Summer 2016

Discover And Analysis Of Grapevine Vein-Clearing Virus In Ampelopsis Cordata

Sylvia Marie Petersen

DISCOVERY AND ANALYSIS OF *GRAPEVINE VEIN-CLEARING VIRUS* IN *AMPELOPSIS CORDATA*

A Masters Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Plant Science

By

Sylvia M Petersen

July 2016

DISCOVERY AND ANALYSIS OF *GRAPEVINE VEIN-CLEARING VIRUS* IN *AMPELOPSIS CORDATA*

Agriculture

Missouri State University, July 2016

Master of Science

Sylvia M Petersen

ABSTRACT

A recent threat to the sustainability of grape production is *Grapevine vein-clearing virus* (GVCV), the first DNA virus discovered in grapevines. Infection with GVCV leads to vine decline, lower quality berries, and eventual death of the grapevine. Since GVCV was discovered in cultivated grapevines, research has been dedicated to investigating its range and origin. The entire genome of the first GVCV isolate from a grape cultivar 'Chardonel' has been deposited in GenBank and is used as a reference genome. More recently, two GVCV isolates were found in native *Vitis rupestris* in Missouri. In this thesis project, I applied polymerase chain reaction (PCR) assays to screen for GVCV in native *Ampelopsis cordata*, which is also in the Vitaceae family. I found GVCV in two accessions of this wild plant species. The entire genomes of the two GVCV isolates, GVCV-AMP1 and GVCV-AMP2, from *A. cordata* were sequenced. The GVCV-AMP1 genome is composed of double-stranded DNA, 7,749 bp long, while GVCV-AMP2 is 7,765 bp long. Genomic analysis indicated that they are new isolates with signature 9-base pair inserts in open reading frame II. A survey of GVCV in seventeen *A. cordata* plants around the Springfield area found that five were infected with GVCV, suggesting high incidence of GVCV among these native plants. These results demonstrated that GVCV spreads among species across genera in native habitats, and yielded crucial clues on origin and epidemics of GVCV. These findings will aid in developing new strategies for the management of GVCV-associated disease.

KEYWORDS: grape, virus, wild plants, GVCV, *Ampelopsis*, viral genome, *Vitis*

This abstract is approved as to form and content

_____

Wenping Qiu, PhD
Chairperson, Advisory Committee
Missouri State University

DISCOVERY AND ANALYSIS OF *GRAPEVINE VEIN-CLEARING VIRUS* IN *AMPELOPSIS CORDATA*

By

Sylvia M Petersen

A Masters Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Plant Science

July 2016

Approved:

_____
Wenping Qiu, PhD

_____
Melissa Remley, PhD

_____
Anson Elliott, PhD

_____
Julie Masterson, PhD: Dean, Graduate College

ACKNOWLEDGEMENTS

I would like to thank my supportive and patient family, Jon, Lillian, Parker, and Mazie Petersen. You all believed in me so much that you even washed dishes from time to time. Thank you mom and dad, Dixie and Elmer Gregg, for teaching me critical thinking, love, and self-sacrifice. Thank you my siblings, Micah and Kamon, for always accepting me just as I am. I am grateful to Pat, Sheli, and Carley Petersen for leading by example, for showing me how strong women are, and for introducing me to academia.

I appreciate every teacher I have ever had. They all had a gift to give and I am grateful. Thank you Dr. Wenping Qiu for your passion, curiosity, and inspiring nature. Thank you Dr. Elliott for helping me find my place at Missouri State University and especially in the warm environment you have created at the William H. Darr School of Agriculture. Thank you Dr. Melissa Remley for teaching me how to read journal articles all those years ago, and for superb editorial suggestions for this thesis. Thank you all at the Graduate College for your untiring assistance for graduate students. Lastly, I am grateful to Ru Dai, Steven Beach, and Michael Kovens for sharing their lab knowledge and experience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

"Help!" cried the farmer. "My crop plants have symptoms of disease, and they are dying!"

