridae*: time scale and mode of virus speciation." J. Virol. **84**(12): 6177-6187.

Pfeiffer, P. (1998). "Nucleotide sequence, genetic organization and expression strategy of the double-stranded RNA associated with the '447 cytoplasmic male sterility trait in *Vicia faba*." Virology **79**: 2349-2358.

Pickersgill, B. (1997). "Genetic resources and breeding of *Capsicum* spp." Euphytica **96**: 129-133.

Pritchard, J., et al. (2000). "Inference of population structure using multilocus genotype data." Genetics **155**: 945-959.

Qin, C., et al. (2014). "Whole-genome sequencing of cultivated and wild peppers provides insights in *Capsicum* domestication and specialization." Proc. Natl. Acad. Sci. USA **111**(14): 5135-5140.

Ronquist, F. and J. P. Huelsenbeck (2003). "mrBayes 3: Bayesian phylogenetic inference under mixed models." Bioinformatics **19**(12): 1572-1574.

Roossinck, M. J. (1997). "Mechanisms of Plant Virus Evolution." Annu. Rev. Phytopathol. **35**: 191-209.

Roossinck, M. J. (2010). "Lifestyles of plant viruses." Phil. Trans. R. Soc. B **365**: 1899-1905.

Roossinck, M. J. (2012). "Plant virus metagenomics: Biodiversity and ecology." Annu. Rev. Genet. **46**: 357-367.

Roossinck, M. J. (2014). "Metagenomics of plant and fungal viruses reveals an abundance of persistent lifestyles." Frontiers in Microbiology **5**: 767.

Roossinck, M. J. and A. Ali (2007). Mechanisms of plant virus evolution and identification of genetic bottlenecks: Impact on disease management. Biotechnology and Plant Disease Management. Z. K. Punja, S. H. DeBoer and H. Sanfaçon. Wallingford, CABI**:** 109-124.

Roossinck, M. J., et al. (2011). "The remarkable evolutionary history of endornaviruses." J. Gen. Virol. **92**: 2674-2678.

Shirasawa, K., et al. (2013). "Development of *Capsicum* EST-SSR markers for species indentificaiton and *in silico* mapping onto the tomato genome sequence." Mol. Breeding **31**: 101-110.

Tewksbury, J. J., et al. (1999). "In situ conservation of wild chiles and their biotic associates." Conservation Biology **13**(1): 98-107.

Valverde, R. A. and D. L. Gutierrez (2007). "Transmission of a dsRNA in bell pepper and evidence that it consists of the genome of an endornavirus." Virus Genes **35**: 399-403.

Valverde, R. A. and D. L. Gutierrez (2008). "Molecular and biological properties of a putative partitivirus from jalapeño pepper (*Capsicum annuum* L.)." Rev. Mex. Fitopat. **26**(1): 1-6.

Valverde, R. A., et al. (1990). "Indigenous double-stranded RNA from pepper (*Capsicum annuum*)." Plant Science **67**(2): 195-201.

Villanueva, F., et al. (2012). "Complete genome sequence of a double-stranded RNA virus from avocado." J. Virol. **86**(2): 1282-1283.

Wakarchuk, D. A. and R. I. Hamilton (1990). "Partial nucleotide sequence from enigmatic dsRNAs in *Phaseolus vulgaris*." Plant Mol Biol **14**: 637-639.

Walsh, B. M. and S. B. Hoot (2001). "Phylogenetic relationships of *Capsicum* (Solanaceae) using DNA sequences from two noncoding regions: the chlrorplast *atpB-rbcL* spacer region and nuclear *waxy* introns." Int. J. Plant Sci. **162**(6): 1409-1418.

Zabalgogeazcoa, I. A. and F. E. Gildow (1992). "Double-stranded ribonucleic acid in 'Barsoy' barley." Plant Science **83**: 187-194.

Chapter 2

Evolution of two partitiviruses in *Capsicum* spp.

Abstract

Most well studied plant viruses are acute viruses that cause disease in their host. However, plants are very frequently infected with cytoplasmic RNA viruses that persist for many generations through nearly 100% vertical transmission without producing any obvious symptoms.

Peppers are perennial plants, native to the Americas, and as domesticated plants human selection accelerated their evolution, so codivergent timelines should be easier to follow. Jalapeño and Hungarian Wax peppers (*Capsicum annuum*) are infected with *Pepper cryptic virus* 1 (PCV 1) and *Pepper cryptic virus* 2 (PCV 2), respectively. Both viruses belong to the genus *Deltapartitivirus* from *Partitiviridae* family. The evolution of these viruses was investigated by dsRNA extraction from almost one hundred different pepper cultivars/landraces/wild plant materials including *C. annuum*, *C. annuum* var. *glabriusculum, C. chacoense*, *C. chinense*, *C. frutescens*, *C. pubescens* and *C. bacccutum.* The presence of PCV 1 and PCV 2 was tested by RT-PCR using specific primers. The nucleotide sequence of the RT-PCR products was determined and their phylogenies have been analyzed. Here evidence for a remarkably slow evolution rate in both viruses is presented.

Introduction

The most common persistent virus family in plants, and indeed the most common virus family overall, is the *Partitiviridae* (Roossinck 2012, Roossinck 2014). These persistent viruses are found in many crop and ornamental plants, including bean, beet, black raspberry, blueberry,

carrot, cherry, Chinese pear, clover, dill, fig, grape, hops, hemp, mulberry, orchid, peppers, persimmon, pine, primrose, radish, rose, strawberry, and watermelon. Since their infections are asymptomatic they have not been looked for, but are likely present in many other crops as well. These are viruses that Boccardo first reported as cryptic viruses (Boccardo, Lisa et al. 1987). These are isometric viruses with a segmented dsRNA genome (Tang, Ochoa et al. 2010). The family consists of five genera, *Alphapartitivirus*, *Betapartitivirus*, *Cryspovirus*, *Deltapartitivirus*, and *Gammapartitivirus* (Nibert, Ghabrial et al. 2014). The *Partitiviridae* family includes viruses from different hosts: plants, fungi, and protozoan (Nibert, Woods et al. 2009). Members of *Gammapartitivirus* and *Deltapartitivirus* exclusively infect fungi and plants, respectively. Viruses in the genera *Alphapartitivirus* and *Betapartitivirus* infect plants and fungi, and the genus *Cryspovirus* has been reported from protozoa. Based on the phylogenetic analysis of the RdRp, there is no evidence for co-divergence with their host at a deep evolutionary level, and it is suggested that these viruses could be transmitted between plant and fungi (Roossinck 2010). It is also hypothesized that these are fungal viruses and use plants as their vectors (Roossinck 2010). They are very persistent in plants and various treatments could not cure plants of paritiviruses (Szegö, Tóth et al. 2005).

In 1991 several high molecular weight dsRNAs were reported in pepper (Valverde and Fontenot 1991), which are assumed to be the endornavirus and the partitivirus. Then, in 2008 the partial genome sequence of *Pepper cryptic virus 1* (PCV1) was reported from Jalapeño M (Valverde and Gutierrez 2008). And finally, in 2011 Sabanadzovic and Valverde published the complete genome sequence of PCV1 from Jalapeño M and *Pepper cryptic virus* 2 from Hungarian Wax (Sabanadzovic and Valverde 2011). As with all persistent plant viruses, PCV 1 and PCV 2 cannot be transmitted by mechanical inoculation or grafting, while ovule or pollen transmits the virus to over 98 % of the next generation (Valverde and Gutierrez 2008). PCV1 and PCV 2 belong to the genus *Deltapartitivirus* (Nibert, Ghabrial et al. 2014); these two pepper

partitiviruses are phylogenetically closer to other members of this genus than to each other (Sabanadzovic and Valverde 2011). Their genomes consist of two linear, monocistronic dsRNAs that are encapsidated in separate isometric particles about 30 nm in diameter. The genome of PCV 1 is slightly smaller; RNA 1 (1,563 bp) and RNA 2 (1,512 bp) encode the RdRp (479 aa) and the coat protein (CP, 412 aa), respectively (Sabanadzovic and Valverde 2011). The genomic dsRNAs of PCV 2 are 1,609 bp (RNA 1, RdRp 478 aa) and 1,525 bp (RNA 2, CP 430 aa) in size. Their genome organization is shown in Figure 2-1. Using specific primers for each RNA evolution of PCV1 and PCV 2 in different peppers is investigated.

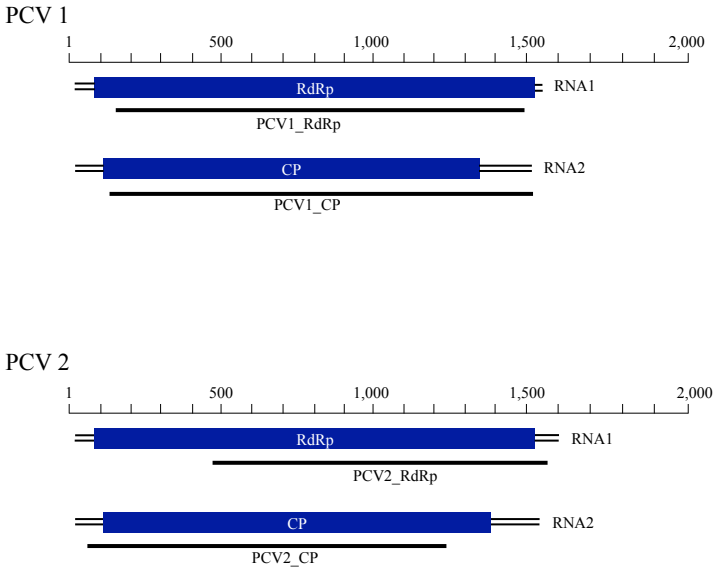

Figure **2-1**: Schematic representation of PCV1 and PCV2 genome organization including the position of amplified fragments.

Materials and Methods

Pepper collection

I collected 97 different peppers including *C. annuum*, *C. chinense*, *C. frutescens*, *C. bacccutum, C. pubescens* and *C. chacoense* with different improvement levels (cultivars/landraces/wild materials). The seed sources and information are found in Table S1. Seeds were germinated in wet paper towels and then transferred to plastic pots containing Sun-Gro Horticulture soil and grown in an insect-free environmental room at 24°C and fluorescent light (16:8 light:dark photoperiod).

Screening *Capsicum* spp. for the presence of PCV 1 and PCV 2

To screen plants for the presence of partitiviruses, dsRNAs were extracted from 5g of fresh weight leaves using the protocol described by Márquez *et al* (Márquez, Redman et al. 2007).

For PCV 1, based on the available sequences in GenBank (JN117276.1 and JN117277.1), primer pairs specific to the RdRp (PCV1_RdRp) and CP (PCV1_CP) were designed to amplify 1,563 bp and 1,512 bp fragments, respectively.

According to the PCV2 sequences in GenBank (JN117278.1 and JN117279.1), specific primers for the RdRp (PCV2_RdRp; 1,097 bp) and CP (PCV2_CP; 1,262 bp) were designed as well. Primer sequences and their positions are provided in Table 2-1 and Figure 2-1. For the RT reaction, used about 2 µg of dsRNA mixed with 2 µM reverse primer, 0.5 mM of Tris-EDTA (pH 8.0) and nuclease-free water to a final volume of 12 µl, boiled for 2 min. The mixture was incubated on ice for 2 min and 8 µl of Reverse Transcriptase (RT) mix [200 U of M-MuLV reverse transcriptase (New England Biolabs), 2 µl of 10X M-MuLV buffer (supplied by the

manufacturer) and 10 mM dNTPs] was added and incubation continued at 42° C for 1 h. Then, cDNA was incubated with 10 µg of boiled RNase A (Sigma) for 15 min at room temperature and cleaned with E.Z.N.A Cycle Pure Kit (Omega Bio-tech) according to the manufactures instruction. About 0.5 µg of cleaned cDNA was used as a template for a 25 µl polymerase chain reaction (PCR) with Taq DNA Polymerase (ThermoFisher Scientific), buffer (30 mM MgCl2, Idaho Technologies), 2 mM dNTPs, and 0.2µM forward and reverse primers for each fragment. The PCR reactions were completed in capillary tubes with an Idaho Technologies Rapid Cycler for 40 cycles (94°C denaturation for 0 s, 48°C annealing for 0 s, and 72°C extension for 45 s). The RT-PCR products were separated on a 1.2% Agarose gel in 0.5X TBE, and appropriate bands were excised and purified using E.Z.N.A.® Gel Extraction Kit (Omega Bio-tech). Sequence analysis of the PCR products was done by the Genomic Core Facility of Pennsylvania State University, University Park, PA. The sequences have been deposited in GenBank under accession numbers listed in Table S6.

Table **2-1**: List of the specific primers for PCV 1 and PCV 2

Primers	Sequence	Position	Size (bp)[1]
PCV1_RdRp			**1,563**
Forward	5' CTCACCGACACCCTCATG 3'	157-173	
Reverse	5' TACCTCTTCTTCTGAAGCCG 3'	1,477-1,496	
PCV1_CP			**1,512**
Forward	5' ACAGTCGTCCCTCACCAAGC 3'	129-148	
Reverse	5' AGCAGGGTGCAATACAG 3'	1,496-1,512	
PCV2_RdRp			
Forward	5' CCATGGACCAAAAGGACCCA 3'	498-518	**1,097**
Reverse	5' GCAGCCACTCCGACTTCAA 3'	1,577-1,595	
PCV2_CP			
Forward	5' TGGCGACACCAGTTAGTGAC 3'	79-99	**1,262**
Reverse	5' TCCGTCTCTTTTCTGAGCGG 3'	1,322-1,341	

[1] size of amplicon

Nucleotide sequences were aligned with Clustal W using default settings in the program Geneious 10.0.9 (Drummond, Ashton et al. 2011). The alignment was visually corrected as necessary. Phylogenetic analysis was performed using MrBayes (Ronquist and Huelsenbeck

2003) implemented as a Geneious 10.1 plug-in. The nucleotide sequences of two PCV 1 isolates, and four PCV 2 isolates available in GenBank (listed in Table S7), and two other members of *Deltapartitivirus* genus (*Pittosporum cryptic virus* 1 (PitCV 1) and *Raphanus sativus cryptic virus* 3 (RsCV3)) were included in the phylogenetic analysis. RsCV3 was used as an outgroup. For both the RdRp and CP phylogenies GTR + G substitution model were selected as the best fit models according to JModelTest (Darriba, Taboada et al. 2012) with 110,000 chain length and 100,000 burn-in.

Results and Discussion

Screening *Capsicum* spp. for the presence of partitiviruses

After dsRNA extraction from peppers, the presence of PCV 1 and PCV 2 was tested by RT-PCR using specific primers for both viruses and the gel extracted RT-PCR products were sent for Sanger sequencing. PCV 1 was detected in one out of 33 different *C. annuum* peppers, in cultivar, Long red cayenne. Three chiltepins (Chiltepin30, Chiltepin 53, Chiltepin54) out of eight accessions of *C. annuum var. glabriusculum* also were positive for PCV 1 (Table 2-2). PCV 2 was detected in ten different peppers; including five of *C. annuum* cultivars: Hinkelhatz hot, Long red cayenne, Neapolitan, Jalapeño criollo and Ember. Diente de perro, which was denoted as *C. frutescens*, and four chiltepins (Chiltepin53, Chiltepin54, Chiltepin55, and Chiltepin102 (Table 2-2). dsRNA profile of samples infected with PCV 1 and PCV 2 is shown in Figure 2-2. Both viruses were not detected in other non-*annuum* species. Chiltepin30, Jalapeno criollo and Diente de perro were collected from Guatemala; Chiltepin53, 54, 55, and 102 from Mexico, and the rest of positive partitiviruses samples were from the USA. The distribution map of samples is shown in Figure 2-3, showing both viruses identified in the same locations.

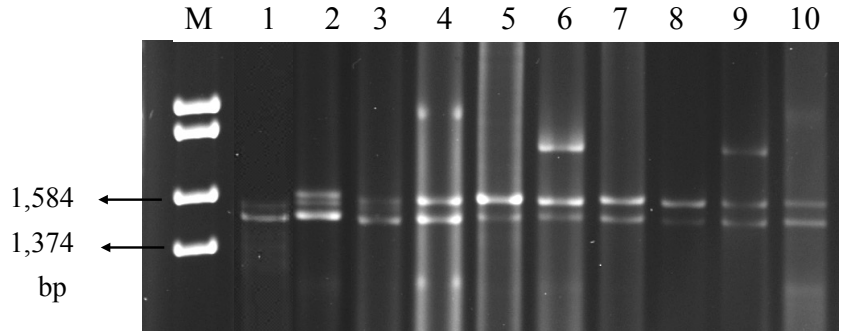

Figure **2-2**: DsRNA profile of 10 representative isolates of partitivirus: 1: Jalapeño (PCV1), 2: Long red cayenne (PCV 1 and PCV 2), 3: Chiltepin54 (PCV1 and PCV2), 4: Hinkelhatz hot (PCV 2), 5: Jalapeño criollo (PCV 2), 6: Chiltepin55 (PCV2), 7: Ember (PCV2), 8: Diente de perro (PCV 2), 9: Chiltepin102 (PCV 2). DsRNA were electrophoresed on 1.2 % agarose gel, using digested lambda DNA with *Eco*RI and *Hind*III as marker (M).

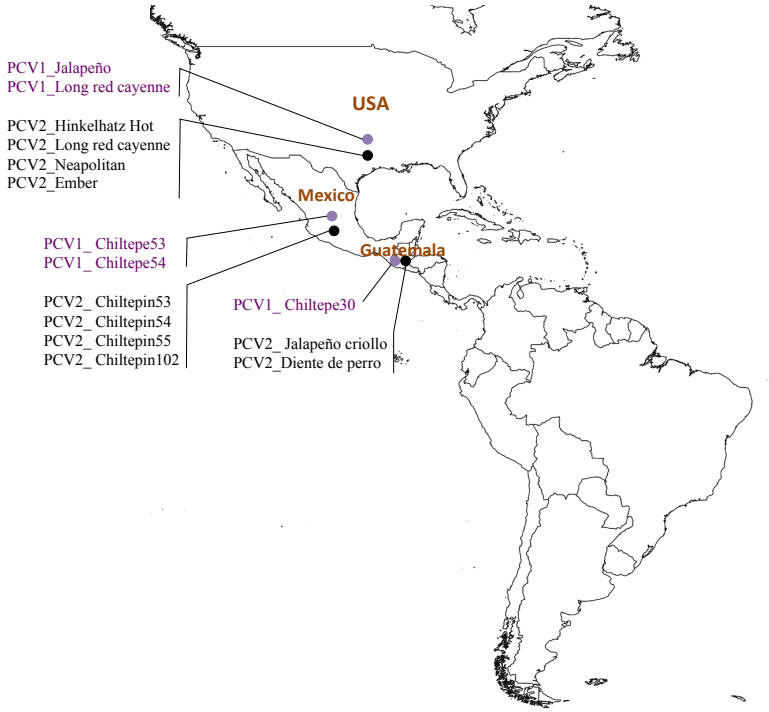

Figure **2-3**: The distribution map of pepper partitiviruses identified in study. Dots reflect the country, and not a specific loaction within the country.

