Detection of Single-Stranded DNA Viruses in Insects and Wild Vegetation

Marco Padilla-Rodriguez

A thesis submitted in partial fulfillment
of the requirements of the
University Honors Program
University of South Florida - St. Petersburg

May 7th 2012

Thesis Director: Mya Breitbart Ph.D.
Assistant Professor, College of Marine Science

Committee Member: Karyna Rosario Ph.D.
Post-Doctoral Research Assistant, College of Marine Science

For my loving parents
Inger Padilla and John Rodriguez

And my loving godparents
Susan and Raegen Carr

My life is beautiful because of them

- Abstract -

The smallest viruses known to infect eukaryotic organisms contain circular, single-stranded (ssDNA) genomes. Viruses with these characteristics infect economically important animals (*Circoviridae*) and food crops (*Geminiviridae*). Circoviruses were traditionally thought to only infect vertebrates; however, recent studies have characterized circoviruses in dragonflies, suggesting that these viruses may be widespread within insect populations. Since there has not been an effort to further characterize circoviruses in these populations, the first objective of this study is to survey various insect species for the presence of circoviruses. A Polymerase Chain Reaction (PCR) assay revealed a novel circovirus genome (~1700 nucleotides) in a *Eurycotis floridana* (palmetto bug) sampled from Tarpon Springs, Fl, which had a 70% nucleotide identity to circoviruses found in bat and human feces . By describing the second circovirus from insects, this study supports the idea that circoviruses are more widespread in invertebrates than previously recognized. While the presence of circoviruses in insects has only recently been discovered, the role of insect vectors in the transmission of geminiviruses is well established. The feeding behavior of these insect vectors, which feed on both wild and cultivated vegetation, may lead to the transmission and emergence of new virulent plant virus strains. Therefore, surveying wild plants for viruses may serve as a preventive approach for controlling geminivirus-related outbreaks. For this reason, the second objective of this study is to survey under-sampled plants, specifically weeds and grasses, for the presence of geminiviruses. Rolling-circle amplification (RCA) followed by restriction enzyme digest revealed a known Euphorbia mosaic virus (EuMV) in a *Euphorbia heterophylla* (a wild weed) sampled from Homestead, Florida. This is the first time that the EuMV genome has been sequenced from the United States, as this virus is primarily found in the Caribbean and Mexico. Differences between the Florida EuMV genomic components, DNA-A and DNA-B, and those reported from the Caribbean and Mexico suggests that these two components are exposed to different evolutionary pressures.

- Introduction -

Viruses are the most abundant and genetically diverse biological entities on earth, found in all examined ecosystems. These obligate intracellular parasites, which rely on a host to replicate, have important ecological and societal effects. The majority of viruses are phages (viruses that infect bacteria). In marine ecosystems, viruses play an important role in the microbial loop, in which unicellular marine organisms cycle energy and dissolved organic matter through the food web (Wilhelm *et al.* 1999). In these ecosystems, virus-mediated infection plays a critical role in the microbial loop by releasing stored carbon and nutrients from microorganisms (Wilhelm *et al.* 1999).

Although viruses play an important role in ecosystem function, they are mainly known as pathogenic agents. Viral diseases are of great concern to public health agencies and the agricultural sector. Health agencies, such as the Centers for Disease Control and Prevention and the U.S. Department of Health and Human Services, allocate significant resources for viral surveillance and disease management. Influenza epidemics, for example, are estimated to cost the U.S. economy $71-167 billion per year (World Health Organization 2003). Since viruses cause devastating diseases in all staple food crops, including potatoes, maize, rice, barley, and wheat, they represent a threat to food security (Rojas *et al.* 2005). In addition, virus-related illnesses cause significant losses to livestock (Crowther 1999). Agricultural institutions have focused research efforts on understanding and managing viral diseases, costing millions of dollars every year in the U.S. alone. Viruses seriously affect plant, animal, and human health, thus burdening the global economy.

Traditionally, virological research occurred in reaction to epidemics affecting

human health and food production. However, over the past ten years, the scientific community has realized the importance of monitoring environmental reservoirs for effective prevention. This approach is known as ecosystem-based surveillance, in which viruses are monitored in reservoirs (e.g., water), and vectors (e.g., mosquitoes) (Epstein 1995). The morphological characteristics and high genetic diversity of viruses make their study difficult. Viral particles are remarkably small (on the scale of nanometers) and microscopy techniques provide limited resolution. Since viruses need a host to replicate, culturing techniques are limited to viruses whose hosts can be grown in the lab. Therefore most virological research has been performed through molecular techniques. Viruses can carry either DNA or RNA within their protein capsids and, unlike cellular organisms, do not have a universal genetic marker (Campbell *et al*. 2008). Due to the absence of a genetic marker, molecular techniques, such as Polymerase Chain Reaction (PCR), are limited to specific groups of viruses. Therefore most investigative techniques require prior knowledge of the targeted viruses.