That sentence may sound like it is straight out of a children's story, but for anyone involved in agriculture this is the stuff of nightmares. Their crop is their livelihood, their passion, the representation of their collaboration with and struggle with nature, the product of all of their hard work, and a gift of sustenance and pleasure to humankind from them. This cry for aid from any farmer in the world has almost always triggered a response from scientists dedicated to the field of agriculture. The cumulative effect of each answer made to a farmer's call for assistance has benefited agriculture in these ways; improved pest and disease management, enhanced efficiency of land and water use, increased plant hardiness, provided greater crop quantity and quality, and ultimately brought outstanding productivity to the agricultural system. Despite all of the advancements in agriculture, there are always more improvements to make in order to feed, clothe, and house an ever-growing human population.

According to Anderson et al. (1), the largest portion of emerging disease in cultivated plants is caused by viruses, and grapevines are no exception. It is not surprising then, that a "Help! My grapevines have symptoms of disease, and some are dying!" call from Missouri grape growers in 2004 was the impetus behind the discovery of the first DNA virus infecting grapevines, *Grapevine vein-clearing virus* (GVCV) a badnavirus in the family *Caulimoviridae*.

The symptoms that had alarmed vineyard owners were dramatic. The grapevine leaves were chlorotic and the major veins were translucent, the leaf edges were curling, the plant overall looked stunted in comparison with unaffected vines; and maybe most importantly, the berries were deformed and of low quality. Plant scientists from Missouri State University responded to the grape growers' call for help. After testing for all of the usual suspects that might cause such symptoms, they realized they were quite possibly dealing with a previously undescribed virus. Using modern biotechnological techniques, such as deep sequencing of small interfering (si)RNAs, researchers indeed found a new virus in grapevines. In 2009, the first DNA virus ever discovered to infect grapevines was documented. By 2011, the entire genome of the new virus was sequenced, deposited in GenBank (NC_015784.2), and named *Grapevine vein-clearing virus* in acknowledgment of the most notable symptom (2). This is the reference sequence for GVCV and is referred to as "GVCV-CHA" because it was isolated from a Chardonel grapevine.

Since that time GVCV infected grapevines have also been reported in Arkansas, Illinois, and Indiana. The virus has been detected in grape cultivars of *Vitis vinifera* and its hybrids including, Cabernet sauvignon, Chardonnay, Valvin muscat, Cabernet franc, Riesling, Chardonel, and Vidal blanc (3). The symptoms of infection are visually noticeable. The signature translucent veins with chlorotic streaks or mottling along the veins are easily seen when scouting a vineyard. Other symptoms include shortened, zigzag shaped internodes, which contribute to the overall dwarfed nature of the vine, reduced berry set, deformed berries with a stony texture, and curling or warped leaves.

Infected grapevines succumb to the stress of the virus over a period of years and eventually die.

When a virus threatens human populations, the most commonly asked questions are: Where did the virus originate? What is the range of the virus? How is the virus spread? These are the same questions grape growers had about GVCV. To help elucidate these questions, investigations were initiated into the status of Missouri's native *Vitis* species concerning GVCV. *Vitis rupestris*, or 'sand grape' as it is locally known, was found to be infected with GVCV. This was the first wild grapevine in which GVCV was detected. The GVCV isolate from an infected *V. rupestris* from Taney County, Missouri was chosen for sequencing. This was the second GVCV genome sequenced and the results were deposited in GenBank (KJ725346.1) in 2014. This GVCV isolate is referred to as "GVCV-VRU-1" for *V. rupestris*. Interestingly, a novel symptom of necrotic flecks occurs along with the vein clearing during an infection of GVCV-VRU-1. Something else was significant about GVCV-VRU-1; it contained a 9 base pair (bp) insertion in its genome. By 2016 the second isolate found in *V. rupestris* had been sequenced and submitted to GenBank (KT907478.1), and is referred to as "GVCV-VRU-2". Though this isolate was also found in a *V. rupestris*, the genome did not contain the 9 bp insert that occurred in VRU-1. Despite all the new information gathered, the three questions still remained. The incidence of GVCV in native *Vitis* species was low, other areas of the country with native *Vitis* had not detected GVCV in their vineyards, and the vector had not been identified.