Table **2-2**: List of peppers positive for partitiviruses

Name in Tree[1]	Species	Location	PI[2]	V
PCV1_Jalapeño	*C. annuum*	USA	-	P
PCV1_Long red cayenne	*C. annuum*	USA	-	P
PCV1_Chiltepe30	*C. annuum var. glabriusculum*	Guatemala	PI 632932	P
PCV1_Chiltepe53	*C. annuum var. glabriusculum*	Mexico	-	P
PCV1_ Chiltepe54	*C. annuum var. glabriusculum*	Mexico	-	P
PCV2_Hinkelhatz Hot	*C. annuum*	USA	-	P
PCV2_Long red cayenne	*C. annuum*	USA	-	P
PCV2_Neapolitan	*C. annuum*	USA	-	P
PCV2_Jalapeño criollo	*C. annuum*	Guatemala	PI 666462	P
PCV2_Ember	*C. annuum*	USA	PI 273426	P
PCV2_Chiltepin53	*C. annuum var. glabriusculum*	Mexico	-	P
PCV2_Chiltepin54	*C. annuum var. glabriusculum*	Mexico	-	P
PCV2_Chiltepin55	*C. annuum var. glabriusculum*	Mexico	-	P
PCV2_Chiltepin102	*C. annuum var. glabriusculum*	Mexico	-	P
PCV2_Diente de perro	*C. frutescens*	Guatemala	PI 666589	P

[1] Virus name_cultivar name or plant designation
[2] PI, Plant Identification number

Interestingly PCV 1 and PCV 2 were detected in chiltepin (*C. annuum* var. *glabriusculum*), the wild progenitor of domestic peppers in *C. annuum*, indicating that during domestication viruses were transmitted from chiltepin to other peppers. In addition, all the peppers positive for both partitiviruses were hot peppers.

Virus phylogenies

Pairwise identity in PCV1_RdRp nt sequences was 98-100%, and in PCV1_CP was 97-100%. There is very little variation between Chiltepin isolates and cultivated isolates. Pairwise identity in PCV2 nt sequences was 94-100% in the RdRp, 95-100% in the CP.

A Bayesian phylogenetic tree was constructed using nt sequences of pepper partitivirus RdRp and CP regions, using PdPV as an outgroup (Figure 2-4). As was expected the tree contains two main clades: PCV1 isolates are grouped together, and the second clade contains the PCV2 isolates for both the RdRp and CP phylogenies. PCV1 isolates for both CP and RdRp fragments are closer to PitCV 1 than PCV2; which is in agreement with previous study on these two viruses (Sabanadzovic and Valverde 2011). This indicates that these two pepper partitiviruses do not share an immediate ancestor, and were introduced separately to peppers, even to the same host, and adapted within the hosts.

A. B.

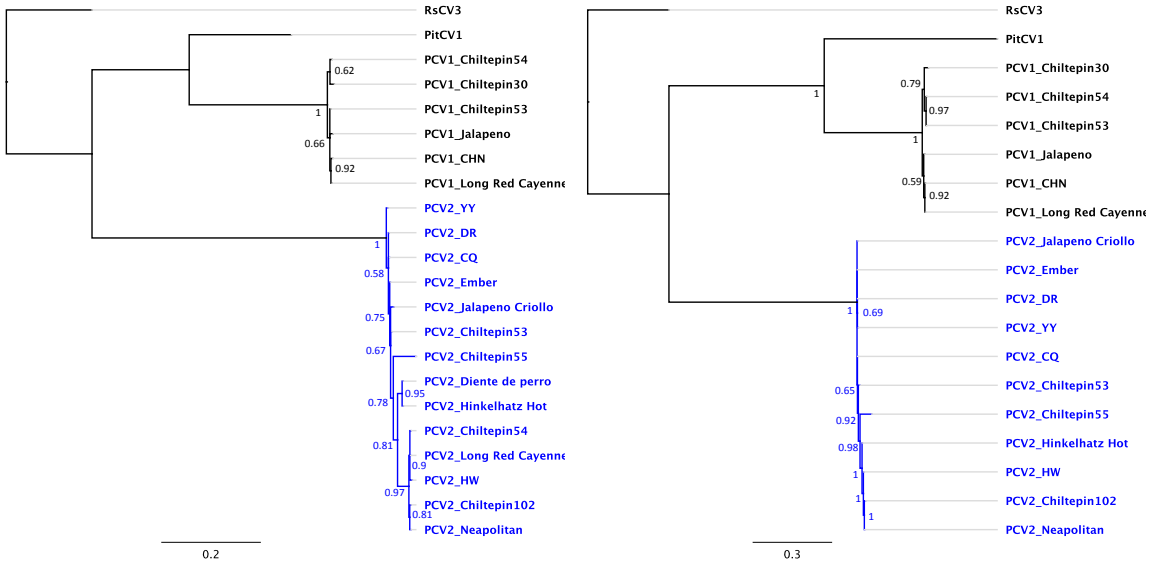

Figure **2-4**: Bayesian analysis of the relationships among aligned nt sequences of the RdRp fragment (**A**) and CP (**B**) of Pepper partitiviruses. RsCV3 was used as an outgroup.

Using the PCV2_CP primers could not amplify the coat protein from Chiltepin54, Long Red Cayenne and Diente de perro isolates. In three samples (Chiltepin53, Chiltepin54, and Long Red Cayenne) the co-infection of PCV 1 and PCV 2 were observed. In addition, co-infection of

PCV 2 and endornaviruses also observed in Neapolitan and Ember peppers. Co-infection of *Lentinula edodes partitivirus* 1 (LePV1) and *L. edodes mycovirus* HKB (LeV-HKB), two unrelated virus has been reported in edible fungus *Lentinula edodes*. This co-infection led to 18.5-fold higher expression of the RdRp to CP ration in LeVP 1, while in the single infection of LeVP 1 the ratio of the RdRp to CP was 3.3-fold higher. I have not investigated how the co-infection of PCV 1 and PCV 2 might affect their replication; however, that could be the reason why the PCV2_CP in Chiltepin 54, and Long red cayenne isolates were not amplified (Guo, Bian et al. 2017).

In PCV 1 and PCV 2 there are less than 3% and 6% variation (in both CP and RdRp) between isolates from chiltepin and domesticated peppers. *C. annuum* was domesticated in Mexico from chiltepin (*C. annuum* var. *glabriusculum*) around 10,000 years ago (González-Jara, Moreno-Letelier et al. 2011, Kraft, Brown et al. 2014, Qin, Yu et al. 2014). Therefore, the evolution rate in PCV 1 and PCV 2 would be 3×10^{-6} and 6×10^{-6} changes/nt/yr, which are remarkably slow rates of evolution for RNA viruses. There are some other examples of a slow evolution rate in RNA viruses. In rice, the persistent virus *Oryza sativa endornavirus* is found in all japonica cultivars, but not in indica cultivars (Horiuchi, Moriyama et al. 2003). *Oryza rufipogon* is the ancestor of domesticated japonica rice and has a related persistent virus, *Oryza rufipogon endornavirus*. Around 10,000 years ago, the hosts of these two viruses diverged during the cultivation of rice (Molina, Sikora et al. 2011). Therefore, assuming the virus was already present at the time of divergence, these viruses have diverged by about 24% during 10,000 years (24×10^{-5} changes/nt/yr). In addition, *Zea mays* chrysovirus 1 has been isolated from ancient corn (~1000 years old sample) and also modern corn, with just 3% variation between the isolates, which would be 3×10^{-4} changes/nt/yr in this virus (Peyambari et al 2017 Unpublished). These rates are all surprisingly slow for RNA viruses with previously reported rate of 10^{-3} to 10^{-4} substitution/site/year (Roossinck 1997, Roossinck and Ali 2007, Pagán and Holmes 2010).

Hence, variation in plant persistent viruses is limited. This might be because of strong purifying selection that with very long virus-host relationships the sequence is fine-tuned and does not tolerate changes (implies mutualism). It could be also explained by replication strategies of these viruses. Most persistent viruses have dsRNA genomes (except endornaviruses), and they replicate by stamping mode of replication, as explained in Chapter 3.

However, *Pseudogymnoascus destruct*ans partitivirus-pa (PdPV-pa) is a mycovirus from *Gammapartitivirus* genus isolated from causal agent of white nose syndrome (*Pseudogymnoascus destruct*ans) in bats (Thapa, Turner et al. 2016). Variation in this virus during 10 years shows a rate of 0.03 changes/nt/yr, four orders of magnitude greater than the variation in pepper partitiviruses. It could be because of recent introduction of this virus to *P. destruct*ans, or it could have been related to their host kingdom. Mycoviruses can be transmitted through hyphal anastomosis between vegetatively compatible strains of the same fungi (Ghabrial 1998). However, there is no evidence of horizontal transmission in plant persistent viruses (Roossinck 2010).

Pepper phylogenies and population structure

To examine the population structure in the *Capsicum* spp. using SNPs data from the KASP analysis Bayesian clustering was implemented in STRUCTURE software. The distribution of the 38 *Capsicum* spp. lines into three distinct clusters is shown in Figure 2-5 (red=cluster 1; green=cluster 2; blue=cluster 3). In cluster 1, Long red cayenne (infected with both PCV 1 and PCV 2), Chiltepin 102 (infected with PCV 2), and Diente de perro (infected PCV2) are clustered together. Jalapeño (PCV 1), Chiltepin53 (PCV1 and PCV2), Chiltepin54 (PCV1 and PCV2), and Jalapeño criollo (PCV2) are grouped together in cluster 2. Cluster 3 includes Chiltepin30, which was infected with PCV 1 and the other species of non-*annuum* species of *Capsicum*.

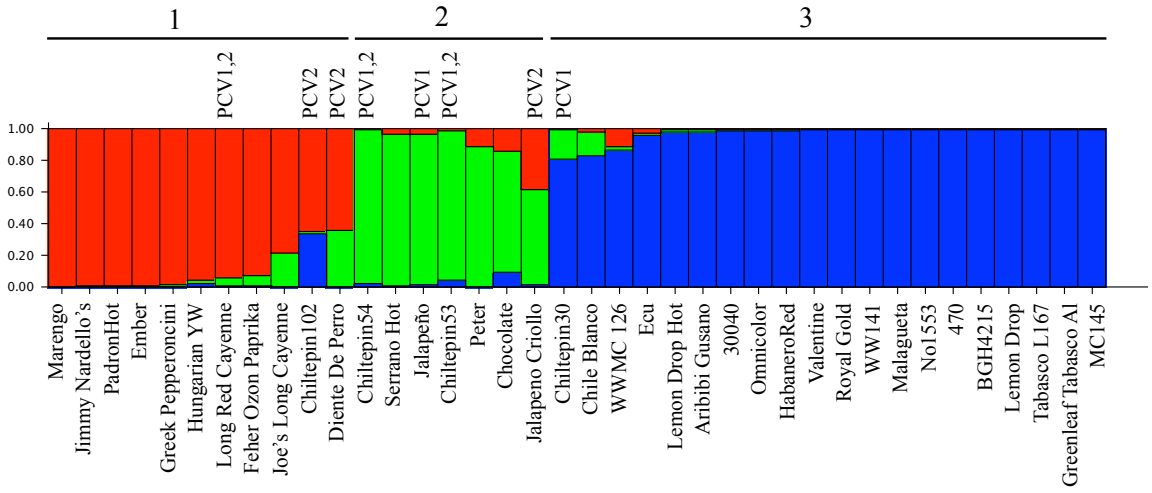

Figure **2-5.** Population structure analysis of SNPs data using STRUCTURE software (2.3.4).

Population structure analysis revealed admixture in Joe's long cayenne and Diente de perro from cluster 1 with peppers in cluster 2; Jalapeño criollo peppers have admixtures with genotypes in cluster 1. Peppers in cluster 1 and 2 belong to *C. annuum*, expect Diente de perro, which has been identified as *C. frutescens* (based on its morphological characteristics), but this results suggested it belongs to *C. annuum*. It is consistence with partitivirus distribution, which was limited to wild and cultivated *C. annuum*, and Diente de perro was the only non-*annuum* pepper positive for PCV 2. Pepper lines in cluster 3 corresponded to peppers positive for CFEV 1 except Chiltepin30, which showed admixed genome (<20%) with cluster 2 (For more details refer to Chapter 1).

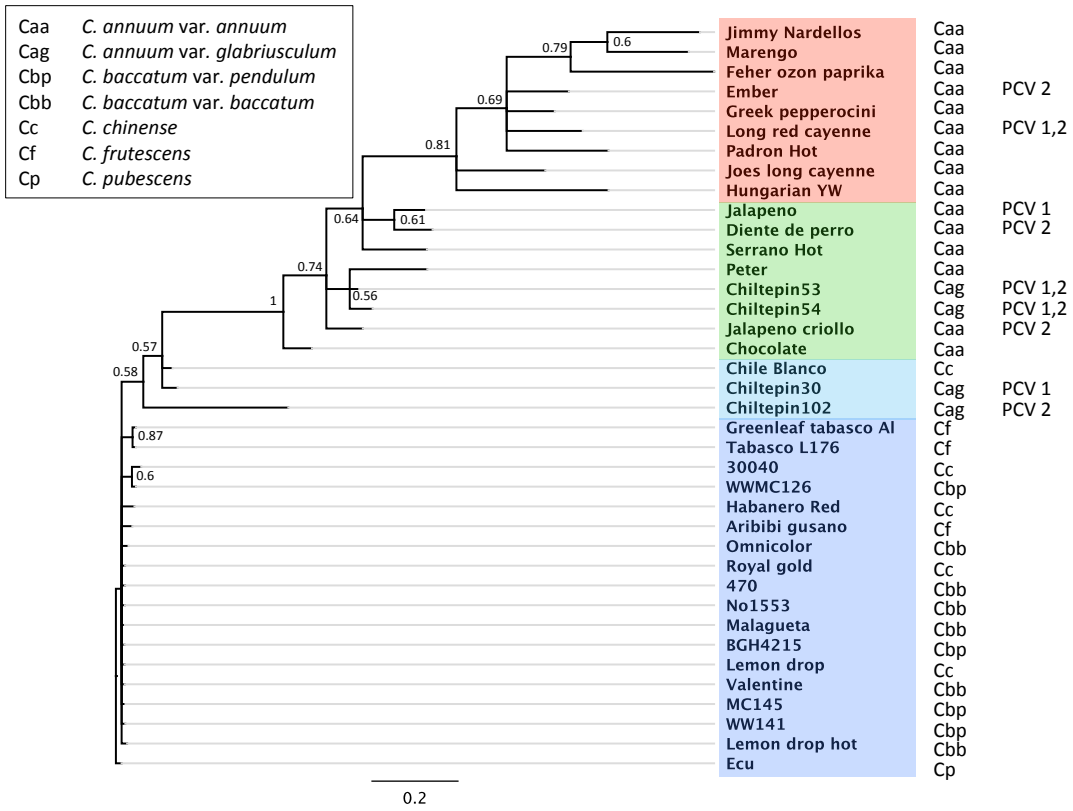

Figure **2-6**: Bayesian analysis of the relationships among SNPs in different pepper lines.

The Baysian phylogenetic tree for the 38 pepper samples was constructed using the MrBayes plug-in in Geneious 10.0.9. Ecu pepper (*C. pubescens*) was used as an outgroup. Pepper partitivirus hosts have been distributed to three highlighted areas (Red, green, and light blue), or *annuum* clades; and Chiltepin102 is probably their ancestor (Figure 2-6). The final grouping includes the various non-*annuum* species of *Capsicum*, where PCV1 and PCV 2 were not detected in any of them. This data is in agreement with previous studies that showed Chiltepin is the ancestor of cultivated *C. annuum* peppers (Kraft, Brown et al. 2014, García, Barfuss et al. 2016). So, it confirmed that PCV 1 and PCV 2 have been introduced from chiltepin to cultivated peppers.

Cited Literature

Boccardo, G., et al. (1987). "Cryptic plant viruses." Adv. Virus Res. **32**: 171-214.

Darriba, D., et al. (2012). "jModelTest 2: more models, new heuristics and parallel computing." Nat. Methods **9**(8): 772.

Drummond, A. J., et al. (2011). Geneious.

García, C. C., et al. (2016). "Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum*, Solanaceae)." Annals of Botany **118**: 35-51.

Ghabrial, S. A. (1998). "Origin, adaptation and evolutionary pathways of fungal viruses." Virus Genes **16**: 119-131.

González-Jara, P., et al. (2011). "Impact of human management on the genetic variation of wild pepper, *Capsicum annuum* var. *glabriusculum*." PLoS ONE **6**(12): e28715.

Guo, M., et al. (2017). "Biological and molecular chracteristics of a novle partitivirus infecting the edilbe fungus *Lentinula edodes*." Plant Dis. **101**: 726-733.

Horiuchi, H., et al. (2003). "Inheritance of *Oryza* sativa endornavirus in F1 and F2 hybrids betwween *japonica* and *indica* rice." Genes Genet Syst **78**: 229-234.

Kraft, K. H., et al. (2014). "Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico." Proc. Natl. Acad. Sci. USA **111**(17): 6165-6170.

Márquez, L. M., et al. (2007). "A virus in a fungus in a plant --three way symbiosis required for thermal tolerance." Science **315**: 513-515.

Molina, J., et al. (2011). "Molecular evidence for a single evolutionary origin of domesticated rice." Proc. Natl. Acad. Sci. USA **108**(20): 8351-8356.

Nibert, M. L., et al. (2014). "Taxonomic reorganization of family *Partitviridae* and other recent progress in partitivirus research." Vir. Res. **188**: 128-141.

Nibert, M. L., et al. (2009). "Cryspovirus: a new genus of protozoan viruses in the family Partitiviridae." Arch. Virol. **154**: 1959-1965.

Pagán, I. and E. C. Holmes (2010). "Long-term evolution of the *Luteoviridae*: time scale and mode of virus speciation." J. Virol. **84**(12): 6177-6187.

Qin, C., et al. (2014). "Whole-genome sequencing of cultivated and wild peppers provides insights in *Capsicum* domestication and specialization." Proc. Natl. Acad. Sci. USA **111**(14): 5135-5140.

Ronquist, F. and J. P. Huelsenbeck (2003). "mrBayes 3: Bayesian phylogenetic inference under mixed models." Bioinformatics **19**(12): 1572-1574.

Roossinck, M. J. (1997). "Mechanisms of Plant Virus Evolution." Annu. Rev. Phytopathol. **35**: 191-209.

Roossinck, M. J. (2010). "Lifestyles of plant viruses." Phil. Trans. R. Soc. B **365**: 1899-1905.

Roossinck, M. J. (2012). "Plant virus metagenomics: Biodiversity and ecology." Annu. Rev. Genet. **46**: 357-367.

Roossinck, M. J. (2014). "Metagenomics of plant and fungal viruses reveals an abundance of persistent lifestyles." Frontiers in Microbiology **5**: 767.

Roossinck, M. J. and A. Ali (2007). Mechanisms of plant virus evolution and identification of genetic bottlenecks: Impact on disease management. Biotechnology and Plant Disease Management. Z. K. Punja, S. H. DeBoer and H. Sanfaçon. Wallingford, CABI**:** 109-124.

Sabanadzovic, S. and R. A. Valverde (2011). "Properties and detection of two cryptoviruses from pepper (*Capsicum annuum*)." VIrus Genes **43**: 307-312.

Szegö, A., et al. (2005). "Detection of high molecular weight dsRNA persisting in *Dianthus* species." Acta Biologica Szegediensis **49**(1-2): 17-19.