The smallest viruses known to infect eukaryotic organisms contain circular, single-stranded (ssDNA) genomes. These viruses are known to infect plants, animals, and humans and are classified under four families: *Circoviridae, Nanoviridae, Anelloviridae,* and *Geminiviridae*. Members of the *Nanoviridae* and *Geminiviridae* families infect plants, while *Circoviridae* and *Anelloviridae* infect vertebrates. Unlike circoviruses and anelloviruses, plant pathogens may have multipartite genomes. *Nanovirus* genomes contain 6 to 11 components that are individually encapsidated, whereas geminiviruses may have one or two components (Shirasawa-Seo *et al*. 2004, Rojas *et al*. 2005). Most of the research on geminiviruses, nanoviruses, and

circoviruses has been done by the agricultural sector, as these viruses infect staple crops and livestock (Li *et al.* 2010, Rojas *et al.* 2005). However, recent studies suggest that these viruses are more widespread than previously recognized.

- Research Background and Objectives -

The research presented in this thesis focuses on investigating the diversity of ssDNA animal viruses (i.e., circoviruses) and plant pathogens (i.e., geminiviruses) in previously unexplored hosts and geographical areas, respectively.

One goal of this study was to assess the presence of novel circoviruses in insects, which have been overlooked as potential hosts. Circoviruses are the smallest viral pathogens (17-20 nm) known to infect eukaryotic organisms. Circovirus genomes range from 1700-2000 nucleotides (nt) and contain two major open reading frames (ORFs), encoding for capsid (Cap) and replication (Rep) proteins (Todd 2000). Known circoviruses affect a variety of birds and mammals and through causing immunosuppression may be fatal in swine (pigs) and psittacines (parrots). Due to their economic importance, a majority of circovirus research has focused on viral transmission and the biological effects brought on by infection. In parrots, Psittacine Beak and Feather Disease Virus (BFDV) causes birds to develop a deformed and brittle beak, which may lead to starvation. It also causes birds to loose a large amount of feathers, leading to hypothermia (Todd 2000). In swine, Porcine Circovirus (PCV) causes Post-Weaning Multisystematic Waste Syndrome, which allows for additional infections as well as malnutrition/dehydration caused by prolonged diarrhea (Allan and Ellis 2000). Since most circoviruses cannot be propagated in cultures, molecular analyses have been used to further understand this group of viruses.

Novel variants of *Circoviridae* have been found in the feces of a number of mammals including chimpanzees, bats, chickens, humans, and in the tissues of fish, sheep, cows, and camels (Blinkova *et al.* 2009, Ge *et al.* 2011, Lörincz *et al.* 2011). Some of these novel viruses were classified in the proposed *Cyclovirus* genus (Li *et al.* 2010). Inquiry into the presence of novel circoviruses in hosts other than vertebrates has been extremely limited. Metagenomic analyses have recently expanded the host range of circoviruses (Li *et al.* 2010, Rosario *et al.* 2009), and in 2011, the first circovirus was found within invertebrates, specifically dragonflies, suggesting that members of the *Circoviridae* family may infect insects (Rosario *et al.* 2011). Since there has not been an effort to further investigate the diversity of circoviruses in insects, the first objective of this study is to survey insects for novel circoviruses and completely sequence their genomes.

While the presence of circoviruses in insects has only recently been discovered, the role of insect vectors in the transmission of geminiviruses is well established. First characterized in the late 1970's, geminivirus-related diseases are currently some of the most economically devastating, infecting important food crops such as maize, tomatoes, beans, cotton, and cassava (Rojas *et al.* 2005). For example, the continent of Africa experiences estimated losses of $1.9-2.7 billion per year due to the pandemic of cassava mosaic disease (Tiendrébéogo *et al.* 2012). Historically, the range of geminiviruses has been limited to the geographical distribution of their respective insect-vectors. Today, the reach of geminiviruses is associated with human activities, such as the transportation of plant material across continents. Agricultural practices, such as monocultures and the use of pesticides, have exacerbated geminivirus-related diseases

within local environments (Rojas *et al.* 2005).

The *Geminiviridae* family includes four genera, *Begomovirus*, *Curtovirus*, *Mastrevirus,* and *Topocuvirus.* Begomoviruses can exist as both monopartite and bipartite genomes, while curtoviruses*,* mastreviruses*,* and topocuviruses are exclusively monopartite. The bipartite components of begomoviruses, DNA-A and DNA-B, efficiently establish systematic infection of dicot plant species. The DNA-A component encodes for replication, transcription, and encapsidation, while the DNA-B component is responsible for viral movement and symptom development. Both genome components are similar in size (~2500-3000 nt) with the total genome size ranging from 2500-5000 nt (Rojas *et al*. 2005).