Ampelopsis cordata Michx. is a perennial, woody vine that is native to the Midwestern and southeastern United States. The leaves of *A. cordata* are heart-shaped,

alternate, and serrated, and can be mistaken for grape leaves. The flowers, unlike grapes, are in cymes rather than panicles, and the colorful fruit is edible only for birds. *A. cordata* is in the Vitaceae family, just as cultivated and native *Vitis* species. When mild GVCV-associated symptoms were noticed on *A. cordata* near Linn Creek, Missouri, samples were collected for testing. The hypothesis was that these symptoms were GVCV related. To determine if this were the case, total DNA would be extracted from the plant and tested in the laboratory. If indeed a second genus of native plants were infected with GVCV the objectives and hypothesis would broaden. This native plant could be a reservoir population for GVCV, perhaps even the source. If *A. cordata* populations were the origin of GVCV, this would help delineate the range of the virus, and would hopefully help narrow the search for the vector of transmission. If GVCV were detected in this plant, the viral DNA would need to be purified and sequenced so that it might be compared to the previously described isolates. Phylogenomic analysis could then be conducted to get a clearer picture of the evolution of GVCV and help to clarify whether native plants or cultivated grapes first hosted the virus. The implications of discovering this devastating virus in another native plant population would be critical for grape growers. This information could affect pest and disease management and vineyard location planning. Providing support for farmers to grow healthier crops is the ultimate goal of this work. To this end, we conducted this project and investigation of GVCV status in wild *A. cordata*.

LITERATURE REVIEW

Viruses in Wild Plants

Viruses are ubiquitous in the plant kingdom; from algae to flowering plants, from moss to trees, all can play host to plant viruses (4-6). Knowledge and research concerning viruses in wild plant populations is limited. When research has been conducted on the incidence of viruses in wild plant populations, the results are revealing. When surveys are conducted, viruses are found to be commonplace in natural plant populations. Yahara and Oyama (7) found as many as 69% of the native *Eupatorium chinense* were infected by a geminivirus, *Tobacco leaf curl virus* (TLCV). A study of *Tobacco mild green mosaic virus* (TMGMV) infecting the wild plant *Nicotiana glauca* in Spain, found as high as a 77% infection rate with TMGMV in plants surveyed (8). When wild cabbage populations were sampled over a 25 km stretch of countryside in Dorset, England, *Cauliflower mosaic virus* (CaMV) was found in 60% of the natural cabbage population (9).

Though infection rates vary, the consensus is that viruses maintain a strong presence in wild plant populations (10). However, it is agreed that there is under-abundance of research into viruses affecting wild plant populations (11). The reasons are many and may be best understood by first examining why such studies *would* be conducted, and most of those reasons are obviously concerning agricultural practices. A driving force for exploring wild plants and viruses would be to discover an alternative host, a source, or a reservoir for a virus that is having an adverse impact on crop production (12). Another motivation would be drastic symptoms being expressed in wild

plants catching the eye of an observant scientist; though, as will be discussed shortly, drastic symptoms are rare in wild plants (13). Recent research into wild plants and their viruses has been prompted by the goal of finding genetic sources for virus resistance or tolerance that may be incorporated into crop plants (14, 15). Another plausible incentive for inquiry into this matter would be phytosanitary measures as a prophylactic for one region from the introduction of a virus infecting native species of another region (16). From the perspective of why scientists would take the time and effort to study viruses in wild plants, the reasons are almost entirely limited to solving problems directly affecting agriculture, yet there are a multitude of viruses in wild populations researchers have never considered, simply because they are as yet a non-issue for plant commerce (1).

As mentioned in the previous paragraph, one factor in the lack of ardent investigation of viruses in wild plants is the typical absence of apparent symptoms during viral infection (10). In large part, this is attributed to high tolerance, or even resistance, to viruses in the natural plant population. In this regard, wild plants have definite advantages over domesticated plants. As plants and viruses have co-evolved over millions of years their relationship has affected both at the genotypic level (17). This is evidenced in part by the resistance genes (R genes) of plants and the effector proteins of viruses (18). Native populations have high genetic diversity supplying them with greater chances of tolerance or resistance to viral infections, thus reducing obvious viral symptoms (19). Not only would genetic diversity serve these plant populations well in resistance to viruses, but also in resistance to the insects or nematodes that transmit viruses.