Tang, J., et al. (2010). "Structure of Fusarium poae virus 1 shows conserved and variable elements of partitivirus capsids and evolutionary relationships to picovirnavirus." J. Struct. Biol. **172**: 363-371.

Thapa, V., et al. (2016). "Using a novel partitivirus in *Pseudogymnoascus destructans* to understand the epidemiology of white-nose syndrome." PLoS Pathog. **12**(12): e1006076.

Valverde, R. A. and J. F. Fontenot (1991). "Variation in double-stranded ribonucleic acid among pepper cultivars." J. Amer. Soc. Hort. Sci. **116**(5): 903-905.

Valverde, R. A. and D. L. Gutierrez (2008). "Molecular and biological properties of a putative partitivirus from jalapeño pepper (*Capsicum annuum* L.)." Rev. Mex. Fitopat. **26**(1): 1-6.

Chapter 3

How does the genome structure and lifestyle of a virus affect its population variation?

Reproduced from Maliheh Safari, Marilyn J Roossinck, How does the genome structure and lifestyle of a virus affect its population variation?, In Current Opinion in Virology, Volume 9, 2014, Pages 39-44, ISSN 1879-6257, https://doi.org/10.1016/j.coviro.2014.09.004.

Abstract

Viruses use diverse strategies for their replication, related in part to the genome structure (double-stranded or single-stranded; positive sense or negative sense; RNA or DNA). During each round of replication, mutations are introduced in the viral genome and the mode of replication (stamping machine and geometric replication) may affect the population dynamics of the progeny virus. Our understanding of the relationships among genome strandedness, mode of replication and the population variation is still limited. Here we will review what is known about virus replication by stamping machine or geometric modes, and how that relates to the biology of single stranded versus double stranded RNA genomes. We will present how this may affect the mutation frequency and population dynamics. Finally the potential importance of the population dynamics in acute viruses and persistent viruses will be discussed.

Introduction

Error-prone replication, large population sizes, and rapid replication lead to large clouds of mutants in RNA viruses (Holland, Spindler et al. 1982), and RNA viruses are often characterized by high levels of genetic variation known as quasispecies. The quasispecies nature

of viruses leads to many biological consequences. Having a large mutant cloud size provides an accessible pool of mutants that could benefit the virus during adaptation to a new environment or escape from the host defenses, sometimes resulting in emerging viral diseases (Domingo, Biebricher et al. 2001). Conversely, highly diverse populations subjected to repeated bottlenecks can lose fitness through a process known as Muller's ratchet (Muller 1964), as demonstrated in a number of studies with *Vesicular stomatitis virus* (Holland, delaTorre et al. 1991, Clarke, Duarte et al. 1993, Novella, Duarte et al. 1995, Novella, Elena et al. 1995).

The viral RNA dependent RNA polymerases (RdRps) are thought to have low fidelity due to a lack of proofreading (Steinhauer, Domingo et al. 1992). DNA dependent DNA polymerases require basepairing of a primer to initiate polymerization, but this is not required for transcription initiation in RNA polymerases, and RNA viruses may be more relaxed in this level of fidelity as well (Roossinck 1997). While the replicase fidelity determines the mutation rate of viruses, their genetic variation also is governed by natural selection and genetic bottlenecks. Here we use "mutation rate" to refer to how often the polymerase makes a mistake, while "mutation frequency" refers to the accumulation of mutations after (unknown) rounds of replication, bottlenecks and selective events (Roossinck 1997). Schneider and Roossinck (Schneider and Roossinck 2001) showed that mutation frequency is host associated, and the population variation of *Cucumber mosaic virus* (CMV) in pepper is higher than that in tobacco. Additional studies show the role of host factors in virus replication (Li, Pogany et al. 2010, Li, Wei et al. 2013). The impact of translation elongation factors in RNA virus replication was reported first in bacteriophage Qβ by Blumenthal et al. (Blumenthal, Young et al. 1976). Later, the effect of the translation elongation factor 1A (eEF1A) in the synthesis of *Tombusvirus* negative strand was demonstrated (Li, Pogany et al. 2010). Some single stranded (ss) DNA viruses, which use a host polymerase, have similar levels of mutation frequency as RNA viruses (Isnard, Granier et al. 1998, Ge, Zhang et al. 2007). Thus, the difference between RNA and DNA virus population

variation does not necessarily reflect what is known about the polymerase fidelity. Beside the error rate of viral polymerase, there are other factors affecting the frequency of mutant viruses during an infectious process, including the amplification dynamics of RNA positive and negative strands and purifying mechanisms acting after the transcription. Here, we focus on the dynamics of different RNA strands. We will briefly explain the replication cycle in RNA viruses with different genome structures and their mode of replication, and discuss how the dynamics of positive and negative strand RNA can affect the mutation frequency, and how all these factor are related to the lifestyle of viruses.

Replication cycle in RNA viruses

Viral genomes are either RNA or DNA, which can be double stranded (ds) or single stranded (ss). Based on the polarity of their genomic RNA, ssRNA viruses are classified into positive (+) and negative (-) sense viruses.

In (+) ssRNA virus, once the virus enters the cytoplasm of an infected host cell, it is uncoated (Figure 3-1A), and then becomes immediately available for translation as an mRNA. During translation the structural and non-structural proteins, including RdRp and other required proteins for virus replication, are produced. The third step is transcription; the (-) strands (antigenomic strands) are copied from the genomic strand. These (-) strands are used as templates for (+) RNA synthesis as progeny genomes or amplified mRNAs. The replication process in (+) ssRNA viruses is usually asymmetric, and a large excess of positive over negative strands is produced (Ishikawa, Kroner et al. 1991, Ahlquist 1992, Buck 1999, Sanfacon 2005). Finally, in most cases genomes are encapsidated to form progeny virions that may acquire a lipid envelope upon exiting the cell.

For (-) ssRNA viruses after entering the host cells and uncoating, the next step is transcription to (+) strands (Figure 3-1B). Transcription is performed by the viral RdRp, which is packaged in the incoming virion. Following translation of structural proteins and viral proteins involved in replication the (-) strand RNA progenies are copied from the (+) strand RNAs and in the encapsidation step mature virions are produced that contain at least one copy of the RdRp (Jackson, Dietzgen et al. 2005, Ahlquist 2006, King, Adams et al. 2012). Details about (-) RNA virus polymerases have only recently been elucidated, and have some unique features, including the use of a single very large protein rather than a protein complex for many of these viruses (Kranzusch and Whelan 2012), although notably influenza virus uses a protein complex. The dsRNA virus replication cycle is significantly different from that of the ssRNA viruses (Figure 3-1C). The virion is delivered into the cytoplasm, and the (+) sense RNAs are transcribed by a viral RdRp in the virion, and extruded from the virion into the cytoplasm. Translation of proteins is followed by packaging each (+) ssRNA to form a new immature particle. Then, each (+) RNA serves as a template for the synthesis of a (-) strand to produce a mature dsRNA progeny virion. The unique feature in this replication is that the dsRNA genomes are transcribed within the viral particles and never exposed directly to the cytoplasm (Patton and Spencert 2000, Nibert and Schiff 2001, Ahlquist 2006).

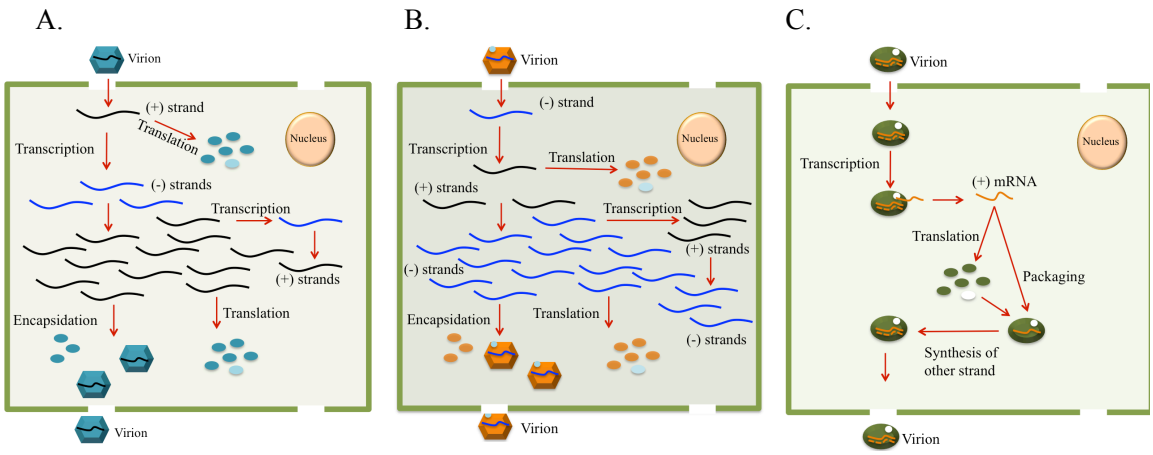

Figure **3-1**: Life strategies for RNA viruses with different genome types (A) (+) ssRNA virus life cycle. (B) (-) ssRNA virus life cycle. (C) Life cycle for dsRNA viruses.

https://doi.org/10.1016/j.coviro.2014.09.004

Mode of Replication

Based on the dynamics of strand synthesis, two modes of replication have been described in RNA viruses: stamping machine and geometric replication.

In geometric, or exponential replication, genomic (+) strands produce multiple (-) strands that are used as templates to produce multiple progeny (+) strands (Figure 3-2A). The key feature is that large numbers of complementary RNAs can be produced at each step, and multiple rounds of (+) to (-) to (+) strand synthesis can occur, leading to an exponential growth of the virus in the cell (Luria 1951, Sardanyés, Solé et al. 2009, Thébaud, Chadœuf et al. 2010). This replication strategy maximizes the intracellular growth rate of the virus (Thébaud, Chadœuf et al. 2010). A similar replication strategy can occur for (-) RNA viruses, although there is considerable variation in replications strategies for these viruses, and in some cases they may use a strategy more like dsRNA viruses. Some studies have shown that for (+) ssRNA viruses the number of (+) strand

copies is much higher than (-) strands (Ishikawa, Kroner et al. 1991, Ahlquist 1992, Sanfacon 2005, Ahlquist 2006), suggesting that actual replication may be semi-geometric (Roossinck 2008).

For the stamping machine scenario, consider the typical replication cycle of a dsRNA virus (Figure 3-2B) as an example. The (+) RNA is synthesized in the parental virion and released into the cytoplasm, translated to produce proteins, and encapsidated. Subsequently, the (-) strand is copied to form the progeny dsRNA virus in the premature particle. Therefore, the parental genome is the only template used for the production of the progeny. This mode of replication was first proposed by Luria in 1951 (Luria 1951). Analyzing the distribution of spontaneous mutations of bacteriophage phi6 demonstrated that it replicates mostly by the stamping machine mode of replication (Chao, Rang et al. 2002). French and Stenger (French and Stenger 2003) suggested that *Wheat streak mosaic virus* [(+) ssRNA] replicates by the stamping machine mode. In addition, this mode of replication has been reported in phage phiX174 and Qβ (Denhardt and Silver 1966, Eigen, Biebricher et al. 1991). On the other hand, the geometric mode of replication is believed to be dominant in DNA phage T2 (Luria 1951). Martinez et al. (Martínez, Sardanyés et al. 2011) investigated the dynamics of (-) and (+) RNA of *Turnip mosaic virus* in the protoplast of the *Nicotiana benthamiana*, and concluded that the mode of replication in this virus is a mixed strategy of geometric and stamping machine, but 90% of the produced genomes were derived from the stamping machine mode. No studies have clearly demonstrated either mode of replication in (-) ssRNA viruses, but the ratio of (-) to (+) strands is very high, and no double-stranded replicative intermediates have been found for these viruses, which could support a stamping machine model.

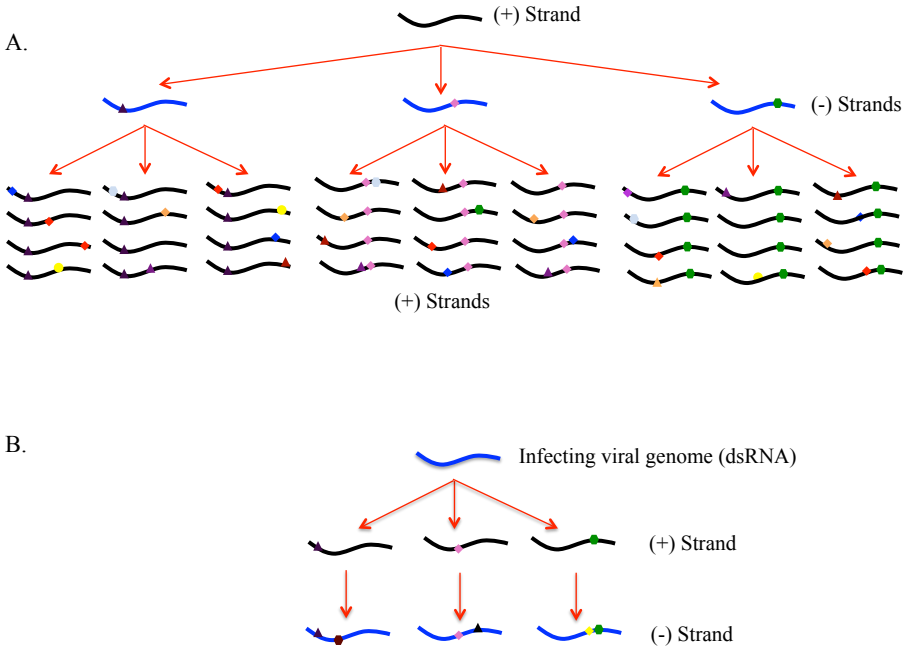

Figure **3-2**: Modes of replication for RNA viruses (A) Geometric mode of replication in (+) ssRNA virus. Each (+) strand becomes a template to repeat this cycle. Mutations may accumulate more rapidly in this mode of replication. (B) Stamping machine mode of replication in dsRNA viruses, and the accumulation of mutations in the genome.

https://doi.org/10.1016/j.coviro.2014.09.004

Impact of replication mode on viral population variation

Mutation frequency is the accumulation of mutations after multiple rounds of replication, and thus, is affected by the mode of replication. In the stamping machine mode, multiple copies are made sequentially from the same template, while the progenies do not become templates until they infect another cell. During the synthesis of RNA from the parental genome, some mutations may be introduced that are inherited in the progeny (Figure 3-2B). Since there is only one cycle of strand copying per infected cell, one would expect lower mutation frequencies in the viral populations. On the other hand, in the geometric mode of replication, mutations are introduced

and propagated at each half round of replication. Since there are several cycles of strand copying per infected cell, greater mutation frequencies may be expected in the population (Figure 3-2A). However, other forces such as natural selection, bottlenecks and reassortment act upon the mutant spectrum, and limit the quasispecies diversity (Roossinck 1997).

Schneider and Roossinck (Schneider and Roossinck 2000) investigated the quasispecies variation in three (+) ssRNA viruses: CMV, *Tobacco mosaic virus* (TMV) and *Cowpea chlorotic mottle virus* (CCMV). The population diversity levels were analyzed over ten consecutive passages in a common host. While the mutation frequency level of each virus was different, the levels were relatively constant. The viral populations rapidly reached a level of diversity in the initially inoculated plant, which was maintained over the course of passaging. Furthermore, the level of diversity in each of these viruses was correlated to their host range ranges: CMV showed the highest, TMV the intermediate, and CCMV had no measurable level of diversity (Schneider and Roossinck 2000). In a subsequent study inoculation of CMV and TMV on a number of hosts revealed that quasispecies cloud size is not constant in different hosts. The population cloud size in a given host remained constant, while shifting the virus between hosts changed the cloud size to the level associated with the new host (Schneider and Roossinck 2001). In a specific host the viral replicase may make fewer or more errors depending on host factors involved in replication, such as host proteins, nucleotide concentrations, and pH or ionic strength in a specific environment. In one of the only direct measurements of polymerase fidelity in an intact host, the rate of indel mutations in CMV was significantly higher in pepper *versus* tobacco (Pita, deMiranda et al. 2007). On the other hand, mutation accumulation may be limited by the purifying mechanisms of a particular host. The levels of diversity in protoplasts were much higher than in intact plants (Schneider and Roossinck 2001). The CMV domains that control genetic variation are controlled by RNAs 1 and 2 that encode the viral replicase proteins (Pita and Roossinck 2013). This implies that the virus replicates by the geometric mode and generates rapid

diversity in its population, but then other factors determine the final population variation in the intact plants (Schneider and Roossinck 2001). For example, genetic bottlenecks during systemic movement may limit the amount of variation (Li and Roossinck 2004, Ali and Roossinck 2010).

Virus lifestyle in plant viruses

Most well studied plant viruses are acute viruses that cause disease in their host by rapid replication. These viruses can infect hosts systemically and are transmitted vertically and horizontally. While the ssRNA genome is the most common genome in acute viruses, viruses with other genome structures are described in this group. In plants, infections by acute viruses can be resolved by recovery, death, or conversion to chronic infection. In contrast, plant persistent viruses replicate in their hosts for many generations, perhaps thousands of years. They do not cause obvious disease. The name "persistent" comes form their lifestyle in the host (Roossinck 2010, Roossinck 2012, Roossinck 2013) and it is different from persistently transmitted viruses (Gray and Banerjee 1999), which refers to their transmission by vector. Plant persistent viruses are distributed to all host cells through host cell division. Neither movement between plant cells, nor transmission through grafting has been observed in these viruses. Hence, there is no horizontal transfer in these viruses, though they are vertically transmitted via the gametes to seeds at nearly 100% (Boccardo, Lisa et al. 1987, Blanc 2007, Roossinck 2010). So far, four families of RNA viruses are reported in this group: Partitiviridae, Endornaviridae, Totiviridae, Chrysoviridae. Only viruses in Endornaviridae family have ssRNA genome, the remainders have dsRNA. Fungi are hosts for viruses in each of these families as well, and they seem to have similar persistent lifestyles in fungi.

The roles of persistent plant viruses have not been studied thoroughly. Their very long-term relationships with their hosts, and their strict vertical transmission suggest beneficial

interactions. One study has shown that coat protein gene of the White clover cryptic virus 2 affects the nodulation regulation in clover and in a model legume, Lotus japonicus (Nakatsukasa-Akune, Yamashita et al. 2005). The discovery of expressed integrated persistent virus sequences in plant genomes is further evidence of the importance of these viruses (Liu, Fu et al. 2010, Chiba, Kondo et al. 2011).

Several quasispecies population surveys have been done on acute viruses in both plants and animals; however, there is no study of the population of plant persistent viruses. In Curvularia thermal tolerance virus (CThTV), a dsRNA persistent virus in fungi (Márquez, Redman et al. 2007), no detectable diversity could be found in the population (Roossinck 2010). This may be explained by its genome type, that is most likely replicated via a stamping machine mode, or by its persistent lifestyle, that may have different forces limiting variation.

It is probable that plant persistent viruses have low diversity in their population. Many of these viruses contain the dsRNA genomes resulting in the stamping machine mode of replication. Their lifestyle does not include many of the potential advantages of large quasispecies clouds. They do not move to new hosts (or at least very rarely) and they may have no need to evade the host adaptive immune system in plants, RNA silencing, as they are found in meristem where silencing eliminates most acute viruses (Martín-Hernández and Baulcombe 2008, Roossinck 2012).