Insect vectors are the primary mode of transmission for geminiviruses. Since vectors, such as whiteflies and hoppers, can feed on a variety of food sources including food crops and wild vegetation, the feeding behavior of these insects may lead to the emergence of new virulent strains by moving geminiviruses to and from wild vegetation. Therefore, the second goal of this study investigates the presence of geminiviruses in wild vegetation samples, such as weeds and grasses, as these may serve as a reservoir and potential source for recombination leading to potential viral outbreaks within food crops.

- Methods -

Sample Processing and DNA Extraction

A variety of different insects and plants samples were tested for the presence of circoviruses and geminiviruses respectively. Insect samples were dissected, placed into bead beating tubes, and 1 ml of SM buffer (0.1 M NaCl, 50 mM Tris-HCl (pH 7.4), 10

mM $MgSO_4$) was added for homogenization. Small insects were homogenized whole, while only chosen segments were processed for larger insects. Samples were homogenized for two minutes in a bead-beater and centrifuged for 6 minutes at 8000xg. The viral particles present in the supernatant were then partially purified through filtration. For this purpose, the supernatant was filtered through a 0.45 µm Sterivex™ filter using a 60 ml syringe. The filtrate containing viral particles was stored at 4°C until further processing. The QIAamp® MinElute DNA Spin Kit (Qiagen®) was used to extract DNA from ~200 µl of partially purified viral particles. For plant samples, DNA was directly extracted from ~100 mg of plant material using the DNeasy Plant Mini Kit (Qiagen®).

Detection of circoviruses via Polymerase Chain Reaction (PCR)

A nested PCR assay was carried out to detect the presence of circovirus-like sequences in DNA extracts using the degenerate primers developed by Li *et al.* 2010. The first PCR reaction consisted of 1.5 mM $MgCl_2$, 1X NH_4 Buffer, 0.5 µM CVF_1 primer, 0.5 µM CVR_1 primer, 0.2 mM dNTPs, 1 U Apex® Taq DNA Polymerase Enzyme, and 2 µl template (sample) DNA. Amplification with CVF/R_1 primers proceeds as follows: initial denaturation at 95°C for 5 min, followed by 40 cycles of denaturation at 95°C for 1 min, annealing at 52°C for 1 min, and extension at 72°C for 1 min, followed by a final extension at 72°C for 10 min. For the second round, 1 µL µl of PCR product from the first round was amplified using CVF_2 and CVR_2 primers. Amplification with CVF/R_2 primers remained the same as the previous amplification except for the annealing step, which was increased to 56°C for 1 min. PCR products were verified through electrophoresis on 1.5% agarose gels stained with ethidium bromide. Products

approximately 400 nt long were cloned using the TOPO TA cloning kit (Invitrogen™) and sequenced. These sequences were used to design back-to-back primers for inverse PCR to obtain complete genomes. The following back-to-back primers were used to obtain the full novel genome of the palmetto bug virus: PalmBug_3Fout (5'-AAGATGGATTGCGTTATCG-3') and PalmBug_3Rout (5'-TTGAAAAGGCAAATGGAT-3').

Detection of ssDNA viruses through Restriction Digest

Since PCR assays may not recover novel circoviruses or geminiviruses, restriction enzyme digest was selected as an alternative method. Through this approach, circular ssDNA genomes are amplified through rolling circle amplification (RCA) to obtain double-stranded (dsDNA). The RCA product can then be digested with restriction enzymes that cut the genome once to obtain linear, full length genomes that can be cloned and sequenced (Haible *et al.* 2006). To obtain dsDNA templates via RCA, insect and plant DNA extracts were amplified using the Illustra TempliPhi™ 100 Kit (GE Healthcare®). For each sample, 5 μl of sample buffer was mixed with 0.5 μl of template DNA. After incubating for 3 min at 95°C, mixture was cooled for 2 min on ice. Five microliters of a master mix, containing 5 μl of Reaction Buffer and 0.2 μl of Enzyme Mix, was added to the 5.5 μl mixture of sample buffer and template DNA from the previous step. This final mixture was incubated for 18 hours at 30°C followed by 10 min at 65°C. The TempliPhi™ products were digested using either XmnI or BamHI restriction enzymes (New England BioLabs®) in 30 μl reaction mixtures. This reaction included a master mix consisting of 21.7 μl of sterile H_2O, 3 μl of 10X Enzyme Buffer, 0.3 μl 100X BSA, 1 μl of XmnI or BamHI enzyme (20,000 U/mLml), and 4 μl of TempliPhi™ product.