Examples of viral infections in wild plants species with unapparent symptoms include reported cases dating back to the 1960's (20) to present studies (21). A comprehensive review of this subject was presented in the American Journal of Botany (22), that listed occurrences such as wild plantain, *Plantago*, with a 64% virus infection rate of randomly sampled plants in Great Britain, with the majority of the infected plants having no visible symptoms (23). Another example of asymptomatic wild plant populations that are infected by viruses is found in wild grasses that were infected with *Barley yellow dwarf viruses* (BYDVs). Small grain crops are susceptible to BYDV infection and the consequences are severe if infected. Symptoms range from reproductive harm, such as reduced seed set, to stunting, even to early senescence or premature death of the grain crop. In a 2002 study of BYVDs in three weedy grass species growing near crop field margins in upstate New York, Remold found that it is difficult, if not impossible, to visually diagnose BYDV infection in the wild grasses used in this study, though a majority were infected (13). *Peanut stunt virus* (PSV), as the name implies, causes severe dwarfing of peanut plants and up to 50% yield loss; however, no clear symptoms were observed on PSV infected wild clover. Furthermore, in a random sampling of white clover in the southeastern United States, PSV was detected in 21% of the plants assayed (24).

Besides the aid of genetic diversity in combatting viral diseases, wild plants have other advantages in diminishing the impact, and thus the symptoms, of plant viral diseases. The non-uniform nature of the growing environment of wild plants lends protections such as buffers of non-host plant species between susceptible species, environmental and climactic variations that may not favor transmission of the virus,

varying plant population densities that could discourage viral transfer, natural predators and parasites of insect, nematode, or wind borne vectors of the virus, varied ages of plants in the wild population influencing susceptibility, varied proximity to viral sources, and even antagonism between indigenous viral populations. Because of all the above factors, a virus that has minimal impact on wild populations can prove to be devastating if and when it spills over to plant monocultures cultivated for human use (25). Conditions in nature mitigate the severity of viral invasions in wild plants, while designed breeding and agricultural practices exacerbate viral invasion in agriculturally significant plants (21).

Wild plant populations can serve as a reservoir for viruses that have detrimental effects on agroecosystems. Viruses that move from wild to cultivated plants (or conversely, from cultivated plants to wild plants) are referred to as emerging viruses; in other words, they are newly discovered and are likely increasing in incidence in plants that are valuable for food, feed, or fiber (or in the case of wild plants, have ecological importance) (26, 27). The definition of an emerging virus may also include a viral population that has evolved in such a way that it is able to infect new hosts or vectors and spread where it has not traditionally been epidemic (28). *Pepino mosaic virus* (PepMV) causes leaf chlorosis and mottling on tomato plants, along with economically devastating symptoms of fruit discoloration and reduced fruit size (29). Recent research in Peru has supported, on a molecular level, the hypothesis that PepMV emerged from wild *Solanum* species plant hosts to infect cultivated tomatoes (30).

At times, when a plant virus emerged from a wild plant population into crop species, one of the management practice employed is destruction of proximal wild hosts

of the detrimental virus, thus reducing incidence of viral infection spreading from the wild plants to the cultivated crop. In Melbourne, Australia, *Lettuce necrotic yellow virus* (LNYV) was harming lettuce production. It was discovered that wild sow thistle was a source of the virus. A study was conducted to assess the effectiveness of destroying wild sow thistle in a 470-yard by 170-yard area surrounding a lettuce field using manual and chemical methods. Using an alternative lettuce field, which had not been cleared of bordering sow thistle for comparison, researchers found that this management practice could account for a dramatic reduction in rate of infection of cultivated lettuce. The incidence of LNYV in the untreated field was 75.6% of lettuce plants lost to the virus, whereas in the treated field loss was 6% (31).