Conclusions

While some RNA viruses clearly have highly variable populations or quasispecies, the genome type and lifestyle may have important impacts on mutation rates and frequencies. Viruses with dsRNA genomes likely all replicate by a stamping machine mode, which would limit the generation of variation in the populations, in contrast to geometic replication. However,

population studies of dsRNA viruses are lacking. Differences in lifestyle may also affect the population diversity of a virus. A virus that has very long association with a single host, rather than undergoing frequent species jumping, may be at an advantage with a narrow population. Population variation affects virulence in acute viruses. In CMV a less pathogenic isolate has lower mutation frequencies (Pita and Roossinck 2013). A similar relationship is found poliovirus (Vignuzzi, Stone et al. 2006); though, in West Nile virus an opposite association has been reported (Jerzak, Bernard et al. 2007).

Cited Literature

Ahlquist, P. (1992). "Bromovirus RNA replication and transcription." Curr. Op. Genet. Develop. **2**: 71-76.

Ahlquist, P. (2006). "Parallels, among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses." Nat. Rev. Microbiol. **4**(5): 371-381.

Ali, A. and M. J. Roossinck (2010). "Genetic bottlenecks during systemic movement of *Cucumber mosaic virus* vary in different host plants." Virology **404**: 279-283.

Blanc, S. (2007). Virus transmission--getting in and out. Viral Transport in Plants. E. Waigman and M. Heinlein. Berlin Heidelberg, Springer-Verlag**:** 1-28.

Blumenthal, T., et al. (1976). "Function and structure in phage Qß RNA replicase." J. Biol. Chem. **251**(9): 2740-2743.

Boccardo, G., et al. (1987). "Cryptic plant viruses." Adv. Virus Res. **32**: 171-214.

Buck, K. W. (1999). "Replication of tobacco mosaic virus RNA." Phil. Trans. R. Soc. B **354**: 613-627.

Chao, L., et al. (2002). "Distribution of spontaneous mutants and inferences about the replication mode of the RNA bacteriophage ø6." J. Virol. **76**(7): 3276-3281.

Chiba, S., et al. (2011). "Widespread endogenization of genome sequences of non-retroviral RNA viruses into plant genomes." PLoS Pathogens **7**(7): e1002146.

Clarke, D., et al. (1993). "Genetic bottleneck and population passages cause profound fitness differences in RNA viruses." J. Virol. **67**: 222-228.

Denhardt, D. T. and R. B. Silver (1966). "An analysis of the clone size distribution of øX174 mutants and recombinants." Virology **30**: 10-19.

Domingo, E., et al. (2001). Quasispecies and RNA Virus Evolutuion, Principles and Consequences. Austin, Landes Bioscience.

Eigen, M., et al. (1991). "The hypercycle. Coupling of RNA and protein biosynthesis in the infection cycle of an RNA bacteriophage." Biochem. **30**(46): 11005-11018.

French, R. and D. C. Stenger (2003). "Evolution of Wheat streak mosaic virus: Dynamics of population growth within plants may explain limited variation." Annu. Rev. Phytopathol. **41**: 199-214.

Ge, L., et al. (2007). "Genetic structure and population variability of tomato yellow leaf curl China virus." J. Virol. **81**(11): 5902-5907.

Gray, S. M. and N. Banerjee (1999). "Mechanisms of arthropod transmission of plant and animal viruses." Microbiol. and Mol. Biol. Reviews **63**(1): 128-148.

Holland, J., et al. (1982). "Rapid evolution of RNA genomes." Science **215**: 1577-1585.

Holland, J. J., et al. (1991). "Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses." J. Virol. **65**(6): 2960-2967.

Ishikawa, M., et al. (1991). "Biological activites of hybrid RNAs generated by 3'-end exchanges between tobacco mosaic virus and brome mosaic viruses." Journal of VIrology **65**(7): 3451-3459.

Isnard, M., et al. (1998). "Quasispecies nature of three maize streak virus isolates obtained through different modes of selection from a population used to assess response to infection of maize cultivars." J. Gen. Virol. **79**: 3091-3099.

Jackson, A. O., et al. (2005). "Biology of plant rhabdoviruses." Annu. Rev. Phytopathol. **45**: 623-660.

Jerzak, G. V. S., et al. (2007). "The West Nile virus mutant spectrum is host-dependant and a determinant of mortality in mice." Virology **360**(2): 469-476.

King, A. M. Q., et al., Eds. (2012). Virus Taxonomy Ninth Report of the International Committee On Taxonomy of Viruses. San Diego, CA, Elsevier Academic Press.

Kranzusch, P. J. and S. P. J. Whelan (2012). "Architecture and regulation of negative-strand viral enzymatic machinery." RNA Biology **9**(7): 941-948.

Li, D., et al. (2013). "The unexpected roles of eukaryotic translation elongation factors in RNA virus replication and pathogenesis." Microbiol. Mol. Biol. Rev. **77**(2): 253-266.

Li, H. and M. J. Roossinck (2004). "Genetic bottlenecks reduce population variation in an experimental RNA virus population." J. Virol. **78**(19): 10582-10587.

Li, Z., et al. (2010) Translation elongation factor 1A facilitates the assembly of the tombusvirus replicase and stimulates minus-strand synthesis. PLoS Pathogens **6**, 1-14

Liu, H., et al. (2010). "Widespread horizontal gene transfer from double-stranded RNA viruses to eukaryotic nuclear genomes." J. Virol. **84**(22): 11879-11887.

Luria, S. E. (1951). "The frequency distribution of spontaneous bacteriophage mutants as evidence of the exponential rate of reproduction." Cold Spring Harbor Symp. Quant. Biol. **16**: 463-470.

Márquez, L. M., et al. (2007). "A virus in a fungus in a plant --three way symbiosis required for thermal tolerance." Science **315**: 513-515.

Martín-Hernández, A. M. and D. C. Baulcombe (2008). "Tobacco rattle virus 16-kilodalton protein encodes a suppressor of RNA silencing that allows transient viral entry in meristems." J Virol **82**(8): 4064-4071.

Martínez, F., et al. (2011). "Dynamics of a plant RNA virus intracellular accumulation: stamping machine *vs.* geomentric replication." Genetics **188**: 637-646.

Muller, H. J. (1964). "The relation of recombination to mutational advance." Mutat. Res. **1**: 2-9.

Nakatsukasa-Akune, M., et al. (2005). "Suppression of root nodule formation by artificial expression of the *TrEnodDR1* (coat protein of *White clover cryptic virus 2*) gene in *Lotus japonicus*." Mol. Plant-Microbe Interact. **18**(10): 1069-1080.

Nibert, M. L. and L. A. Schiff (2001). Reoviruses and Their Replication. Field's Virology. D. M. Knipe and P. M. Howley. Philadelphia, Lippincott Williams &Wilkins**:** 1679-1728.

Novella, I. S., et al. (1995). "Exponential increases of RNA virus fitness during large population transmissions." Proc. Natl. Acad. Sci. USA **92**: 5841-5844.

Novella, I. S., et al. (1995). "Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness." J. Virol. **69**(5): 2869-2872.

Patton, J. T. and E. Spencert (2000). "Genome replication and packaging of segmented double-stranded RNA viruses." Virology **277**: 217-225.

Pita, J. S., et al. (2007). "Environment determines fidelity for an RNA virus replicase." J. Virol. **81**(17): 9072-9077.

Pita, J. S. and M. J. Roossinck (2013). "Mapping viral functional domains for genetic diversity in plants." J. Virol. **87**(2): 790-797.

Roossinck, M. J. (1997). "Mechanisms of Plant Virus Evolution." Annu. Rev. Phytopathol. **35**: 191-209.

Roossinck, M. J. (2008). Mutant Clouds and Bottleneck Events in Plant Virus Evolution. Origin and Evolution of Viruses. E. Domingo, C. R. Parrish and J. J. Holland. New York, Academic Press**:** 251-258.

Roossinck, M. J. (2010). "Lifestyles of plant viruses." Phil. Trans. R. Soc. B **365**: 1899-1905.

Roossinck, M. J. (2012). Persistent Plant Viruses: Molecular Hitchhikers or Epigenetic Elements? Viruses: Essential Agents of Life. G. Witzany. Dordrecht, Springer**:** 177-186.

Roossinck, M. J. (2013). "Plant Virus Ecology." PLoS Pathogens **9**(5): e1003304.

Sanfacon, H. (2005). "Replication of positive-strand RNA viruses in plants: contact points between plant and virus components." Canadian J. Bot. **83**(12): 1529-1549.

Sardanyés, J., et al. (2009). "Replication mode and landscape topology differentially affect RNA virus mutational load and robustness." J. Virol. **83**(23): 12579-12589.

Schneider, W. L. and M. J. Roossinck (2000). "Evolutionarily related sindbis-like plant viruses maintain different levels of population diversity in a common host." J. Virol. **74**(7): 3130-3134.

Schneider, W. L. and M. J. Roossinck (2001). "Genetic diversity in RNA viral quasispecies is controlled by host-virus interactions." J. Virol. **75**(14): 6566-6571.

Steinhauer, D. A., et al. (1992). "Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase." Gene **122**: 281-288.

Thébaud, G., et al. (2010). "The relationship between mutation frequency and replication strategy in positive-sense single-stranded RNA viruses." Proc. R. Soc. B **277**: 809-817.

Vignuzzi, M., et al. (2006). "Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population." Nature **439**: 344-348.

Chapter 4

The effect of a persistent virus on its host

A persistent plant virus benefits its host by manipulating aphid behavior

Abstract

Plants are frequently infected with cytoplasmic RNA viruses that persist for many generations through nearly 100% vertical transmission without producing any symptoms. Movement between plant cells and horizontal transmission have not been observed in these viruses; instead they are distributed to all host cells through host cell division. Jalapeño peppers (*Capsicum annuum*) are all infected with *Pepper cryptic virus* 1 (PCV 1, *Partitiviridae)*. In wild plants partitivirus infection decreased the likelihood of acute virus infection. We compared the effect of odor cues from PCV 1 infected (J+) and virus free (J-) Jalapeño pepper on the aphid *Myzus periscae*, a common vector of acute plant viruses. Pairwise preference experiments showed a stark contrast to insect-plant interactions in acute virus infections: virus infected plants deterred aphids. The acute plant virus *Cucumber mosaic virus* (CMV) manipulates its host's volatile emission to attract aphid vectors and facilitate its transmission. J+ and J- were inoculated with CMV. Volatiles of J+ and J- CMV infected plants were more attractive to aphids than J+ and J- mock inoculated plants. However, in pairwise preference between J+ CMV- and J- CMV-infected plants, aphids preferred the J- CMV volatile blend. Aphid fecundity on J+ and J- plants was measured as an indicator for the effect of PCV 1 on host quality for aphids. Aphid reproduction on J+ plants was more than two fold lower than J- plants. This study demonstrates a beneficial relationship between PCV 1 and Jalapeño plants by protecting the plants from the vector of acute viruses.

Introduction

Vectors play a crucial role in the transmission of acute plant viruses because the plants are largely immobile (Whitfield, Falk et al. 2015). Plant-feeding insects are the most common plant virus vectors, and among insects, aphids are the most common vectors of agriculturally important plant viruses (Ng and Perry 2004). Aphids probe any plant they land on to determine if it is a suitable food source (Powell, Tosh et al. 2006), and during this probing process many viruses can be transmitted. Because of plant virus dependence on vectors for their transmission, viruses evolved to manipulate some host and vector features to facilitate transmission.

Studies on the impact of acute plant viruses on insect behavior began in 1951; *Aphis fabae* population growth was higher on virus-infected sugar beets than healthy ones (Kennedy 1951). Studies on two viruses from the *Luteoviridae* family (*Barley yellow dwarf virus,* BYDV and *Potato leafroll virus*) and their aphid vectors showed that virus infection changed aphid development time and fecundity (Jiménez-Martínez, Pérea et al. 2004). In addition, virus-induced changes in hosts in both luteoviruses, encouraged aphid feeding on infected plants, which led to rapid population growth of aphids and subsequent dispersal of viruliferous aphids (Montllor and Gildow 1986, Eigenbrode, Ding et al. 2002, Jiménez-Martínez, Bosque-Pérez et al. 2004, Srinivasan, Alvarez et al. 2006). These aphid behavioral changes were consistent with the transmission mode of luteoviruses by aphids. Luteoviruses are acquired by vector feeding for hours to days on the phloem tissues of infected plants (Gray and Gildow 2003, Blanc, Drucker et al. 2014, Whitfield, Falk et al. 2015). Both viruses affected volatile organic compounds (VOCs) of their host plants, and vectors were responsive to host VOCs (Eigenbrode, Ding et al. 2002, Jiménez-Martínez, Bosque-Pérez et al. 2004, Medina-Ortega, Bosque-Pérez et al. 2009). The non-viruliferous aphid vector of BYDV preferred infected host plants, while viruliferous aphids preferred uninfected hosts; these changes promote pathogen spread (Ingwell, Eigenbrode et al.

2012). Using membrane-fed viruliferous aphids demonstrated that aphid behavioral alterations are mediated directly by acquisition the virus (Ingwell, Eigenbrode et al. 2012).

Cucumber mosaic virus (CMV) is a widespread plant pathogen with more than 1200 hosts species (Edwardson and Christie 1991). Aphid vectors can acquire CMV during brief probing of infected plant cells and transmit them effectively if the vector disperses rapidly to a new host plant (Martín, Collar et al. 1997, Blanc, Drucker et al. 2014, Whitfield, Falk et al. 2015). CMV infected squash plants, despite their reduced size, have elevated levels of VOCs that attract aphid vectors to the infected plants (Mauck, DeMoraes et al. 2010). CMV infection initially attracts aphids, but then reduces quality and palatability of plants so that aphids disperse quickly after virus acquisition (Mauck, DeMoraes et al. 2010). Salicylic acid- and jasmonic acid-mediated host defense responses are not involved in diminishing the plant quality for aphids; instead, CMV infection induces changes in the concentrations of carbohydrates and free amino acids in the plant tissue that causes reduction of plant quality for aphids (Mauck, DeMoraes et al. 2014). These virus-induced changes in the plants are in favor of CMV transmission: they attract aphids to feed on the infected plant, and disperse quickly to efficiently transmit the virus (Mauck, DeMoraes et al. 2014).

Recently it was shown that although CMV infection induced quantitative and qualitative changes to the VOC emission of tobacco plants, it did not change aphid preference for infected or uninfected tobacco plants (Tungadi, Groen et al. 2017). Inoculating the squash isolate of CMV on another host, pepper, did not follow the pattern observed on the native squash host. In addition, the virus-mediated changes in the new host were not consistent with CMV-vector interactions to aid its transmission (Mauck, DeMoraes et al. 2014). Hence virus-host-vector interactions appear to be adaptive.

Based on their lifestyles plant viruses are divided into two main groups; acute viruses and persistent viruses (Roossinck 2010). Most well-studied plant viruses are acute viruses that cause

disease in domestic plant hosts. In plants, infections by acute viruses can be resolved by recovery, death, or conversion to chronic infection. By contrast, plant persistent viruses replicate in their hosts for many generations, perhaps thousands of years. They do not cause obvious disease. The name 'persistent' comes from their lifestyle in the host, and differs from viruses that are persistently transmitted by a vector (Roossinck 2010, Roossinck 2013, Safari and Roossinck 2014). The persistent viruses are distributed to all plant cells through host cell division. They are not able to move between plant cells, and even grafting cannot transmit these viruses; there is no evidence for their horizontal transmission by a vector, while they are vertically transmitted to seeds at rates close to 100% (Boccardo, Lisa et al. 1987, Blanc 2007, Roossinck 2010). Unlike animal persistent viruses, plant persistent viruses cannot convert to an acute lifestyle. The most common persistent virus family in plants, and indeed the most common plant virus family overall, is the *Partitiviridae* family (Roossinck 2012, Roossinck 2014). The partitiviruses were first reported as cryptic viruses (Boccardo, Lisa et al. 1987). The *Partitiviridae* family includes viruses from diverse hosts: plants, fungi, and protozoan (Nibert, Woods et al. 2009). Phylogenetic analysis of their RNA dependent RNA polymerase implies that these viruses could be transmitted between plant and fungi (Roossinck 2010). Plants cannot be cured of these viruses using various treatments (Szegö, Tóth et al. 2005), although occasionally a virus-free plant will develop due to a rare lack of vertical transmission. *Pepper cryptic virus 1* (PCV 1) belongs to the *Deltapartitivirus* genus and consists of two linear, monocistronic double-stranded (ds) RNAs that are encapsidated in separate isometric particles (Valverde and Fontenot 1991, Valverde and Gutierrez 2008, Sabanadzovic and Valverde 2011). PCV 1 was first reported from the cultivar Jalapeño M, and as with all persistent plant viruses, there is no evidence for its transmission by any vector, mechanical inoculation or grafting; while ovule or pollen transmits the virus to over 98 % of the next generation (Valverde and Gutierrez 2008, Sabanadzovic and Valverde 2011).

Persistent viruses are very common in crop plants; hence it seems likely that they were selected during domestication. In this case they must provide some advantages over virus-free plants. While there are many studies on the effect of acute viruses on plant volatiles that impact insect behavior, no studies have looked at the effects of persistent viruses. Here the co-incidence of partitiviruses and acute viruses in a plant virus biodiversity inventory was analyzed, and used three isogenic lines of Jalapeño plants with and without PCV 1 to assess their attraction and plant quality traits for aphids.

Materials and Methods

Analysis of plant virus biodiversity data

Details of the plant virus biodiversity inventory in the Tallgrass Pairie Preserve (TGPP) of northeastern Oklahoma, including sampling, sample processing, sequence analysis and data analyses have been reported previously (Roossinck, Saha et al. 2010, Thapa, McGlinn et al. 2015). Plants with partitivirus-like RNA sequences were considered infected. Mixed infection with acute viruses was assessed for each of partitivirus positive plants.

Plants and insects

Three isogenic lines of Jalapeño peppers (*Capsicum annuum*) with and without PCV 1 (denoted as J1+/-, J2+/-, and J3+/-) were obtained from Dr. Rodrigo Valverde (Department of Plant Pathology and Crop Physiology, Louisiana State University). Plants were sown in 10 cm plastic pots containing Sun-Gro Horticulture soil and grown in an insect-free environmental room at 24°C and fluorescent light (16:8 light:dark photoperiod).

A colony of *M. persicae* was obtained from Gary Thompson (Penn State University), and maintained on turnips (*Brassica rapa*). To produce sufficient aphids for experiments the aphid colony was transferred to fresh turnip plants, and colonies were maintained with a natural photoperiod at 25°C. Aphid colonies were re-established by moving two aphid-infested leaves to a new turnip plant every 10-15 days.

CMV inoculation

Two-week-old pepper plants were kept in the dark for 24 hours and then mechanical inoculated using purified virus particles of Fny-CMV (Roossinck and Palukaitis 1990) diluted in 50 mM $NaHPO_4$ (pH 7.0). Leaves were dusted with carborundum to increase inoculation efficiency, followed by rubbing with acid-etched glass. Mock inoculation employed the same protocol but with buffer only. Plants were used for aphid preference experiments 20 days after inoculation, when infected seedlings showed virus symptoms.