The digestion reaction was incubated at 37°C for 2 hours and digested products were examined through electrophoresis on 1.5% agarose gels stained with ethidium bromide. Products larger than 1500 nt were excised from the gel and cleaned using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research). Cleaned products were cloned using the CloneJET™ pJet 1.2 cloning kit (Fermentas) and sequenced. If sequencing of digested bands did not result in a full genome, inverse PCR was preformed using primers designed from the digested fragment. The begomovirus DNA-A sequence reported in the results was obtained using Eub-A F1out (5'-CAACGCCGTCAAGAATTACA-3') and Eub-A R1 out primers (5'-GTGTCTTTGGCTCCTTGAA-3').

- Results and Discussion -

The focus of this study was to test the presence of ssDNA viruses in under-sampled hosts (insects and wild vegetation). In 2011, a circovirus was described in dragonfly samples (Rosario *et al.* 2011), the first characterized in invertebrates. For this reason, insects were chosen to assess the diversity of circoviruses, which is currently unknown. Additionally, wild grasses and weeds serve as an important environmental reservoir for new virulent strains, specifically geminiviruses (Rojas *et al.* 2005). Monitoring the viruses present in these environmental reservoirs is critical to understanding viral diversity and preventing viral outbreaks in staple food crops.

Circovirus Detection within Insect Samples

A total of 506 insect samples (Table 1) were processed and tested (via nested PCR assay) for circoviruses (product size ~400 nt). Initial findings showed circovirus-related sequences in four insect samples: a walking stick, waxworm, earthworm, and

palmetto bug. For each sample, inverse PCR was performed using back-to-back primers designed from the initial sequences to obtain a complete genome (~1700-2000 nt). From this assay, a complete circovirus genome (~1700 nt) was obtained from the palmetto bug sample. Sequences obtained from the walking sticks, waxworms, and earthworms had significant similarities to known circoviruses, but the inverse PCR assay did not work even though several primer sets were attempted. Since back-to-back primers only work to amplify circular genomes, the circovirus-like sequence fragments (~400 nt) found within these samples could either be a part of a linear viral genome or may be integrated within the genome of the organisms (Katzourakis *et al*. 2010).

The novel palmetto bug genome is depicted in Figure 1. Compared to the genomes of known circoviruses (e.g., Porcine Circovirus 1 genome) (AF071879) and cycloviruses (cyclovirus genome found in bat feces) (JF938079), the palmetto bug virus genome is consistent with cycloviruses. While all three genomes have a highly conserved, nonanucleotide sequence (stem loop motif) and two ORFs that code for encapsidation and replication, the ORFs of the PCV genome have a reversed orientation from those of the bat and palmetto bug viral genomes. In addition, both the bat and palmetto bug viral genomes have only one intergenic region (area of no ORFs). The completed novel palmetto bug viral genome showed a 70% identity at the nucleotide level to cycloviruses found in bat (GQ404854) and human (JF938079) feces over the Rep ORF. For these reasons, the sequenced palmetto bug viral genome is believed to be a novel cyclovirus, with the proposed name Palmetto bug cyclovirus GS140.

Circovirus and circovirus-like genomes have and continue to be described in previously un-sampled vertebrate species and environments through metagenomic analysis (Li *et al.* 2010, Rosario *et al.* 2009). However, the recent discovery of a novel circovirus in dragonflies (Rosario *et al.* 2011) along with this novel palmetto bug "cyclovirus" suggest that circoviruses may play a previously unrecognized role in insect populations. These findings may point to an evolutionary link between circoviruses found within vertebrates and invertebrates. Since investigation into the presence of circoviruses within insect populations has been insufficient, further epidemiological and bioinformatic analysis is needed.

Geminivirus Detection within Wild Vegetation Samples

A total of 56 wild grass and weed samples, all with visible signs of infection, were collected (Guanica, Puerto Rico), processed, DNA enriched for circular viral ssDNA (via RCA), and digested using a restriction enzyme. Along with these collected weeds and grasses, the Department of Plant Pathology (University of Florida) provided a *Euphorbia heterophylla* sample, a weed found throughout the tropical and subtropical regions of the Americas. The infected *E. heterophylla* tissue was the only positive sample revealed through restriction digest, using both XmnI and BamHI enzymes. The XmnI restriction enzyme (recognition site: GAANN↓NNTTC) digestion resulted in a complete (~2500 nt) DNA-B genomic component whereas the BamHI (recognition site: G↓GATCC) product represented a near-complete (~2200 nt) DNA-A component. Inverse PCR was used to complete the DNA-A component (~2600 nt). Both DNA components were similar to genomic sequences of the Euphorbia mosaic virus (EuMV). Figure 2 illustrates the DNA-A component of the completed *E. heterophylla* viral genome, which was 100%

13

identical to a EuMV sequenced from Cuba (HQ 896201), along with the DNA-B

component, which was 97% identical to a different EuMV sequenced from the same

region (FJ 807783).