Removing sources of inoculum in wild plant populations is not always practical because of logistics, differing opinions on land use, cultural or religious ideals, and environmental or ecological implications. There is also the confounding factor of the number of species, genera, and even families that may be present in wild plant populations that can host an agriculturally significant virus. For example, Hawaiian grown tomatoes, lettuce, and bell peppers are adversely impacted by *Tomato spotted wilt virus* (TSWV), causing up to 90% loss of crop production (32). Years of research concerning TSWV has demonstrated that 16 plant families and 44 plant species are natural hosts of this virus (33). Five out of six grass subfamilies have been shown to host BYVDs in 150 species, with the majority of these being wild grasses (34).

Missouri is home to at least five species of wild grapes: *Vitis aestivalis, Vitis cinerea, Vitis riparia, Vitis rupestris, and Vitis vulpina,* and North America has as many as 15 indigenous *Vitis* species (35). According to the USDA plant database there are

around 50 native plants in the Vitaceae family that are native to the United States, with about 21 of those species inhabiting niches in Missouri (36). Viruses have been detected in native grapevines in the United States (37) and in other countries (38, 39). When the nine most crucial grapevine viruses were surveyed in wild grapevines in Sicily, the results showed five populations of wild grapevines were infected with *Grapevine rupestris stem pitting-associated virus* (GRSPaV) and one population with *Grapevine fanleaf virus* (GFLV) (38). In Missouri, GVCV has been detected in native *V. rupestris* plants (Beach et al., unpublished). Clear statements concerning whether wild or cultivated grapevines were the source of spillover of the viruses were not forthcoming in any of these studies.

Viruses in Cultivated Grapevines

Roughly 20 million acres of arable land is devoted to vineyards worldwide (40). Grapes are an extremely valuable agricultural product, and have been cultivated since human history has been recorded (41). In the United States grapes are the highest value fruit crop, and in 2014 over 1 million acres was used for vineyards and $162 billion in revenue was generated from the grape and wine industry (42). Missouri is one of the top ten grape producing states with nearly 2,000 acres devoted to vineyards harvesting over 6,000 tons of fruit in 2013. Revenue from vineyards and related industry totaled $1.76 billion in 2013 for the state of Missouri (43).

Unfortunately, this economically important plant is also host to more viral pathogens than any other woody crop (44). It is difficult to quantify the toll that the aggregation of viral diseases exerts on grapes as an agricultural product worldwide. The

five main RNA viruses affecting grape production are*: Grapevine fanleaf virus, Arabis mosaic virus, Grapevine leafroll-associated virus 1, Grapevine leafroll-associated virus 3* and *Grapevine fleck virus* (45). *Grapevine-leaf roll associated virus* (GLRaV) is estimated to be the most prevalent and damaging grapevine virus in the world (46), and is predicted to cost California at least $60 million annually (47). Interestingly, the virus rarely kills the plant. It is economically detrimental because viral symptoms include reduced photosynthesis and vine productivity, and altered sugar and organic acid composition which affects berry quality and ultimately the end products of the vineyards, especially wines (48).

Grapevines are shown to host more than 65 viral pathogens (44). This is partly because grapevines are perennials, and as such have many opportunities for exposure to pathogens over their lifespan (49). Another cause, and arguably the most culpable, for the high number of viruses infecting grapevines are vegetative propagation methods (50, 51). Most vineyards are established using cuttings, rooted cuttings, grafted plants, or a combination of all three. In all of these cases, viruses can be moved into the new vineyard via the planting material, or introduced to an established vineyard with the addition of new plants (52, 53). As discussed previously, native plants also play a role in viral infections in cultivated grapevines either as a source of spillover or as a reservoir for viruses (54).

There are no cures for viral disease in grapevines established in the field (55-57). Recommendations for control of viral disease include controlling the vector (58), planting only certified virus free vines (47), rouging infected vines to prevent vectors from

spreading the virus to non-infected plants (59), and surveying nearby wild plant populations as potential sources (37, 60).