Aphid preference to volatiles of plants with/without PCV 1

Aphid preference tests to plant volatile cues were performed using the experimental set up shown in Figure 4-1. Two cylinders (13 × 28 cm) containing different treatments were connected to an aphid box (11 × 17 cm) via two aphid collection cages (2.5 × 4 cm). The aphid box was covered to eliminate any visual cues. Each collection cage had one layer of extra fine mesh at the bottom to allow aphids to respond to VOCs without any contact cues, and trap aphids in the cage.

For the pairwise preference tests, 50 wingless non-viruliferous three to five day-old aphids were collected using a fine brush and starved for one hour, after which they were

transferred in the middle of aphid box (Figure 4-1). The number of aphids in each collection cage was recorded 45 min after releasing aphids. The percentage of the aphids that made a choice was calculated for each treatment. Pairwise preference tests were performed in three different control arrangements: (I) empty vs. empty, (II) J+ vs. empty, and (III) J- vs. empty. Aphid preference for three isogenic lines of J+ and J- were tested; (IV-1) J1+ vs. J1-, (IV-2) J2+ vs. J2-, and (IV-3) J3+ vs. J3-. Plants used in these tests were matched for size and development stage (four to six leaf stage). Each pairwise experiment was done 12 times using four to six plants. All experiments were done at 25°C and daylight to ensure that plants released a consistent volatile blend.

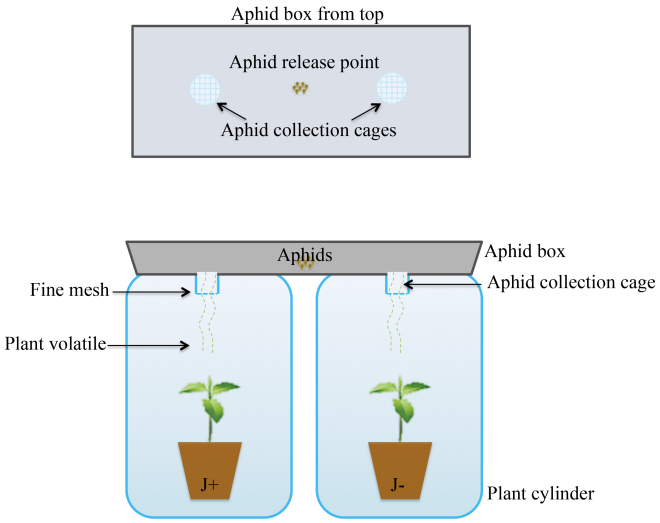

Figure **4-1**: **Experimental design for aphid preference test.** Schematic of experimental set up for aphid pairwise preference to plant volatiles. Wingless non-viruliferous 3 to 5 day old aphids were placed in the middle of aphid box, which was covered to eliminate any visual cues affecting aphid behavior. Aphids can make a choice between two collection cages that were connected to plant cylinders containing treatments. Each collection cage had one layer of fine-mesh to trap aphids in the cage, and the number of aphids in each collecting cage was recorded 45 min after releasing aphids.

Aphid preference to volatiles of plants with/without PCV 1 and CMV

To investigate aphid preference to volatiles of Jalapeño plants in the presence of CMV (an acute virus) and PCV 1 (a persistent virus), the same experimental set up as above was used (Figure 4-1). The pairwise preference tests were performed in 5 different plant arrangements including (V) J- mock vs. J- CMV, (VI) J- mock vs. J+ CMV, (VII) J+ mock vs. J+ CMV, (VIII) J+ mock vs. J- CMV, and (IX) J- CMV vs. J+ CMV. All experimental conditions were as above.

Aphid fecundity

Experiments were done using six- to eight-leaf stage Jalapeño plants in a pesticide-free greenhouse at 23 to 26°C temperature, supplemented with light for a 16-h day length. Eight adult wingless aphids were confined in a small clip cage (2.5 × 4 cm), and placed on a caged plant. Three clip cages were randomly attached to each plant, and for each line four plants were used (Figure 4-2). The adult aphids were left to reproduce inside the cage for 24 h, and then adults were removed and offspring were maintained in clip cages to mature and reproduce. During 14 days, offspring molting to adults, reproducing and producing new offspring were monitored in the clip cages. At day 14, the number of offspring and adults in each individual clip cage was recorded, and aphid fecundity was calculated by dividing the number of offspring produced by the number of adult aphids. This experiment was replicated 12 times for each Jalapeño line, and in total fecundity was measured for more than 60 aphids for each treatment.

$$\text{Aphid fecundity} = \frac{\text{number of offspring}}{\text{number of adults}}$$

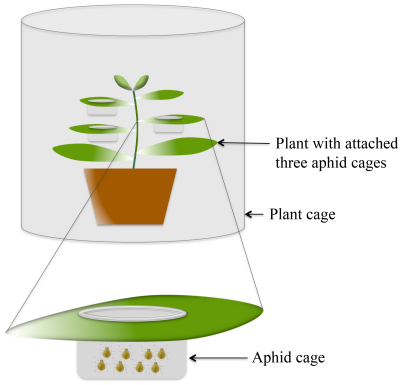

Figure **4-2**: **Experimental design for aphid fecundity tests.** Aphid fecundity was tested at 6 to 8 leaf old Jalapeño plants. Eight adult wingless aphids were confined in a small clip cage, and placed on a caged plant (3 clip cages per plant). These adult aphids were removed after 24 h and offspring were left in each clip cage. They molted to adults and produced new offspring. Aphid fecundity in each clip cage was calculated by dividing the number of offspring produced by the number of adult aphids on day 14.

Statistical analysis

Statistical analysis was performed using 'R', version 3.3.3. A General Linear Model (GLM) with binomial error was used to analyze aphid preference data for all pairwise experiments, and estimated the proportion of aphids that moved to the preferred option, and 95% confidence intervals, for each pairwise test. For statistical analysis of aphid fecundity the lme4 package in R was used and two models were defined. The null model for aphid fecundity included different plants and lines assumed as random effects. A second model included the presence of PCV 1 considered as a fixed effect, with the same random effects as the null model. A likelihood ratio test of the two models was performed (using ANOVA in 'R') to test the significance of the fixed effect.

Results

Co-infections of partitiviruses and acute viruses in wild plants

From 2005 to 2007 wild plants were sampled and assessed for RNA virus infection in the TGPP (Roossinck, Saha et al. 2010, Thapa, McGlinn et al. 2015). A total of 1276 individual plants were analyzed, and 220 (18%) of these had evidence of paritivirus infection based on sequence similarity to known partitiviruses. The incidence of acute viruses in these same plants was 21%. However, co-infection of a partitivirus and an acute virus in these plants was very low. Only six plants (0.03%) were coinfected. In plants infected with other persistent viruses (totiviruses, endornaviruses or chrysoviruses) there was no obvious bias (unpublished data). Four of the six co-infected plants were infected by a tymovirus that was widespread throughout the study area (Min, Feldman et al. 2012), while the other two were infected by a bromovirus and a betaflexivirus. Since insects vector the majority of acute plant viruses, I hypothesized that the presence of a partitivirus might be affecting acute virus infection by deterring insect vectors.

Aphid preference to volatiles of plants with/without PCV 1

To examine the response of *M. persicae* to VOC cues, three pairwise preference experiments were performed in which aphids were exposed to odors of I) two empty cylinders, II) an empty cylinder or a J+ plant, and III) an empty cylinder or a J- plant (Table 4-1, Figure 4-1). In experiment I, there was no significant difference between choosing two empty cylinders, while in experiments II and III, aphids selected cylinders containing plants. Figure 4-3A shows the mean percentage of aphids arrested in collection cages for each experiment, and the pairs involved in each choice test are shown bi-directionally. The 95% confident intervals for the probability of

selecting the empty cylinder were 0.46–0.57 for I, 0.23–0.34 for II, and 0.22–0.33 for III (Figure 4-3B); only in the preference test with both empty cylinders (I: empty vs. empty) was 0.5 (i.e. no preference) within 95% confident intervals. Hence aphid choice was not random and they were responding to volatile cues. In experiments IV-1, IV-2 and IV-3, three isogenic lines of Jalapeño with and without PCV 1 were used to discover how PCV 1 is affecting aphid behavior. A similar pattern was observed in all three lines and aphids demonstrated their preference to volatiles of virus free plants (J-). The 95% confident intervals for selecting the J+ over J- plants were 0.36–0.49 (IV-1: J1+ vs. J1-), 0.37–0.49 (IV-2: J2+ vs. J2-), and 0.31–0.41 (IV-3: J3+ vs. J3-) as it is shown in Figure 4-3B. Aphids showed discrimination against volatiles of PCV 1 infected plants, and 0.5 was not contained in any of the 95% credible intervals. Using three different Jalapeño lines verified that aphid behavior is affected by manipulation of plant volatiles by PCV 1.

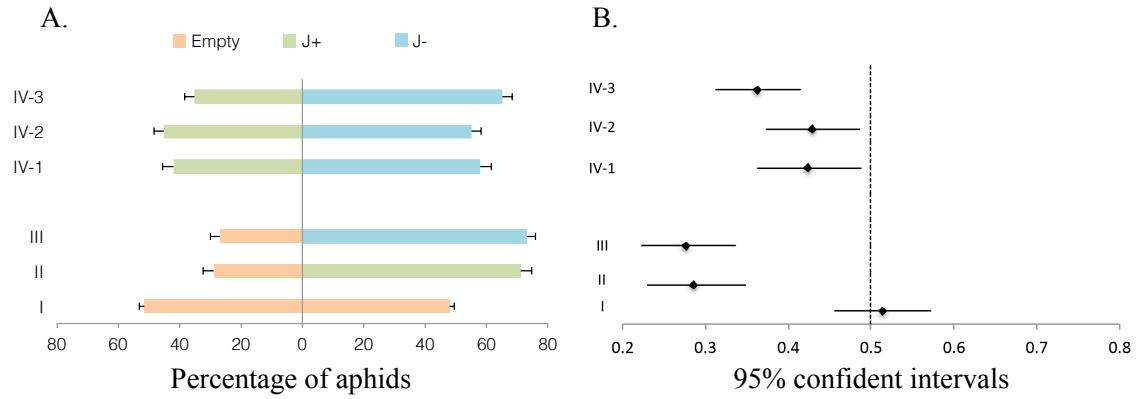

Figure **4-3**: **Aphid preferences to volatiles of plants with/without PCV 1.** Control pairwise preference tests (I to III) were performed to test if aphids were responding to volatile cues: I) Empty vs. Empty; II) Empty vs. J+; and III) Empty vs. J-. In experiments IV-1 to IV-3 aphid preferences to volatiles of three different isogenic lines of Jalapeño were tested: IV-1) J1+ vs. J1-; IV-2) J2+ vs. J2-; and IV-3) J3+ vs. J3-. A. Numbers represent the mean percentage of aphids in collection cages at the end of each treatment and pairs involved in each choice test are shown horizontally. In arrangement I, there is no significant difference between choosing two empty cylinders, while in arrangement II and III, cylinders containing plants were selected. In experiments IV-1, IV-2 and IV-3, three isogenic lines of J+ and J- were used to test aphid attraction to PCV 1 infected or virus free plants. A similar pattern was observed in all three lines, and aphids demonstrated their preference for virus free plants (J-). B. 95% confident intervals for the probability of selecting the empty cylinder in arrangement I to III: 0.46–0.57 (I: Empty vs. Empty); 0.23–0.34 (II: Empty vs. J+); and 0.22–0.33 (III: Empty vs. J-). The 95% confident

intervals for selecting the J+ over J- plants were 0.36–0.49 (IV-1: J1+ vs. J1-); 0.37–0.49 (IV-2: J2+ vs. J2-); and 0.31–0.41 (IV-3: J3+ vs. J3-).

Table **4-1**: Aphid preference experimental arrangements.

I	Empty	vs.	Empty
II	Empty	vs.	J+
III	Empty	vs.	J-
IV-1	J1+	vs.	J1-
IV-2	J2+	vs.	J2-
IV-3	J3+	vs.	J3-
V	J- mock	vs.	J- CMV
VI	J- mock	vs.	J+ CMV
VII	J+ mock	vs.	J+ CMV
VIII	J+ mock	vs.	J- CMV
IX	J+ CMV	vs.	J- CMV

Aphid preference to volatiles of plants with/without PCV 1 and CMV

Several studies have shown that CMV manipulates its host by inducing elevated levels of volatile emissions that attract aphid vectors for its transmission (Mauck, DeMoraes et al. 2010, Mauck, DeMoraes et al. 2014). In order to explore the effect of PCV 1 on aphid attraction to CMV infected plants, J+ and J- plants were inoculated with CMV, and performed aphid pairwise preference studies (Figure 4-1) in experiments V to IX. Aphid attraction to volatiles of mock-inoculated virus free plants (J- mock) vs. CMV infected plants (J- CMV and J+ CMV) was compared in V and VI (Table 4-1). In both experiments aphids preferred volatiles from CMV infected plants (Figure 4-4A). In arrangement VII and VIII, pairwise comparisons were between mock-inoculated PCV 1 infected plants (J+ mock) vs. CMV infected plants (J+ CMV and J- CMV). A similar pattern was observed, and aphids significantly preferred the volatile of CMV

infected plants (Figure 4-4A). The 95% confident intervals for the probability of selecting the mock-inoculated plants in arrangements V to VIII ranged from 0.27-0.41 (Figure 4-4B). However, when pairwise comparison was done between volatile emissions from CMV-infected J+ and J- plants (IX: J+ CMV vs. J- CMV) aphids preferentially selected the volatiles of J- CMV plants (Figure 4-4A). The 95% confident intervals for choosing the J+ CMV was 0.32–0.41 (IX: J+ CMV vs. J- CMV) (Figure 4-4B).

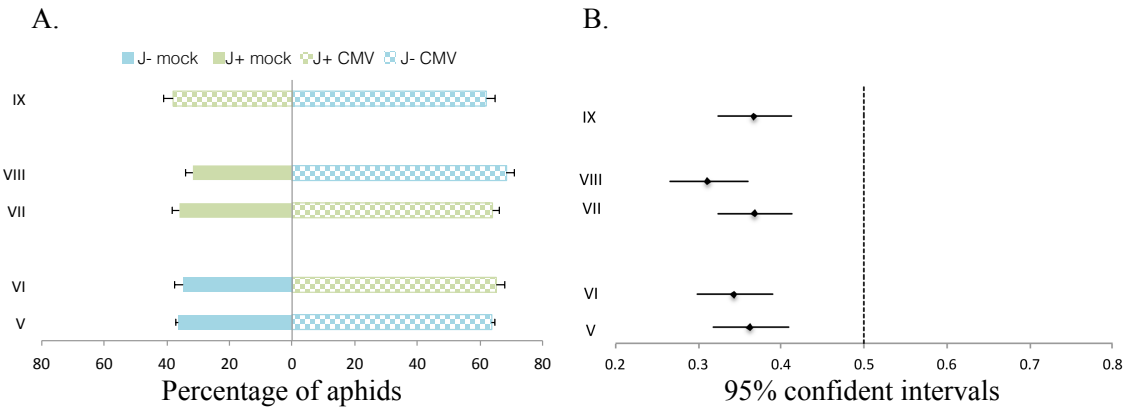

Figure **4-4**: **Aphid attraction to volatiles of mock or CMV-infected J+ and J- plants.** Pairwise preference experiments were: V) J- mock vs. J- CMV; VI) J- mock vs. J+ CMV; VII) J+ mock vs. J+ CMV; VIII) J+ M vs. J- CMV; and IX) J+ CMV vs. J- CMV. A. Numbers represent the mean percentage of aphids in collection cages at the end of each treatment and pairs involved in each choice test are shown horizontally. In arrangements V and VI, J- mock volatiles were less attractive to aphids than J-CMV and J+CMV plants. In arrangements VII and VIII, J+ M volatiles were also less appealing to aphids than J+ CMV and J- CMV plants. In arrangement IX, aphids preferentially chose volatiles of J- CMV to J+ CMV. B. 95% confidence intervals for the probability of selecting the J+ and J- mock in arrangements V to VIII were: 0.32–0.0.41 (V: J- mock vs. J- CMV); 0.30–0.40 (VI: J- mock vs. J+ CMV); 0.32–0.41 (VII: J+ mock vs. J+ CMV); 0.27-0.36 (VIII: J+ mock vs. J- CMV). 95% confident intervals for choosing the J+ CMV was 0.32–0.0.41 (IX; J+ CMV vs. J- CMV).

Aphid fecundity

Aphid fecundity was used as an indicator of the impact of PCV 1 on plant quality. Reproduction of non-viruliferous aphids that developed from day one to adulthood on virus free and PCV 1-infected plants over 14 days was compared. Since it took eight to nine days to complete the first reproduction cycle, only one generation of offspring are born within this period. Average fecundity on J1+ (74 adult aphids) was 7; whereas fecundity of J1- (71 adult aphids) was 13; for J2+ (69 adult aphids) fecundity was 9.5 compared to J2- (71 aphids) where fecundity was 21; for J3+ (63 aphids) fecundity was 9.3, compared to 22 for J3- (61 aphids) (Figure 4-5, Table 4-2). In all three lines, aphid reproduction on the PCV 1 infected plants (J+) was more than two fold lower than virus free plants (J-). Statistical analysis indicated the affect of virus on aphid fecundity was significant (p value < 0.001). This showed the beneficial role of PCV 1 in protecting its host from increasing population growth of aphids. In Line 1 aphid fecundity for both treatments (J1+ and J1-) was lower than in the other two lines. Although the conditions used were the same, the experiments were not all done at the same time, so comparisons are only valid within an experiment, although we cannot rule out distinct plant qualities in the different lines. In all three lines the same trend of higher fecundity on J- in comparison with J+ plants can be seen; Jalapeño plants employed PCV 1 to protect themselves from increasing aphid population.

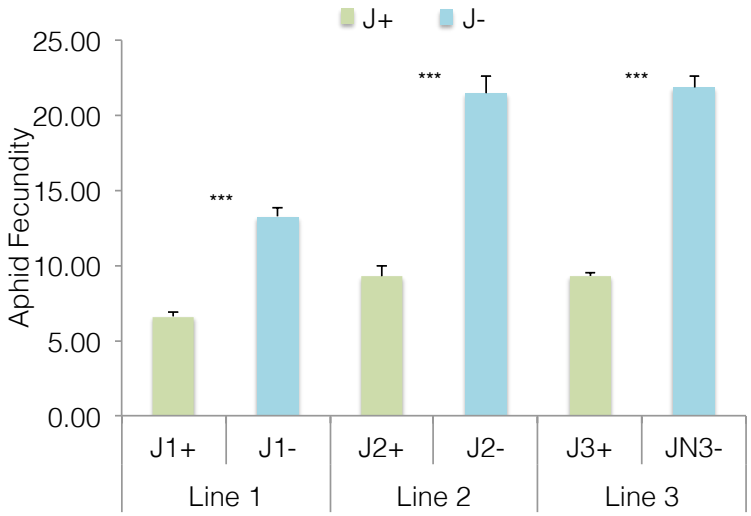

Figure **4-5**: Aphid fecundity on plants with and without PCV1 tested on three isogenic lines of Jalapeño plants. Fecundity was calculated by dividing the number of offspring by the number of the adults on each clip cage after 14 days. The results are the average (±SE) of 12 clip cages per treatment. P values: line 1 (2.485×10-6); Line 2 (5.812×10-5); Line 3 (1.928×10-6). Statistical analysis indicated the affect of virus on aphid fecundity was significant (p value = 2.6×10-10). Asterisks indicate significant differences between control J+ and J- plants (*** $p < 0.001$).