Like other bipartite begomoviruses, the EuMV characterized in the *E. heterophylla*

sample is composed of two ssDNA components, DNA-A and DNA-B. The DNA-A

component has four ORFs: Rep, Coat (similar function as the Cap), transcriptional

activator protein (TrAP), replication enhancement protein (REn), and a protein that may

be involved in symptom expression (AC4). The TrAP ORF is believed to be responsible

for the suppression of gene silencing along with tissue-specific expression of the capsid

protein (Rojas *et al.* 2005, Settlage *et al.* 1996). The DNA-B component has two ORFs

responsible for viral movement across nuclear (NSP) and plasmodial (MP) membranes

(Rojas *et al.* 2005).

The range of begomoviruses is limited to the range of their whitefly vector, which

includes tropical and subtropical regions. While samples showing symptoms of EuMV

infection have been previously described in Florida (Kim *et al.* 1984), this study reports

the first sequenced EuMV genome from this region. The EuMV described in the *E.*

heterophylla sample has been a widespread problem throughout the Caribbean and

Central America (Rojas *et al.* 2005, Garrido-Ramirez *et* al. 1996). Table 2 shows the

genetic similarity between the characterized EuMV genome sampled from Florida and

the sequenced EuMV genomes collected from Cuba, Puerto Rico, Jamaica, and

Mexico. The DNA-A component is nearly identical to EuMVs from Puerto Rico, Jamaica,

and Mexico and completely identical to EuMVs in Cuba. The DNA-B component shows

greater genetic variation between these geographical regions. While DNA-A and DNA-B

are genomic components of the same virus, these findings support that they may be affected by different evolutionary pressures (Briddon *et al*. 2010).

- Concluding Remarks -

Pathogenic agents from the *Circoviridae* and *Geminiviridae* families infect economically important animals and food crops, respectively. The first portion of this study focused on circoviruses, which were thought to only infect vertebrates. Their recent characterization in dragonflies and now palmetto bugs suggests that they may be more widespread in insect populations than previously thought. These findings may point to an evolutionary link between circoviruses found within vertebrates and invertebrates; however, further epidemiological and bioinformatic analysis is needed.

The second objective of this study focused on plant pathogens from the *Geminiviridae* family. The geminivirus EuMV has been found throughout the Caribbean and certain parts of Mexico. This study reports the first sequenced EuMV genome from the United States, specifically Florida. Analyzing the similarities between the Florida EuMV genomic components, DNA-A and DNA-B, and those reported from the Caribbean and Mexico revealed differences between these components based on their geographic distribution. This suggests that these components are exposed to different evolutionary pressures.

The findings of this report emphasize the importance of studying under-sampled organisms for ssDNA viral strains, as they may reveal evolutionary links and environmental reservoirs for viruses pathogenic to economically important plants and animals.

- Acknowledgments -

- References -

Allen, G. M., Ellis, J. A. (2000) Porcine Circoviruses: A Review. Journal of Veterinary Diagnostic Investigation 12, 3-14.

Banda, A., Galloway-Haskins, R. I., Sandhu, T. S., Schat, K. A. (2007) Genetic Analysis of a Duck Circovirus Detected in Commercial Pekin Ducks in New York. Avian Diseases 51, 90-95.

Blinkova, O., Victoria, J., Li, Y., Keele, B. F., Sanz, C., Ndjango, J. N., Peeters, M., Travis, D., Lonsdorf, E. V., Wilson, M. L., Pusey, A. E., Hahn, B. H., Delwart. E. L. (2010) Novel Circular DNA Viruses in Stool Samples of Wild-Living Chimpanzees. Journal of General Virology 91, 74-86.

Briddon, R. W., Patil, B. W., Bagewadi, B., Nawaz-ul-Rehman, M. S. (2010) Distinct Evolutionary Histories of the DNA-A and DNA-B Components of Bipartite Begomoviruses. BMC Evolutionary Biology 10 (97).

Campbell, N. A., Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., Jackson, R. B. (2008) Biology. Pearson Education (8th ed) 381-88.

Collins, A. M., Roye, M. E. (2007) Two New Bipartite Begomoviruses Infecting *Wissadula amplissima* in Jamaica. Plant Pathology 56 (2), 340.

Crowther, J.R. (1999) Detection of Viruses in Livestock. Parasitology 117, 29-40.