Most of the 65 known viruses infecting grapevines are RNA viruses. To date there are only three DNA viruses that have been described. The first DNA virus discovered in grapevines, GVCV, was found in Missouri, and is a double-stranded (ds) DNA virus in the *Caulimoviridae* family (2). A second DNA virus of grapevines was simultaneously explored circa 2010 in California and New York, *Grapevine red blotch-associated virus* (GRBaV), a single-stranded (ss) DNA virus in the family *Geminiviridae* (61, 62). The third DNA virus found to infect grapevines was discovered in Greece, *Grapevine Roditis leaf discoloration-associated virus* (GRLDaV), another dsDNA virus in the *Caulimoviridae* family (63), with a strain of this virus found in Italy in 2016 (60).

Caulimoviridae Badnavirus* and *Grapevine vein-clearing virus

The family of viruses known as *Caulimoviridae* are the only dsDNA pararetroviruses in the plant kingdom (64). R.J. Shepherd made the discovery that *Cauliflower mosaic virus* (CaMV), in the genus *Caulimovirus* was composed of dsDNA (65). Since that time, dsDNA viruses in the *Caulimoviridae* family have been divided into two groups and classified into eight genera (66). The shape of the virion determines which group they are placed in; either rod shaped in the bacilliform group, or roughly spherical in the icosahedral group (67). Only two of the eight genera are in the bacilliform group, *Badnavirus* and *Tungrovirus*. Whereas there is only one species of *Tungrovirus*, there are at least 32 species of *Badnavirus*, and GVCV is one of these (2).

Replication and genome structure. All *Caulimoviridae* replicate through an RNA intermediate and have a reverse transcription step; but they do not require genome integration as do true retroviruses (68). Badnavirus genomes are dsDNAs of 7.2 to 9.2 kb long with single-strand overlaps, or discontinuities, this is referred to as open circular form (66). The overlaps mark the beginning of plus and minus strand reverse transcription at the $tRNA^{met}$-binding site that is the initiation for reverse transcription on the minus strand. The first step in the replication process is when the genomic DNA is imported into the nucleus and the overlaps are covalently closed. The host DNA-dependent RNA polymerase II creates terminally redundant, longer than genome size mRNA which is transported to the cytoplasm and translated into viral proteins or used as a template for reverse transcription to the open circular dsDNA, starting the cycle over (69). The primer for reverse transcription is $tRNA^{met}$, and the virally encoded reverse transcriptase is responsible for polymerization of the first strand (-) DNA, using the virally encoded RNaseH to degrade the RNA template. Small fragments of the degraded RNA are used as primers for synthesis of the second strand (+) DNA, and upon completion the complementary strands form the characteristic relaxed dsDNA molecule of a badnavirus (70).

The badnavirus genome is around 7.5 kb and encodes at least three open reading frames (ORFs), but can have up to five ORFs. GVCV-CHA has 7,753 bp and three ORFs. Current research has identified two ORFs on the antisense strand of at least one badnavirus, but has not identified expression or function of these ORFs (66). The first ORF on the sense strand of badnaviruses ranges from 399 to 927 bp, the second ORF is the shortest at 312 to 561 bp, and the third ORF is the largest at 5100 to 6000 bp. The

reference sequence for GVCV has an ORF I of 627 bp, ORF II of 384 bp, and the third ORF is 5,826 bp. A leader sequence of 686 nucleotides is present in GVCV and four short ORFs are contained in the leader region (71). The first and second ORFs encode products whose function is still being deciphered. The ORF I protein is thought to be associated with both virions and plant components (72), while the ORF II likely encodes a nucleic acid-binding protein (73). The third ORF encodes the reverse transcriptase (RT), RNaseH, capsid protein, aspartate protease, and perhaps a movement protein (74). There is one strong promoter for the GVCV genome, with the core promoter of 341 nucleotides (nt) long and the transcription initiation site at nucleotide 7571, which is 27 nts downstream from the TATA1 box (71).

Classification and diversity. Badnaviruses are classified by the percentage identity of the RT-RNaseH sequence in ORF III. If there is ≥ 20% divergence of nucleotide sequence in this region, then the badnavirus is considered a novel species according to the International Committee on the Taxonomy of Viruses (ICTV) (75).