Table **4-2**: Number of adults, offspring, and aphid fecundity for three isogenic lines of Jalapeño infected with PCV 1 (J+) and virus free (J-) plants.

	J+			J-		
	Adults	Offspring	Fecundity	Adults	Offspring	Fecundity
Line 1	74	506	6.83	71	913	12.86
Line 2	69	655	9.49	71	1490	20.98
Line 3	63	586	9.30	61	1326	21.74

Discussion

Behavioral responses that impact insect dispersal are important because of their effect on viral disease spread. Several studies have shown that acute viruses, in spite of their small size and negative effects on apparent plant quality, manipulate their host volatiles so that infected plants

are more enticing to vectors. Virus-induced changes are in favor of virus transmission: attracting aphids to feed on plants and acquiring the virus for its dispersal (Eigenbrode, Ding et al. 2002, Jiménez-Martínez, Bosque-Pérez et al. 2004, Medina-Ortega, Bosque-Pérez et al. 2009, Mauck, DeMoraes et al. 2010). These virus-induced changes in vector behavior are adaptive (Mauck, DeMoraes et al. 2014). Here it is shown that partitivirus infection also influences aphid behavior (Figure 4-3 IV-1, IV-2, IV-3), but, this interaction is in stark contrast to insect-plant interactions in acute virus infection: the partitivirus deters aphids and protects plants from the vector of acute viruses. During thousands of years of virus replication in the plant host, the virus-plant interaction evolution is beneficial for the plants by discouraging the vector of acute viruses.

The acute plant virus *Cucumber mosaic virus* (CMV) manipulates its host's volatile emission to attract aphid vectors and facilitate its transmission. In inoculated J+ and J- with CMV; volatiles of J+ and J- CMV infected plants were more attractive to aphids than J+ and J- mock inoculated plants (Figure 4-4 V, VI, VII, VIII). This is consistent with other studies that the VOC emissions from CMV-infected plants are more desirable for aphids than those from uninfected plants (Mauck, DeMoraes et al. 2010, Mauck, DeMoraes et al. 2014). PCV 1 did not counter the effect of CMV on the plant VOCs, perhaps because of a strong elevation of VOCs by CMV. However, in pairwise preference between J+ CMV- and J- CMV-infected plants, aphids preferred the J- CMV volatile blend (Figure 4-4). Therefore, the presence of PCV 1 did influence aphid choice when both plants were CMV-infected. This outcome is again consistent with wild plant data and verified our hypothesis that partitivirus infection decreases the likelihood of acute virus transmission, even in the presence of an acute virus.

Finally, the aphid fecundity on J+ and J- plants was measured as an indicator for the effect of PCV 1 on host quality for aphids. Aphid reproduction on J+ plants was more than two fold lower than J- plants (Figure 4-5). There are several factors that can affect aphid fecundity. Turgor pressure can affect aphid feeding and therefore fecundity (Wearing 1967). However, in

this study turgor pressure effects was removed by watering plants regularly, and similarly. Host plant quality during aphid growth and development (both nymph and adult stage) is another key factor of aphid fecundity [for review see (Awmack and Leather 2002)]. In poor host quality condition, aphids resorbed their eggs for their own survival (Leathers, Tanguay et al. 1993, Sequeria and Dixon 1996); Nitrogen also plays an important role in plant quality and impacts aphid fecundity. Early in the growing season phloem amino acid content is high, and aphid fecundity is higher than when leaves mature and phloem amino acids level drop (Dixon 1970, Leather and Dixon 1981, Weibull 1987, Leather, Wade et al. 2005).

Virus infections cause biochemical and physical changes in the host, and consequently affect the fecundity of aphids as well. The population of *S. graminium*, *R. padi* and *Sitobion avenae* increased on BYDV infected plants in comparison with uninfected plants (Montllor and Gildow 1986, Araya and Foster 1987, Fereres, Lister et al. 1989). Moreover, aphid fecundity was higher on wheat cultivars sensitive to BYDV than on BYDV-tolerant wheat cultivar (Fereres, Lister et al. 1989). Aphid population growth was correlated with the virus-vector relationship. The acquisition of BYDV requires continual aphid feeding on the phloem of infected plants. Thus, BYDV-induced changes in plants encourage sustained feeding and lead to aphid settling and rapid population growth followed by dispersal of viruliferous aphids (Montllor and Gildow 1986, Eigenbrode, Ding et al. 2002, Jiménez-Martínez, Bosque-Pérez et al. 2004, Srinivasan, Alvarez et al. 2006). On the other hand, CMV, which is transmitted in a non-persistent manner, induced changes in the concentrations of carbohydrates and free amino acids in the plant tissue that lead to reduction of plant quality for aphids. By reducing plant quality and palatability, aphids are dispersed quickly and efficiently transmit CMV to a healthy plant (Martín, Collar et al. 1997, Mauck, DeMoraes et al. 2010, Blanc, Drucker et al. 2014, Whitfield, Falk et al. 2015).

Hence, while acute plant viruses manipulate the behavior of aphid vectors in their favor, enhancing their transmission, PCV 1, a persistent plant virus that is not horizontally transmitted,

protects its plant host from aphids by reducing the attractiveness of the plant and the quality of the plant as an aphid host.

Effect of PCV 1 on the developmental growth of Jalapeño plant

Acute virus infection can lead to morphological changes in plant, such as chlorosis, stunting, and changes in the plant developmental growth (Hull 2002). There are several plant acute viruses reported from pepper plants. *Cucumber mosaic virus*, *Potato virus Y*, *Tobacco mosaic virus*, *Pepper mottle virus*, *Pepper severe mosaic virus*, *Pepper yellow mosaic virus*, *Tomato spotted wilt virus*, *Chilli veinal mottle virus*, *Pepper veinal mottle virus*, *Chilli ringspot virus*, and *Wild tomato mosaic virus* are some of the viruses identified from pepper plants. Many have been reported with some symptoms on pepper plants such as small and distorted leaves and fruits, leaf and fruit mottle, plant stunting, flower dropping, and yield loss depending on the virus strain, the plant cultivar, environmental conditions, and time of the virus infection (Kenyon, Kumar et al. 2014).

The impact of plant persistent viruses on their host has not been studied thoroughly. Their frequent presence in crops confirms that they have no deleterious effect on their hosts, and implies that they may have beneficial effects that were selected during the domestication of crops. Another study has shown that the coat protein gene of *White clover cryptic virus 2* affects the nodulation regulation in clover and in a legume, *Lotus japonicus* (Nakatsukasa-Akune, Yamashita et al. 2005). The discovery of expressed integrated persistent virus sequences in plant genomes is further evidence of the potential importance of these viruses (Liu, Fu et al. 2010, Chiba, Kondo et al. 2011). Also, there is some evidence that shows a potential role of persistent viruses in the evolution of acute viruses (Roossinck 2005, Rastgou, Habibi et al. 2009). These viruses could be used as epigenetic elements that provide novel functions, or even as a source for new emerging viruses (Roossinck 2010, Roossinck 2012)}. Long-term relationships of plant persistent viruses with their hosts, and their strict vertical transmission suggest beneficial interactions. In most studies on these viruses they have no effect on their host, but finding an isogenic plant without the

virus is not common. So far, there is no report of the effects of persistent plant viruses on the plant phenotype. To test the effect of PCV1 on developmental growth of Jalapeño plants, virus influence on seed germination time, the required time for the emergence of first true leaf, first open flower and the dry biomass of plants in J+ and J- were analyzed.

Materials and Methods

To test the effect of PCV 1 on developmental growth of Jalapeño, first its effect on the germination time of Jalapeño seeds was tested; two hundred seeds from J+ and J- including all three isogenic lines were selected (80, 60, 60 seed for each treatment in Line 1, 2, and 3, respectively), and surface sterilized using sodium hypochlorite (NaClO) 10% for 10 min following three washes with sterilized distilled water. Seeds were placed in the middle of six layers of wet paper towels in a petri dish at 25 °C for each treatment, including ten petri dishes per treatment and twenty seeds per petri dish. They were examined by daily observation, and the germination time was recorded by the appearance of the root radical.

To test the effect of PCV1 on the average time required for the emergence of the first true leaf and open flower, 25 germinated seeds were randomly selected from each treatment (J+ and J-), and transplanted to 10 cm plastic pots containing Sun-Gro Horticulture soil and grown in an insect-free environmental room at 24°C and fluorescent light (16:8 light:dark photoperiod). Seedling were monitored daily, and the day of the first unfurled true leaf emergence, and the first open flower for each plant were recorded. To estimate the effect of virus on the plant dried biomass, the whole above ground plant was placed in an envelope and dried at 65 °C for 48 h, followed by weighing. All these experiments were done from March to June 2017. The significance of the effect of PCV1 on different experiments was assessed using the Student t-test ($p \leq 0.05$).

Results and discussion

There was no any evidence of mosaic, slower growth, stunting, or fruit malformation in virus-infected Jalapeño lines; as it can be seen in Figure 4-6, the PCV1 infected and virus-free Jalapeño lines have no evidence of morphological differences.

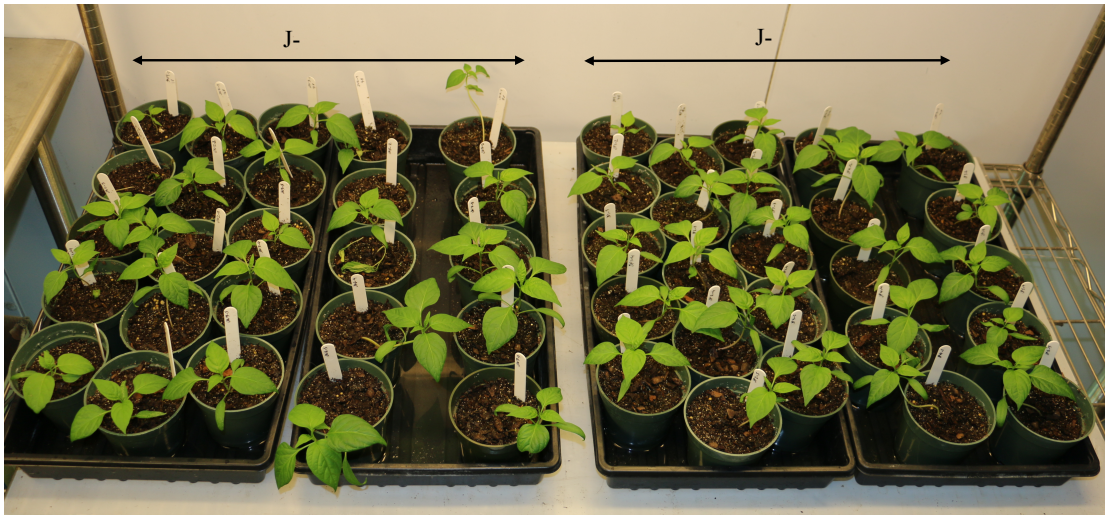

Figure **4-6**: Virus free Jalapeño plants (J-) and PCV 1 infected Jalapeño plants (J+).

However, previously I observed that the PCV 1 infected seeds germinated faster than virus free Jalapeño seeds. To test the effect of PCV 1 on germination time of Jalapeño seeds two hundred seeds from each treatment (J+ and J-) including all three isogenic lines was used. Both J+ and J- seeds were collected recently. Their germination time was recorded by the date of emergence of the root radical. The average seed germination time for J+ and J- was 7.4 ± 0.12 and 7.6 ± 0.13 days; the J+ seeds germinated slightly faster than J-, but the difference was not significant (t-test, p value > 0.05) (Figure 4-7A).

Average required time to emerge the first unfurled leaf for both PCV1 infected and virus-free Jalapeño lines were examined by daily observation of plants. The average time needed for J+

plants to have the first leaf was 26 ± 0.36 days and for J- was 25.5 ± 0.33 days (Figure 4-7B). In the next step, the date of having the first open flower was recorded. The average time required for each J+ and J- to have the first open flower was 88.4 ± 0.97 and 88.0 ± 1.03 days, respectively (Figure 4-7C). At the end of the experiment the dried biomass of plants were measured; the average dried biomass for J+ was recorded 2.87 ± 0.18 g and for J- was 2.66 ± 0.18 g (Figure 4-7D). Statistical analysis for required time to germinate, appearance of the first true leaf, the first open flower, and the plants dried biomass showed that the virus is not significantly affecting the plants developmental growth. The detailed data for each line is presented in Table 4-3.

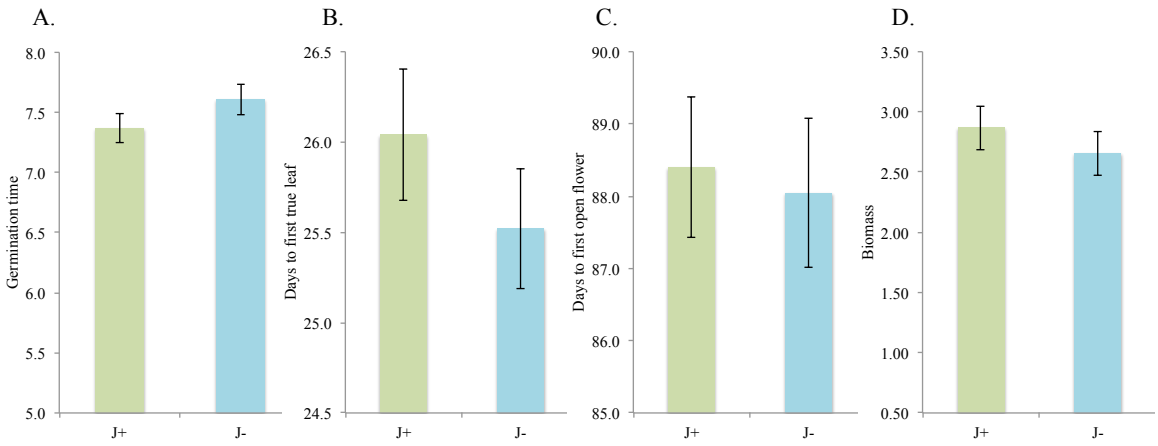

Figure **4-7**: Effect of PCV1 on developmental growth of Jalapeño. A. Germination rate of Jalapeno seed infceted with PCV1 (J+), and virus free (J-). B. Effect of PCV1 on the day of first true leaf apprearance. C. Effect of PCV1 on the day of first open flower. D. Effect of PCV1 on the biomass of J+ and J- at 90 days. There was no significant difference on J+ and J- in four measured developmental growth.

Table **4-3**: The effect of PCV 1 on developmental growth of Jalapeño plants.

	% Germinated	Germination T[1] (Days)	Leaf T[2] (Days)	Flower T[3] (Days)	Biomass[4] (Grams)
J1+	95.0	8.0 ± 0.14	25.5 ± 0.31	88.8 ± 1.79	2.84 ± 0.35
J1-	98.7	8.1 ± 0.19	25.0 ± 0.42	87.3 ± 2.19	3.09 ± 0.36
J2+	95.0	6.9 ± 0.18	25.8 ± 0.36	86.8 ± 0.89	3.02 ± 0.17

J2-	93.3	7.3 ± 0.20	25.5 ± 0.40	86.7 ± 0.79	2.61 ± 0.21
J3+	100	7.0 ± 0.28	27.6 ± 1.47	90.8 ± 2.71	2.63 ± 0.46
J3-	98.3	7.3 ± 0.91	26.6 ± 1.17	93.0 ± 1.08	2.22 ± 0.32

1. The average germination time. 2. The average required time ± SE to the first true leaf.
3. The average required time ± SE to the first open flower. 4. The average aboveground biomass ± SE.

These results are consistent with previous reports and show that not only does PCV1 not cause any symptom in Jalapeño plants, but it also does not affect the developmental growth of plants. Previously, I observed that virus infected seeds germinate two fold faster than virus free ones. The poor germination was apparently related to the age of the seeds, and the seeds viability of PCV 1 infected plants may be extended. Persistent viruses may provide some benefits to their plant hosts as well as additional functional proteins (Villarreal 2009, Roossinck 2010). An example of a three-way symbiosis involving a mutualistic interaction between an obligate mycovirus, *Curvularia thermal tolerance virus*, an endophytic fungus, and a plant has been reported (Márquez, Redman et al. 2007). This three-way interaction conferred, plant tolerance to extremely high soil temperatures in Yellowstone National Park. Here it is shown that PCV 1 protects its plant host from aphids by reducing the attractiveness of the plant and the quality of the plant as an aphid host, while it is not affecting the developmental growth of its host.

Cited Literature

Araya, J. E. and J. E. Foster (1987). "Laboratory study on the effects of barley yellow dwarf virus on the life cycle of *Rhopalosiphum padi* (L.)." Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz **94**(6): 578-583.

Awmack, C. S. and S. R. Leather (2002). Host plant qulaity and fecundity in herbivorous insects. Annu. Rev. Entomol. M. R. Berenbaum, R. T. Carde and G. E. Robinson. Palo Alto, Annual REviews. **47:** 817-844.

Blanc, S. (2007). Virus transmission--getting in and out. Viral Transport in Plants. E. Waigman and M. Heinlein. Berlin Heidelberg, Springer-Verlag**:** 1-28.

Blanc, S., et al. (2014). "Localizing viruses in their insect vectors." Annu. Rev. Phytopathol. **52**: 403-425.

Boccardo, G., et al. (1987). "Cryptic plant viruses." Adv. Virus Res. **32**: 171-214.

Chiba, S., et al. (2011). "Widespread endogenization of genome sequences of non-retroviral RNA viruses into plant genomes." PLoS Pathogens **7**(7): e1002146.

Dixon, A. F. G. (1970). Quality and availability of food for a sycamore aphid population. Animal Populations in Relation to Their Food Sources, Aberdeen, Blackwell Science.

Edwardson, J. R. and R. G. Christie (1991). Cucumoviruses. CRC Handbook of viruses infecting legumes. Boca Raton, CRC Press**:** 293-319.

Eigenbrode, S. D., et al. (2002). "Volatiles from potato plants infected with potato leafroll virus attract and arrest the virus vector *Myzus persicae* (Homptera: Aphididae." Proc. R. Soc. B **269**: 455-460.

Fereres, A., et al. (1989). "Development and reproduction of the English grain aphid (Homoptera: Aphididae) on wheat culrivars infected with barley yellow dwarf virus." Environ. Entomol. **18**(3): 388-393.

Gray, S. and F. E. Gildow (2003). Luteovirus-aphid interactions. Annu. Rev. Phytopathol. K. R. Barker. **41:** 539-566.

Hull, R. (2002). Matthews' Plant Virology. San Diego, Academic Press.

Ingwell, L. L., et al. (2012). "Plant viruses alter insect behavior to enhance their spread." Scientific Reports **2**: 578.

Jiménez-Martínez, E. S., et al. (2004). "Volatile cues influence the response of *Rhopalosiphum padi* (Homoptera: Aphididae) to Barley yellow dward virus-infected transgenic and untransformed wheat." Environ. Entomol. **33**(5): 1207-1216.

Jiménez-Martínez, E. S., et al. (2004). "Life history of the Bird Cherry-oat aphid, *Rhopalosiphum padi* (Homoptera: Aphididae), on transgenic and untransformed wheat challended with *Barley yellow dwarf virus*." J. Econ. Entomol. **97**(2): 203-212.