Epstein, Paul R. (1995) Emerging Diseases and Ecosystem Instability: New Threats to Public Health. American Journal of Public Health 85 (2), 168-172.

Garrido-Ramirez, E.R., Gilbertson, R.L. (1996) Molecular Detection and Characterization of *Bean golden mosaic geminivirus* (BGMV) from Chiapas, Mexico. Phytopathology 86.

Ge, X., Li, J., Peng, C., Wu, L., Yang, X., Wu, Y., Zhang, Y., Shi, Z. (2011) Genetic Diversity of Novel circular ssDNAviruses in bats in China. Journal of General Virology, vir.0.034108-0.

Haible, D., Kober, S., Jeske, H. (2006) Rolling Circle Amplification Revolutionizes Diagnosis and Genomics of Geminiviruses. Journal of Virological Methods 135, 9-16.

Katzourakis, A., Gifford, R. J. (2010) Endogenous Viral Elements in Animal Genomes. Public Library of Science 6 (11).

Kim, K. S., Fulton R. W. (1984) Ultrastructure of *Datura stramonium* Infected with a Euphorbia Virus Suggestive of a Whitefly-Transmitted Virus. Phytopathology 74, 236.

Li, L., Kapoor, A., Slikas, B., Bamidele, O. S., Wang, C., Shaukat, S., Masroor, M. A., Wilson, M. L., Ndjango, J. N., Peeters, M., Gross-Camp, N. D., Muller, M. N., Hahn, B. H., Wolfe, N. D., Triki, H., Bartkus, J., Zaidi, S. Z., Delwart, E. (2010) Multiple Diverse Circoviruses Infect Farm Animals and are Commonly Found in Humans and Chimpanzee Feces. Journal of Virology 84, 1674-1682.

Lörincz, M. Cságola, A., Farkas, S. L., Székely, C., Tuboly, T. (2011) First Detection and Analysis of Fish Circovirus. Journal of General Virology, vir.0.031344-0.

Rojas, M. R., Hagen, C., Lucas, W. L., Gilbertson, R. L. (2005) Exploiting Chinks in the Plant's Armor: Evolution and Emergence of Geminiviruses. Annual Review of Phytopathology 43, 361-94.

Rosario, K., Duffy, S., Breitbart, M. (2009) Diverse Circovirus-like Genome Architectures Revealed by Environmental Metagenomics. Journal of General Virology 90, 2418-2424.

Rosario, K., Marinov, M., Stainton, D., Kraberger, S., Wiltshire, E. J., Collings D. A., Walters, M., Martin, D. P., Breitbart, M., Varsani, A. (2011) Dragonfly Cyclovirus, a Novel Single-Stranded DNA Virus Discovered in Dragonflies. Journal of General Virology 92, 1302-1306.

**Shirasawa-Seo, N., Sano, Y., Nakamura, S., Murakami, T., Seo, S., Ohashi, Y.,**

Hashimoto, Y., Matsumoto, T. (2005) Characteristics of the Promoters Derived from the Single-Stranded DNA Components of Milk Vetch Dwarf Virus in Transgenic Tobacco. Journal of General Virology 86, 1851-1860.

Todd, D. (2000) Circoviruses: Immunosupressive Threats in Avian Species: A Review. Avian Pathology 29, 373-394.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Harimalala, M. A., De Bruyn, A., Villemot, J., Traoré, V., Konaté, G., Traoré, A. S., Barro, N., Reynaud, B., Traoré, O., Lett, J. (2012) Evolution of *African Cassava mosaic virus* by Recombination Between Bipartite and Monopartite Begomoviruses. Virology Journal 9 (67).

Wilhelm, S. W., Suttle, C. A. (1999) Viruses and Nutrient Cycles in the Sea: Viruses Play Critical Roles in the Structure and Function of Aquatic Food Webs. Bio Science 49, 781-788.