Badnaviruses are highly heterogeneous, likely due to the lack of proofreading capacity of the reverse transcriptase and introduced inaccurate nucleotides during replication by reverse transcription (80). Propagation methods also lend to the diversity of badnavirus populations, as frequent exchange of plant material through grafting and cuttings is common in perennial plants, thus possibly introducing several isolates of a virus into one plant or plants in a growing region (26). The genetic diversity of badnaviruses has been documented in banana (81), yams (82), sugarcane (83), and cacao (84), among other crop plants. This phenomenon of genetically diverse viral populations

holds true for GVCV as well, with sequence variations defying phylogenetic grouping by location or grape variety (3).

Transmission and symptomatology. There are three main ways that badnaviruses are transmitted from plant to plant. Vegetative propagation is the primary means for badnaviruses to spread among the propagated plant material, particularly in perennial woody plants such as grapevines (76). Insect transmission, particularly via mealybugs and aphids, is the second most likely means of conveyance of any given badnavirus (77). Finally, a very few badnaviruses are known to be vertically transmitted through the seeds or pollen of infected plants (78), and fewer yet may be mechanically transmitted (79). For many badnaviruses however, the vector remains unknown (66), which is currently the case for GVCV.

Symptoms attributed to badnaviruses include chlorotic mottling and streaking, necrotic streaks and spots, vein clearing, deformed leaves, reduced internode length, swollen shoots, overall stunting of the plant, and death of the infected plant (69) The spectrum of symptoms caused by badnaviruses is diverse for many reasons. Even though most badnavirus species have a limited host range, as a genus they do infect a large variety of plant families, thus symptoms vary according to the type of host plant (66). The host also has an influence on symptoms according to the level of resistance, tolerance or susceptibility that is genetically inherent in the plant (85). Alternative causes of symptom variation are viral titer, or concentration in the host plant, along with the pathogenicity of the specific badnavirus species (86, 87). Studies have shown that environmental factors, such as temperature and humidity, also influence symptoms caused by badnavirus infections. Biotic and abiotic stresses may lend to the type and

severity of symptoms, along with the age of the infected plant (88). Finally, symptoms have been shown to develop, disappear, and appear again over the life of the host, with new growth of the plant, or with the seasons (89).

MATERIALS AND METHODS

Synopsis of Research Methods

The focal points of this study included determining whether GVCV is infecting wild plants outside of the *Vitis* genus in order to gain insight into the origin of the virus, aiding in understanding the range of the virus in native, non-domesticated plants, and conducting genomic comparisons between all known GVCV isolates, including the two obtained from *A. cordata*. To achieve these research goals the work entailed extracting DNA from the *A. cordata* plants (referred to as AMP1P and AMP2P, while the viral isolates obtained are referred to as AMP1 and AMP2), using the widely available technique of polymerase chain reaction (PCR) to amplify specific fragments of the DNA, isolating and purifying the fragments in an agarose gel, cloning the fragments into competent bacterial cells allowing the *E. coli* to make many thousands of copies of the viral DNA, extracting the plasmid DNA from the bacteria and then capturing the nucleotide order by Sanger sequencing. The primer-walking technique was then used to design primers to obtain any unknown portions of the viral sequences that could not be gained using previously designed GVCV specific primers. Once the entire genomes were known, comparative analysis between the two newly discovered GVCV isolates in *A. cordata* and the three previously determined isolates, CHA, VRU1, and VRU2 was conducted using various bioinformatics software.

Collection and Preservation of Samples

The plant sample AMP1P was collected from Linn Creek, Missouri (Figure 1) on June 11, 2015 along with a *V. rupestris* sample (which tested negative for GVCV). While on a scouting trip to investigate viral symptoms in an Augusta, Missouri vineyard, a group from Missouri State University's plant biotechnology laboratory stopped along creek beds to search for samples of native *Vitis* plants for lab testing to determine the status of GVCV in this part of Missouri. The *A. cordata* was collected for the following reasons; mild visible symptoms associated with GVCV were noticed on some *A. cordata* plants, they are in the Vitaceae family as are cultivated and wild grapes, and *A. cordata* shares habitat with native grapes (90).