Kennedy, J. S. (1951). "Benefits to aphids from feeding on galled and virus-infected plants." Nature **168**: 825-826.

Kenyon, L., et al. (2014). "Virus diseases of peppers (*Capsicum* spp.) and their control." Adv. Virus Res. **90**: 297-354.

Leather, S. R. and A. F. G. Dixon (1981). "The effect of cereal growth stage and feeding site on the reproductive activity of the bird-cherry aphid, *Rhopalosiphum padi*." Ann. appl. Biol. **97**: 135-141.

Leather, S. R., et al. (2005). "Plant quality, progeny sequence, and the sex ratio of the sycamore aphid, *Drepanosiphum platanoidis*." Entomol. exp. appl. **115**: 311-321.

Leathers, V., et al. (1993). "A phylogenetically conserved sequence within viral 3' untranslated RNA psuedoknots regulates translation." Mol. and Cell. Biology **13**(9): 5331-5347.

Liu, H., et al. (2010). "Widespread horizontal gene transfer from double-stranded RNA viruses to eukaryotic nuclear genomes." J. Virol. **84**(22): 11879-11887.

Márquez, L. M., et al. (2007). "A virus in a fungus in a plant --three way symbiosis required for thermal tolerance." Science **315**: 513-515.

Martín, B., et al. (1997). "Intracellular ingestion and salivation by aphids may cause the acquisition and inoculation of non-persistently transmitted plant viruses." J. Gen. Virol. **78**: 2701-2705.

Mauck, K. E., et al. (2010). "Deceptive chemical signals induced by a plant virus attract insect vectors to inferior hosts." Proc. Natl. Acad. Sci. USA **107**(8): 3600-3605.

Mauck, K. E., et al. (2014). "Biochemical and physiological mechanisms underlying effects of *Cucumber mosaic virus* on host-plant traits that mediate transmission by aphid vectors." Plant, Cell and Environment **37**: 1427-1439.

Mauck, K. E., et al. (2014). "Evidence of local adapation in plant virus effects on host-vector interactions." Integrative and Comparative Biology **54**(2): 193-209.

Medina-Ortega, K. J., et al. (2009). "*Rhopalosiphum padi* (Hemoptera: Aphididae) responses to volatile cues from Barley yellow dward virus-infected wheat." Environ. Entomol. **38**(3): 836-845.

Min, B.-E., et al. (2012). "Molecular characterization, ecology, and epidemiology of a novel tymovirus in *Asclepias viridis* from Oklahoma." Phytopath. **102**(2): 166-176.

Montllor, C. B. and F. E. Gildow (1986). "Feeding responses of two grain aphids to barley yellow dward virus-infected oats." Entomol. exp. appl. **42**: 63-69.

Nakatsukasa-Akune, M., et al. (2005). "Suppression of root nodule formation by artificial expression of the *TrEnodDR1* (coat protein of *White clover cryptic virus 2*) gene in *Lotus japonicus*." Mol. Plant-Microbe Interact. **18**(10): 1069-1080.

Ng, J. C. K. and K. L. Perry (2004). "Transmission of plant viruses by aphid vectors." Mol. Plant Path. **5**(5): 505-511.

Nibert, M. L., et al. (2009). "Cryspovirus: a new genus of protozoan viruses in the family Partitiviridae." Arch. Virol. **154**: 1959-1965.

Powell, G., et al. (2006). "Host plant selection by aphids: Behavioral, evolutionary, and applied perspectives." Ann Rev. Entomol. **51**: 309-330.

Rastgou, M., et al. (2009). "Molecular characterization of the plant virus genus Ourmiavirus and evidence on inter-kingdom reassortment of viral genome segments as its possible route of origin." J. Gen. Virol. **90**: 2525-2535.

Roossinck, M. J. (2005). "Symbiosis *versus* competition in the evolution of plant RNA viruses." Nature Rev. Microbiology **3**(12): 917-924.

Roossinck, M. J. (2010). "Lifestyles of plant viruses." Phil. Trans. R. Soc. B **365**: 1899-1905.

Roossinck, M. J. (2012). Persistent Plant Viruses: Molecular Hitchhikers or Epigenetic Elements? Viruses: Essential Agents of Life. G. Witzany. Dordrecht, Springer**:** 177-186.

Roossinck, M. J. (2012). "Plant virus metagenomics: Biodiversity and ecology." Annu. Rev. Genet. **46**: 357-367.

Roossinck, M. J. (2013). "Plant Virus Ecology." PLoS Pathogens **9**(5): e1003304.

Roossinck, M. J. (2014). "Metagenomics of plant and fungal viruses reveals an abundance of persistent lifestyles." Frontiers in Microbiology **5**: 767.

Roossinck, M. J. and P. Palukaitis (1990). "Rapid induction and severity of symptoms in zucchini squash (*Cucurbita pepo*) map to RNA 1 of cucumber mosaic virus." Mol. Plant Microbe-Interact. **3**(3): 188-192.

Roossinck, M. J., et al. (2010). "Ecogenomics: Using massively parallel pyrosequencing to understand virus ecology." Mol. Ecol. **19 (S1)**: 81-88.

Sabanadzovic, S. and R. A. Valverde (2011). "Properties and detection of two cryptoviruses from pepper (*Capsicum annuum*)." VIrus Genes **43**: 307-312.

Safari, M. and M. J. Roossinck (2014). "How does the genome structure and lifestyle of a virus affect its population variation?" Curr. Op. Virol. **9**: 39-44.

Sequeria, R. and A. F. G. Dixon (1996). "Life history responses to host quality changes and competition in the Turkey-oak aphid, Myzocallis boerneri (Hemiptera: Sternorrhyncha: Callaphididae)." Eur. J. Entomol. **93**: 53-58.

Srinivasan, R., et al. (2006). "Influence of hairy nightshade *Solanum sarrachoides* (Sendtner) and *Potato leafroll virus* (Luteoviridae: *Polervirus*) on the host preference of *Myzus persicae* (Sulzer) (Homoptera: Aphididae)." Environ. Entomol. **35**(2): 546-553.

Szegö, A., et al. (2005). "Detection of high molecular weight dsRNA persisting in *Dianthus* species." Acta Biologica Szegediensis **49**(1-2): 17-19.

Thapa, V., et al. (2015). "Determinants of taxonomic composition of plant viruses at the Nature Conservancy's Tallgrass Prairie Preserve, Oklahoma." Virus Evolution **1**(1): vev007.

Tungadi, T., et al. (2017). "Cucumber mosaic virus and its 2b protein alter emission of host volatile organic compounds but not aphid vector settling in tobacco." Virol. J. **14**: 91.

Valverde, R. A. and J. F. Fontenot (1991). "Variation in double-stranded ribonucleic acid among pepper cultivars." J. Amer. Soc. Hort. Sci. **116**(5): 903-905.

Valverde, R. A. and D. L. Gutierrez (2008). "Molecular and biological properties of a putative partitivirus from jalapeño pepper (*Capsicum annuum* L.)." Rev. Mex. Fitopat. **26**(1): 1-6.

Villarreal, L. P. (2009). "Persistence pays: how viruses promote host group survival." Curr. Op. Microbiol. **12**: 467-472.

Wearing, C. H. (1967). "Studies on the relations of insect and host plant: II. Effects of water stress in host plants on the fecundity of *Myzus persicae* (Sulz.) and *Brevicoryne brassicae (L.)*." Nature **213**: 1052-1053.

Weibull, J. (1987). "Seasonal changes in the free amino acids of oat and barley phloem sap in relation to plant growth state and grothe of *Rhopalosiphum padi*." Ann. appl. Biol. **111**: 729-737.

Whitfield, A. E., et al. (2015). "Insect vector-mediated transmission of plant viruses." Virology **479-480**: 289-290.

Appendix A

Estimating the limit of virus detection

Based on the published data on the anatomical characterization of *C. annuum*, the number of cells in 5 g of pepper leaves was estimated. DsRNA was extracted from 5g of fresh plant tissue (Marengo and Jalapeño) positive for viruses (BPEV and PCV 1) and from virus free plants. The dsRNA extracted from virus free plants was used as the blank to measure the concentration of dsRNA extracted from virus-infected plant using a spectrophotometer. Then, dsRNA of BPEV and PCV 1 was diluted by 1:1 serial dilutions and used as substrate for performing RT-PCR using RdRp primers (BPEV_RdRp and PCV1_RdRp) to estimated the limit of virus detection.

Estimating the limit of BPEV detection

I estimated 25.2 million cells in 5g of pepper leaves according to the published data on *C. annum* anatomical characterization (Weryszko-Chmielewska and Michalojc 2009, Dias, Gomes et al. 2013). The concentration of dsRNA extracted from Marengo positive was determined 1.5 fmol. As the calculation showed below, I estimated 36 molecules of BPEV per cell. Then, serial dilution of dsRNA and RT-PCR using BPEV_RdRp primers were performed. As it is shown in Figure A-1, at 5^{th} dilution the BPEV was detected by RT-PCR. Calculation revealed that the limit of BPEV detection was estimated 11 molecules per cell.

BPEV concentration: 1.5 fmol

- Genome copy number per cell:

  Mole to molecule: Molar × Avogadro number

  $1.5 \times 10^{-15} \times 6.022 \times 10^{23} = 90 \times 10^{7}$ molecules in 5g

  $90 \times 10^{7} / 25.2 \times 10^{6} =$ <u>36 molecules per cell</u>

- Limit of virus detection:

$0.45 \times 10^{-15} \times 6.022 \times 10^{23} = 27 \times 10^{7}$ molecule in 5g

$27 \times 10^{7} / 25.2 \times 10^{6} =$ 11 molecules per cell

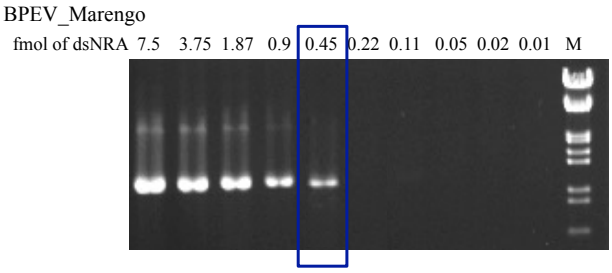

Figure **A-1**: RT-PCR products for detecting the limit of virus using BPEV_RdRp specific primers and dsRNA extracted from Marengo pepper diluted at 1:1. RT-PCR products were electrophoresed on 1.2 % agarose gel, using digested lambda DNA with *Eco*RI and *Hind*III as marker

Estimating the limit of PCV 1 detection

The concentration of dsRNA extracted from Jalapeño positive was determined 13 fmol. As the calculation showed below, I estimated 300 molecules of PCV 1 per cell. Then, serial dilution of dsRNA and RT-PCR using PCV1_RdRp primers were performed. As it is shown in Figure A-2, at 5th dilution the PCV 1 was detected by RT-PCR. Calculation revealed that the limit of PCV 1 detection was estimated 2.6 molecules per cell.

PCV1 concentration: 13 fmol

- Genome copy number per cell:

Mole to molecule: Molar × Avogadro number

$1.5 \times 10^{-15} \times 6.022 \times 10^{23} = 78 \times 10^{8}$ molecules in 5g

$78 \times 10^{8} / 25.2 \times 10^{6} =$ 300 molecules per cell

- Limit of virus detection:

$4 \times 10^{-15} \times 6.022 \times 10^{23} = 66 \times 10^{6}$ molecule in 5g

$66 \times 10^6 / 25.2 \times 10^6 =$ 2.6 molecules per cell

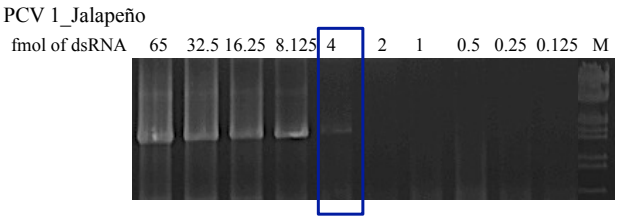

Figure **A-2**: RT-PCR products for detecting the limit of virus using PCV1_RdRp specific primers and dsRNA extracted from Jalapeño pepper diluted at 1:1. RT-PCR products were electrophoresed on 1.2 % agarose gel, using digested lambda DNA with *Eco*RI and *Hind*III as marker

Cited Literature

Dias, G. B., et al. (2013). "Characterization of *Capsicum* species using anatomical and molecular data." Genetics and Molecular Research **12**(4): 6488-6501.

Weryszko-Chmielewska, E. and Z. Michalojc (2009). "Anatomical features of leaves of sweet pepper (*Capsicum annuum* L.) fed with calsium using foliar nutrition." Acta Agrobotanica **62**(2): 155-164.

Appendix B

Supplementary Tables

Table **S1**: Pepper collection list

Species	Name	Location[1]	PI[2]	Seed source[3]
C. annuum	Marengo	USA	-	Roossinck Lab
C. annuum	Marengo	USA	-	Roossinck Lab
C. annuum	Jalapeño	USA	-	Dr. Valverde
C. annuum	Jalapeño	USA	-	Dr. Valverde
C. annuum	Jimmy Nardello's	USA	-	Hudson Valley[4]
C. annuum	Serrano Hot	USA	-	Hudson Valley
C. annuum	Hinkelhatz Hot	USA	-	Hudson Valley
C. annuum	Feher ozon paprika	USA	-	Hudson Valley
C. annuum	Serrano Hot	USA	-	Hudson Valley
C. annuum	Joe's long cayenne	USA	-	Hudson Valley
C. annuum	Anaheim Hot	USA	-	Hudson Valley
C. annuum	Matchbox	USA	-	Hudson Valley
C. annuum	Padron Hot	USA	-	Hudson Valley
C. annuum	Long red cayenne	USA	-	Victory Seeds[5]
C. annuum	Greek Pepperoncini	USA	-	Victory Seeds
C. annuum	Hungarian YW	USA	-	Victory Seeds
C. annuum	Large red Cherry	USA	-	Victory Seeds
C. annuum	Neapolitan	USA	-	Victory Seeds
C. annuum	Peter	USA	-	Victory Seeds
C. annuum	Chile negro	Mexico	PI 511882	USDA-GRIN[6]
C. annuum	WTS-32	Ecuador	PI 595906	USDA-GRIN
C. annuum	Jalapeño criollo	Guatemala	PI 666462	USDA-GRIN
C. annuum	Chocolate	Guatemala	PI 666471	USDA-GRIN
C. annuum	Blanco	Guatemala	PI 666536	USDA-GRIN
C. annuum	Ember	USA	PI 273426	USDA-GRIN
C. annuum	-	Venezuela	-	Dr. Orzolek
C. annuum	Hot	Venezuela	-	Dr. Orzolek
C. annuum	-	Ukraine	-	Dr. Orzolek
C. annuum	-	Ukraine	-	Dr. Orzolek
C. annuum	-	Ukraine	-	Dr. Orzolek
C. annuum	-	Ukraine	-	Dr. Orzolek
C. annuum	Pimento	USA	-	Dr. Orzolek
C. annuum	-	USA	-	Dr. Orzolek
C. annuum var.	Chiltepe	Guatemala	PI 632932	USDA-GRIN

glabriusculum				
C. annuum var. glabriusculum	Chiltepin	Mexico	-	Dr. GarciaArenal
C. annuum var. glabriusculum	Chiltepin	Mexico	-	Dr. GarciaArenal
C. annuum var. glabriusculum	Chiltepin	Mexico	-	Dr. GarciaArenal
C. annuum var. glabriusculum	Chiltepin	Mexico	-	Dr. GarciaArenal
C. annuum var. glabriusculum	Chiltepin	Mexico	-	Dr. GarciaArenal
C. annuum var. glabriusculum	Chiltepin	Mexico	-	Dr. GarciaArenal
C. annuum var. glabriusculum	Chiltepin	Mexico	-	Dr. GarciaArenal
C. baccatum	Lemon drop hot	USA	-	Hudson Valley
C. baccatum var. pendulum	WTS-14	Ecuador	PI 595905	USDA-GRIN
C. baccatum var. pendulum	WWMC 126	Paraguay	PI 632927	USDA-GRIN
C. baccatum var. pendulum	WW 141	Paraguay	PI 633756	USDA-GRIN
C. baccatum var. pendulum	MC 145	Paraguay	PI 633757	USDA-GRIN
C. baccatum var. pendulum	MC 147	Paraguay	PI 633758	USDA-GRIN
C. baccatum var. pendulum		Argentina	PI 337522	USDA-GRIN
C. baccatum var. baccatum		USA	PI 337524	USDA-GRIN
C. baccatum var. pendulum	BGH 4215	Brazil	PI 441589	USDA-GRIN
C. baccatum var. baccatum	470	Peru	PI 215699	USDA-GRIN
C. baccatum var. baccatum	No.1553	Bolivia	PI 238061	USDA-GRIN
C. baccatum var. pendulum	1SCA	Ecuador	PI 257135	USDA-GRIN
C. baccatum var. pendulum	Malagueta	Brazil	PI 260543	USDA-GRIN
C. baccatum var. pendulum	Valentine	Peru	PI 260549	USDA-GRIN
C. baccatum var. pendulum	Omnicolor	Peru	PI 260590	USDA-GRIN
C. baccatum var. baccatum	WWMC 136	Paraguay	PI 633752	USDA-GRIN
C. chinense	Habanero Red	USA	-	Victory Seeds
C. chinense	Chile blanco	Mexico	PI 574545	USDA-GRIN
C. chinense	WWT-1322	Ecuador	PI 593919	USDA-GRIN
C. chinense	30034	Belize	PI 594139	USDA-GRIN
C. chinense	Habanero naranja	Guatemala	PI 666547	USDA-GRIN
C. chinense	Habanero rojo	Guatemala	PI 666556	USDA-GRIN
C. chinense	Blanco	Guatemala	PI 666563	USDA-GRIN
C. chinense	30040	USA	PI 159236	USDA-GRIN
C. chinense	Scarlet Latern	Peru	PI 315008	USDA-GRIN
C. chinense	Royal gold	Peru	PI 315023	USDA-GRIN
C. chinense	Lemon drop	Peru	PI 315024	USDA-GRIN
C. frutescens	Aribibi gusano	Bolivia	PI 573337	USDA-GRIN

C. frutescens	WWT-1323	Ecuador	PI 593920	USDA-GRIN
C. frutescens	chile nan	Guatemala	PI 631144	USDA-GRIN
C. frutescens	Blanco	Guatemala	PI 666579	USDA-GRIN
C. frutescens	Tolito	Guatemala	PI 666580	USDA-GRIN
C. frutescens	Santo Domingo	Guatemala	PI 666593	USDA-GRIN
C. frutescens	Tabasco	USA	PI 586675	USDA-GRIN
C. frutescens	Greenleaf Tabasco	USA	PI 634826	USDA-GRIN
C. frutescens	WWT-1336	Ecuador	PI 593924	USDA-GRIN
C. frutescens	Pima	Vanuatu	PI 639661	USDA-GRIN
C. frutescens	Diente de perro	Guatemala	PI 631142	USDA-GRIN
C. frutescens	Diente de perro	Guatemala	PI 632917	USDA-GRIN
C. frutescens	Chiltepe	Guatemala	PI 632918	USDA-GRIN
C. frutescens	Tabasco-AVRDC	Mexico	PI 645556	USDA-GRIN
C. frutescens	Tabasco L-167	USA	PI 640909	USDA-GRIN
C. frutescens	Diente de perro	Guatemala	PI 666581	USDA-GRIN
C. frutescens	Diente de perro	Guatemala	PI 666589	USDA-GRIN
C. frutescens	Habanero	Mexico	-	Dr. Bello-Bedoy
C. pubescens	Ecu 2260	Ecuador	PI 585264	USDA-GRIN
C. pubescens	Ecu 2262	Ecuador	PI 585265	USDA-GRIN
C. pubescens	Ecu 2263	Ecuador	PI 585266	USDA-GRIN
C. pubescens	Ecu 2272	Ecuador	PI 585273	USDA-GRIN
C. pubescens	Ecu 2243	Ecuador	PI 585277	USDA-GRIN
C. pubescens	80018	Guatemala	PI 593616	USDA-GRIN
C. pubescens	80020	Guatemala	PI 593617	USDA-GRIN
C. pubescens	Ecu 2252	Ecuador	PI 585262	USDA-GRIN
C. pubescens	80072	Guatemala	PI 593642	USDA-GRIN
C. chacoense	-	-	-	Dr. Stephenson
C. chacoense	-	-	-	Dr. Stephenson