World Health Organization (2003) Influenza Overview. http://www.who.int/mediacentre/factsheets/2003/fs211/en/
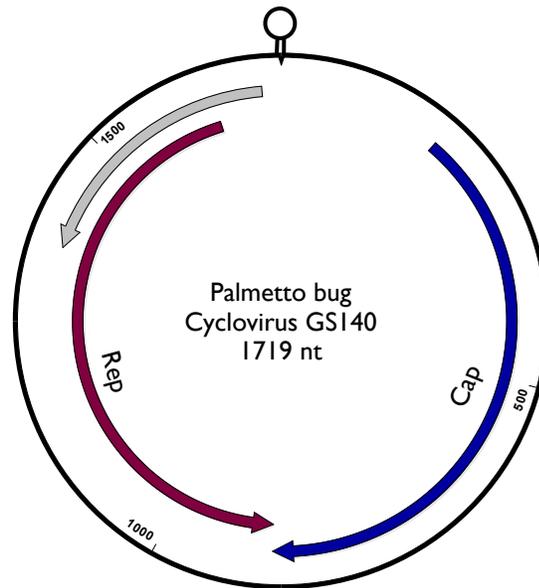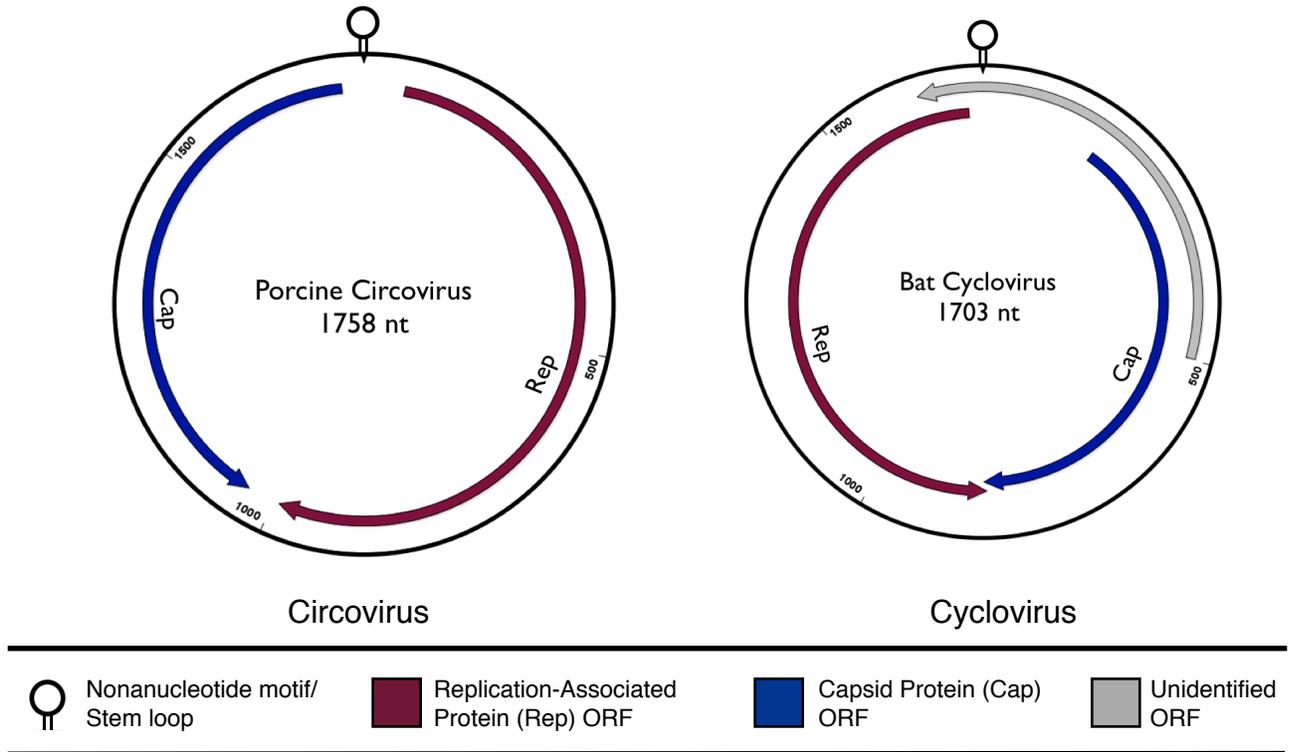
Figure 1: Schematic of genome organization of known circoviruses, cycloviruses, and the ssDNA virus discovered in this study. Each schematic shows the different open reading frames (ORFs) identified in each genome, including the replicating-associated (Rep) and capsid (Cap) proteins. Note that the virus discovered in palmetto bugs exhibits similar characteristics to known cycloviruses. The palmetto bug cyclovirus genome shares 64% identity with a cyclovirus identified in bat feces.
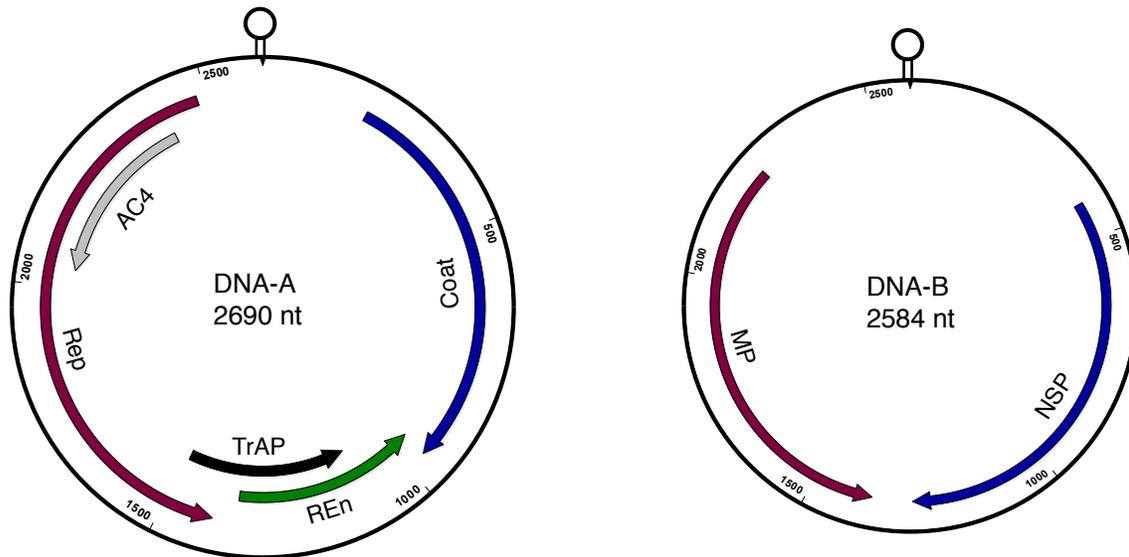
Figure 2: Schematic showing the Euphorbia mosaic virus (EuMV) genome sequenced from a *Euphorbia heterophylla* sample collected from Homestead, FL. EuMV contains a bipartite genome consisting of DNA-A and DNA-B components. DNA-A encodes for coat (CP), replication-associated (Rep), transcriptional activator (TrAP), and replication enhancement (REn) proteins. DNA-B encodes for movement (MP) and nuclear-shuttle (NSP) proteins. EuMV symptoms have been observed in Florida, however this study reports the first sequenced EuMV genome from this region. The FL EuMV genome is 99% identical to EMV isolates identified in Havana, Cuba.

Table 1: Sampling list of insects processed in this study, along with their corresponding location and the number of individuals processed for that sample type.

Sample (Common Name)	Sampling Location	Samples Processed
Superworms	Animal House Pet Store, St. Petersburg FL	6
Spiders	Girl Scout Camp Wai Lani, Tarpon Springs	3
Meal worms	Animal House Pet Store, St. Petersburg FL	4
Waxworms	Animal House Pet Store, St. Petersburg FL	4
Crickets	Girl Scout Camp Wai Lani, Tarpon Springs FL	4
Palmetto bugs	Girl Scout Camp Wai Lani, Tarpon Springs FL	4
Wasps	Girl Scout Camp Wai Lani, Tarpon Springs FL	2
Walking sticks	Girl Scout Camp Wai Lani, Tarpon Springs FL	3
Black beetles	Girl Scout Camp Wai Lani, Tarpon Springs FL	2
Large black beetle	Girl Scout Camp Wai Lani, Tarpon Springs FL	1
Centipede	Girl Scout Camp Wai Lani, Tarpon Springs FL	1
Worm (unidentified)	Girl Scout Camp Wai Lani, Tarpon Springs FL	1
Stink bug	Girl Scout Camp Wai Lani, Tarpon Springs FL	1
Spiders	Hillsborough County FL	6
Butterfly	Hillsborough County FL	1
Sea spider	Eckerd College Pier, St. Petersburg FL	1
German cockroach	Carolina Biological Supply, NC	6
Crickets	Carolina Biological Supply, NC	4
Lady bugs	Carolina Biological Supply, NC	12
Bean beetles	Carolina Biological Supply, NC	60
Horsefly pupae	Carolina Biological Supply, NC	60

Scarcophaga	Carolina Biological Supply, NC	8
Cricket nymphs	Carolina Biological Supply, NC	4
Jewel wasps	Carolina Biological Supply, NC	4
Earthworms	Carolina Biological Supply, NC	2
Red worms	Carolina Biological Supply, NC	2
Waxworms	Petsmart, St. Petersburg FL	15
Love Bugs	St. Petersburg FL	284

Table 2: BLASTn results for characterized EuMV genome sampled from Florida (*Euphorbia heterophylla*)

DNA-A[1] component				DNA-B[2] component			
Cuba	Puerto Rico	Jamaica	Mexico	Cuba	Puerto Rico	Jamaica	Mexico
100%	**99%**	**98-99%**	**93-96%**	**96-97%**	**-**	**96%**	**86-93%**

[1] The EuMV DNA-A component collected in Florida is nearly identical to EuMV sequenced from Puerto Rico, Jamaica, and Mexico and completely identical to sequences from Cuba.

[2] The DNA-B component shows slightly lower similarities to sequences characterized in Cuba and Jamaica, and larger variation compared to samples identified in Mexico.