1. The locations that seeds were collected from.
2. Plant identification number
3. The source that seeds were obtained
4, 5. Commercial seed companies
6. USDA- Germplasm Resources Information Network (GRIN)

Table **S2:** The accession numbers of Endornavirus sequences deposited in GenBank

Names in the tree[1]	Domain	Accession number
BPEV_Marengo2	-	MG545489
BPEV_Marengo1	RdRp	MG545530
BPEV_Marengo1	Hel	MG545496
BPEV_Jimmy Nardello's	RdRp	MG545532
BPEV_Jimmy Nardello's	Hel	MG545498
BPEV_Feher ozon paprika	RdRp	MG545535
BPEV_Feher ozon paprika	Hel	MG545501
BPEV_Joe's long cayenne	RdRp	MG545531
BPEV_Joe's long cayenne	Hel	MG545497
BPEV_Greek pepperoncini	RdRp	MG545534
BPEV_Greek pepperoncini	Hel	MG545500
BPEV_ Hungarian YW	RdRp	MG545533
BPEV_ Hungarian YW	Hel	MG545499
BPEV_Padron Hot	RdRp	MG545528
BPEV_Padron Hot	Hel	MG545494
BPEV_Neapolitan	RdRp	MG545529
BPEV_Neapolitan	Hel	MG545495
BPEV_Peter	RdRp	MG545527
BPEV_Peter	Hel	MG545493
BPEV_Chocolate	RdRp	MG545536
BPEV_Chocolate	Hel	MG545502
CFEV1_Habanero Red	RdRp	MG545517
CFEV1_Chile blanco	RdRp	MG545521
CFEV1_30040	RdRp	MG545524
CFEV1_30040	Hel	MG545492
CFEV1_Royal gold	RdRp	MG545509
CFEV1_Lemon drop	RdRp	MG545516
CFEV1_Aribibi gusano	RdRp	MG545523
CFEV1_Greenleaf tabascoAL	RdRp	MG545519
CFEV1_Greenleaf tabascoAL	Hel	MG545491
CFEV1_Tabasco L-167	RdRp	MG545508
CFEV1_Tabasco L-167	Hel	MG545490
CFEV1_Habanero	RdRp	MG545518
CFEV1_Ecu	RdRp	MG545520
CFEV1_Lemon drop hot	RdRp	MG545515
CFEV1_WTS-14	RdRp	MG545506
CFEV1_WWMC126	RdRp	MG545504
CFEV1_WW141	RdRp	MG545505
CFEV1_MC145	RdRp	MG545513
CFEV1_MC147	RdRp	MG545512
CFEV1_Malagueta	RdRp	MG545514
CFEV1_Valentine	RdRp	MG545507
CFEV1_Omnicolor	RdRp	MG545510
CFEV1_82	RdRp	MG545526
CFEV1_BGH 4215	RdRp	MG545522
CFEV1_470	RdRp	MG545525

CFEV1_No 1553	RdRp	MG545511
HPEV_Ember	RdRp	MG545503

[1] Virus name_cultivar or plant designation.

Table **S3**: GenBank accession number for sequences used in phylogenetic trees

Name in tree[1]	Pepper type	Virus	Location[2]	Accession number
BPEV_Yolo Wonder	Sweet	BPEV	USA	JN019858
BPEV_Maor	Sweet	BPEV	USA	KP455654
BPEV_lj	Sweet	BPEV	China	KF709944
BPEV_IS	Sweet	BPEV	Israel	JQ951943
BPEV_Santa Fe	Pungent	BPEV	Colombia	KX977568
BPEV_Penol	Sweet	BPEV	Colombia	KX977569
BPEV_Kyosuzo	Sweet	BPEV	Japan	AB597230
BPEV_Healey	Sweet	BPEV	Canada	KT149366
BPEV_TW	Pungent	BPEV	Taiwan	KU923756
BPEV_PJ	Pungent	BPEV	India	KU923755
HPEV_CS	Pungent	HPEV	South Korea	KR080326
PvEV2	-	PvEV 2	-	AB719398

[1] Virus name_cultivar of plant designation.
[2] Location of the pepper isolate from which the sequence was derived.

Table **S4**: Peppers used for KASP analysis

Name [1]	Species	Location [2]	PI [3]
Marengo	*C. annuum*	USA	-
Jalapeno	*C. annuum*	USA	-
Serrano hot	*C. annuum*	USA	-
Long red cayenne	*C. annuum*	USA	-
Jalapeno criollo	*C. annuum*	Guatemala	PI 666462
Jimmy Nardello's	*C. annuum*	USA	-
Feher ozon paprika	*C. annuum*	USA	-
Joe's long cayenne	*C. annuum*	USA	-
Greek pepperoncini	*C. annuum*	USA	-
Hungarian YW	*C. annuum*	USA	-
Padron hot	*C. annuum*	USA	-
Peter	*C. annuum*	USA	-
Chocolate	*C. annuum*	Guatemala	PI 666471
Ember	*C. annuum*	USA	PI 273426
Chiltepin30	*C. annuum var. glabriusculum*	Guatemala	PI 632932
Chiltepin53	*C. annuum var. glabriusculum*	Mexico	-
Chiltepin54	*C. annuum var. glabriusculum*	Mexico	-
Chiltepin102	*C. annuum var. glabriusculum*	Mexico	-
Habanero Red	*C. chinense*	USA	-
Chile blanco	*C. chinense*	Mexico	PI 574545
30040	*C. chinense*	USA	PI 159236
Royal gold	*C. chinense*	Peru	PI 315023
Lemon drop	*C. chinense*	Peru	PI 315024
Aribibi gusano	*C. frutescens*	Bolivia	PI 573337
Greenleaf tabascoAL	*C. frutescens*	USA	PI 634826
Tabasco L-167	*C. frutescens*	USA	PI 640909
Ecu	*C. pubescens*	Ecuador	PI 585262
Lemon drop hot	*C. baccatum*	USA	-
WWMC126	*C. baccatum* var. *pendulum*	Paraguay	PI 632927
WW141	*C. baccatum* var. *pendulum*	Paraguay	PI 633756
MC145	*C. baccatum* var. *pendulum*	Paraguay	PI 633757
Malagueta	*C. baccatum* var. *pendulum*	Brazil	PI 260543
Valentine	*C. baccatum* var. *pendulum*	Peru	PI 260549
Omnicolor	*C. baccatum* var. *pendulum*	Peru	PI 260590
BGH 4215	*C. baccatum* var. *pendulum*	Brazil	PI 441589
470	*C. baccatum* var. *baccatum*	Peru	PI 215699
No 1553	*C. baccatum* var. *baccatum*	Bolivia	PI 238061

[1] Cultivar name or other reference name
[2] Location of the seed source
[3] Plant Identification number

Table **S5**. SNPs information

SNP #	SNP	Map_Chr [1]	Annotation
1	T/C	CHR08	Aquaporin, MIP family, NIP subfamily
2	T/C	CHR08	Chloroplast polyphenol oxidase
3	A/C	CHR08	Acetyltransferase, putative
4	T/G	Psuedo1or8	Sentrin/sumo-specific protease, putative
5	A/C	Psuedo1or8	Copper binding protein 3
6	T/C	Psuedo1or8	Pyruvate, phosphate dikinase, chloroplastic
7	A/C	Psuedo1or8	Heat shock protein
8	A/G	CHR01	Protein transport protein SEC23
9	A/G	CHR01	Expressed protein
10	T/C	CHR01	Detected protein of confused Function
11	A/G	CHR01	Putative transposon MuDR mudrA-like protein, identical
12	T/C	CHR01	Detected protein of confused Function
13	T/G	CHR01	Detected protein of confused Function
14	A/G	CHR01	Cytochrome P450
15	T/G	CHR02	Detected protein of unknown function
16	A/G	CHR02	Short-chain dehydrogenase, putative
17	T/C	CHR02	Hop-interacting protein THI034
18	A/G	CHR02	Class III HD-Zip protein 4
19	T/C	CHR02	Putative growth regulator
20	A/C	CHR02	DAG protein
21	T/G	CHR02	2-oxoglutarate-dependent dioxygenase (Fragment)
22	T/C	CHR02	Basic leucine zipper transcription factor
23	T/C	CHR02	Chloroplast ferredoxin-NADP+ oxidoreductase (Precursor)
24	A/G	CHR02	Ser/Thr protein kinase

25	A/G	CHR02	Detected protein of confused Function
26	C/T	CHR02	Kip-related protein
27	T/C	CHR02	Detected protein of unknown function
28	A/G	CHR02	Probable galacturonosyltransferase 13
29	T/C	CHR02	Non-cell-autonomous protein pathway1
30	T/C	CHR02	Calcium homeostasis regulator CHoR1
31	A/C	CHR03	Pectinesterase
32	T/C	CHR03	Putative receptor kinase-like protein, identical
33	T/C	CHR03	GRAS family transcription factor
34	T/C	CHR03	Detected protein of unknown function
35	A/G	CHR03	Cyclin-L1, putative
36	A/G	CHR03	Cation efflux protein/ zinc transporter, putative
37	T/C	CHR03	Chloroplast Trx
38	T/C	CHR03	Detected protein of unknown function
39	T/C	CHR03	Detected protein of confused Function
40	T/G	CHR03	Heavy metal transport/detoxification domain-containing protein
41	A/G	CHR03	Detected protein of unknown function
42	T/C	CHR03	Replication factor-A protein 1
43	A/G	CHR03	Serine/threonine-protein kinase ATM
44	T/C	CHR03	Phosphatidic acid phosphatase-related protein
45	A/G	CHR03	RNA polymerase II mediator complex subunit, putative
46	A/G	CHR03	Histidine-containing phosphotransfer protein, putative
47	T/G	CHR04	Glutamate receptor
48	T/G	CHR04	Pentatricopeptide repeat-containing protein
49	A/G	CHR04	Detected protein of unknown function
50	T/G	CHR04	Blue copper protein, putative
51	A/G	CHR04	ATP binding protein, putative
52	T/C	CHR04	Inositol-1,4,5-triphosphate-5-phosphatase

53	A/C	CHR04	Os02g0537900 protein
54	T/C	CHR04	Citrate-binding protein
55	A/G	CHR04	60S acidic ribosomal protein P0
56	T/C	CHR04	Calcium-dependent protein kinase
57	T/C	CHR04	Water channel protein
58	T/G	CHR04	Detected protein of unknown function
59	A/G	CHR04	Regulator of nonsense transcripts-like protein
60	T/C	CHR04	Purine transporter, putative
61	T/C	CHR04	26S proteasome regulatory subunit S3, putative
62	T/C	CHR05	Strictosidine synthase-like protein
63	T/C	CHR05	Glucan endo-1,3-beta-glucosidase
64	A/G	CHR05	MYB-like DNA-binding protein
65	A/G	CHR05	Glycogenin-1
66	T/C	CHR05	Alpha-glucosidase
67	T/C	CHR05	Tobamovirus multiplication 1 homolog
68	A/G	CHR05	DnaJ homolog protein
69	A/G	CHR05	Putative RNA binding protein
70	T/C	CHR05	Structural maintenance of chromosomes 6 smc6, putative
71	A/G	CHR05	Hydrolase, putative
72	T/C	CHR05	Speckle-type POZ protein, putative
73	A/G	CHR05	Cellulose synthase-like protein CslE
74	A/G	CHR05	C3HL domain class transcription factor
75	A/G	CHR05	Zinc finger protein, putative
76	T/G	CHR06	Pentatricopeptide repeat-containing protein At2g29760, chloroplastic
77	A/G	CHR06	Protein DEK
78	A/G	CHR06	Expressed protein
79	T/C	CHR06	Detected protein of confused Function
80	T/C	CHR06	Patellin 1

81	T/C	CHR06	Detected protein of unknown function
82	T/C	CHR06	Triacylglycerol lipase, putative
83	A/C	CHR06	SWI/SNF complex subunit SMARCC2
84	T/G	CHR06	Detected protein of unknown function
85	G/A	CHR06	CCR4-NOT transcription complex subunit
86	A/G	CHR06	Bile acid Na+ symporter family protein
87	T/C	CHR06	Protein translocase, putative
88	A/G	CHR06	Detected protein of confused Function
89	T/C	CHR06	Detected protein of confused Function
90	T/C	CHR06	Detected protein of unknown function
91	T/C	CHR07	V-type proton ATPase subunit H
92	A/G	CHR07	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase
93	T/C	CHR07	Expressed protein
94	T/C	CHR07	Retrotransposon protein
95	T/C	CHR07	Protein disulfide isomerase family
96	A/G	CHR07	Protein binding protein, putative

[1] CHR, Chromosome number; Pseudo 1 or 8, on either chromosome 1 or 8, precise position not determined.

Table **S6**: The accession number for Partitivirus sequences deposited in GenBank

Name in tree[1]	RNA	Accession number
PCV1_Long red cayenne	RNA 1	MG542614
PCV1_Long red cayenne	RNA 2	MG542609
PCV1_Chiltepe30	RNA 1	MG542613
PCV1_Chiltepe30	RNA 2	MG542612
PCV1_Chiltepe53	RNA 1	MG542616
PCV1_Chiltepe53	RNA 2	MG542611
PCV1_ Chiltepe54	RNA 1	MG542615
PCV1_ Chiltepe54	RNA 2	MG542610
PCV2_Hinkelhatz Hot	RNA 1	MG542633
PCV2_Hinkelhatz Hot	RNA 2	MG542622
PCV2_Long red cayenne	RNA 1	MG542632
PCV2_Neapolitan	RNA 1	MG542631
PCV2_Neapolitan	RNA 2	MG542621
PCV2_Jalapeño criollo	RNA 1	MG542630
PCV2_Jalapeño criollo	RNA 2	MG542620
PCV2_Ember	RNA 1	MG542626
PCV2_Ember	RNA 2	MG542618
PCV2_Chiltepin53	RNA 1	MG542629
PCV2_Chiltepin53	RNA 2	MG542623
PCV2_Chiltepin54	RNA 1	MG542628
PCV2_Chiltepin55	RNA 1	MG542627
PCV2_Chiltepin55	RNA 2	MG542619
PCV2_Chiltepin102	RNA 1	MG542624
PCV2_Chiltepin102	RNA 2	MG542617
PCV2_Diente de perro	RNA 2	MG542625

[1] Virus name_cultivar or plant designation.

Table **S7**: GenBank accession number for sequences used in partitivirus phylogenetic trees

Name in tree[1]	RNA	Location[2]	Genus	Accession number
PCV1_Jalapeño	1	USA	*Deltapartitivirus*	JN117276.1
PCV1_Jalapeño	2	USA	*Deltapartitivirus*	JN117277.1
PCV1_CHN	1	China	*Deltapartitivirus*	KX765307.1
PCV1_CHN	2	China	*Deltapartitivirus*	KX765306.1
PCV2_HW	1	USA	*Deltapartitivirus*	JN117278.1
PCV2_HW	2	USA	*Deltapartitivirus*	JN117279.1
PCV2_YY	1	South Korea	*Deltapartitivirus*	LC195294.1
PCV2_YY	2	South Korea	*Deltapartitivirus*	LC195295.1
PCV2_CQ	1	China	*Deltapartitivirus*	KX905077.1
PCV2_CQ	2	China	*Deltapartitivirus*	KX905078.1
PCV2_DR	1	Dominican Republic	*Deltapartitivirus*	KX525268
PCV2_DR	2	Dominican Republic	*Deltapartitivirus*	KX525269.1
PitCV1	1	Italy	*Deltapartitivirus*	LN680393.2
PitCV1	2	Italy	*Deltapartitivirus*	LN680394.2
RsCV3	1	China	*Deltapartitivirus*	FJ461349.1
RsCV3	2	China	*Deltapartitivirus*	FJ461350.1

[1] Virus name_cultivar of plant designation.
[2] Location of the pepper isolate from which the sequence was derived.

VITA

Maliheh Safari

EDUCATION:

Ph.D. Plant Pathology & Environmental Microbiology, Penn State University
M.Sc. Plant Pathology, University of Tehran
B.Sc. Plant Protection, University of Tehran

AWARDS:

Lester P. Nichols Memorial Award: College of Agricultural Science, Penn State, 2017-2018.
Second place poster presentation: 32nd Annual Graduate Exhibition, Penn State, 2017.
Graduate Student Travel Awards: PPEM, Penn State, April 2016, February 2017, May 2017.
Lester P. Nichols Memorial Award: College of Agricultural Science, Penn State, 2014-2015.

SELECTED PUBLICATIONS/PRESENTATIONS:

Journals:

M. Safari, MJ. Roossinck. Evolution of endornaviruses in *Capsicum* spp. Under review.
M. Safari, MJ. Roossinck. A virus benefits its host by manipulating aphid's behavior. Under review.
M. Safari, MJ. Roossinck. Evolution in *Partitiviridae* family. In preparation.
M. Safari, MJ. Roossinck. How does the genome structure and lifestyle of a virus affect its population variation? Current Opinion in Virology, 9 (2014): 39-44.
M. Safari, et al. Inhibition of spore development and mycelial growth of *Rhizopus stolonifer* by essential oil of *Satureja richingeri* from Iran, Journal of Essential Oil Research, 23 (2011): 5–10.

Oral presentation:

M. Safari, MJ. Roossinck. Co-evolution of Two Persistent Viruses with Their Plant Hosts. 36th Annual meeting of American Society of Virology, June 2017.
M. Safari, MJ. Roossinck. How a persistent virus has evolved in *Capsicum* spp. 9th Virus Evolution Workshop, March 2017.
M. Safari, MJ. Roossinck. Evolution of *Bell pepper endornavirus* in peppers. 35th Annual meeting of American Society of Virology, June 2016.
M. Safari, MJ. Roossinck. Evolution of a persistent virus in different pepper cultivars. 19th Environmental Chemistry and Microbiology Student Symposium. April 2016.

Seminar Talks:

Persistent viruses in peppers: their evolution and aphid behavior modification. Cornell University, April 2017.
Evolution of an endornavirus in *Capsicum* spp. Virology@PSU, Penn State, April 2017.
Evidence for slow evolution in an RNA virus, and its modification of a vector behavior. Center of Infectious Disease Dynamics. Penn State, April 2017
Plant persistent viruses, genome structure and population diversity. Virology @PSU, March 2015.