The plant sample AMP2P was collected from Close Memorial Park in Springfield, Missouri (Figure 1) in the summer of 2014. This sample was collected by Dr. Wenping Qiu, Missouri State University, for similar reasons as AMP1P; the mild, GVCV-like symptoms present on the vine leaves. The plants were placed in a plastic bag with a moist paper towel, labeled, placed in a cooler with ice, and transported back to the lab. Both plants were vegetatively propagated and clones of each are in the greenhouse at the MSU Fruit Experiment Station in Mountain Grove, Missouri. Leaf samples of each were kept in the original plastic bag and placed in the 4°C freezer for no more than three days, at which time three 100 mg portions of leaf tissue from each plant (AMP1P and AMP2P) were weighed, wrapped in foil, labeled, and frozen in liquid nitrogen. These were stored in a -80°C freezer.

Extraction and Visualization of DNA

Leaf samples were removed from the freezer and immediately placed in liquid nitrogen. Using a mortar and pestle, the samples were ground to a fine powder constantly exposed to liquid nitrogen. Using Qiagen DNeasy© Plant Mini Kit, DNA was extracted from the powdered leaf samples following Qiagen protocol and eluted in 15 µL of autoclaved distilled water. DNA was quantified using the Thermo Scientific NanoDrop 2000 spectrophotometer and the quality was assessed by sample absorbance at 260/280 nm. The minimum acceptable concentration was 10 ng/µL, and all samples were diluted to this concentration prior to polymerase chain reaction (PCR).

The initial assay to detect the presence of GVCV in the Ampelopsis DNA samples was conducted using the primer set designed specifically for the GVCV genome to generate an 835 bp amplicon. The primers are defined as follows: 1101 F (5'-CTGAAAGGTAGATCTCCACG-3') and 1935R (5'-TCGGTGTAGCACTTCTATTCT-3'). To ensure the presence of quality DNA from the plant samples 16S ribosomal RNA gene primers were included with an expected amplicon size of 105 bp. The 16S primers used are as follows: forward primer (5'-TGCTTAACACATGCAAGTCGGA-3') and reverse primer (5'-AGCCGTTTCCAGCTGTTGTTC-3'). The PCR master mix consisted of all necessary reagents shown in Table 1 in the noted concentrations, excluding the DNA, which is independently added to each PCR tube. Reactions were placed in a Verti 96 Well thermocycler with the program parameters shown in Table 2.

After completion of the above settings in the thermocycler 10 µL of each sample were loaded into a GelRed stained 1% agarose gel and subjected to electrophoresis at 100 volts in a 1X Tris-Borate EDTA buffer for 30 minutes. The gel was viewed under UV

light to assess bands and an image of the gel was saved. Because the bands for 16S and the 835 bp amplicon specific to GVCV were detected in this initial assay, the procedure to capture and sequence the entire genome of each putative GVCV isolate in both AMP1P and AMP2P samples was implemented.

Generation of Overlapping Fragments by PCR

The known sequence of the reference genome, GVCV-CHA, was used to design primers for acquiring the unknown sequences of the GVCV isolates in AMP1P and AMP2P. Dividing the reference sequence of 7,753 bp into three overlapping fragments of roughly 3000 bp each, three sets of primers were designed and used in PCR to amplify each of these fragments. Specific primers and fragment sizes are shown in Figure 2. The primers shown in Figure 2 are named according to their numerical position on the GVCV-CHA reference genome which begins with 1 at the transfer RNA (tRNA) binding site.

Using the primers shown in Figure 2, three large fragments of the previously detected viral genomes were amplified with PCR. Because of the length of the fragments and the downstream applications of cloning and sequencing the DNA amplified, Platinum® *Taq* DNA Polymerase, High Fidelity was used. This polymerase possesses 3' to 5' proofreading capability to ensure a more accurate transcription of longer stretches of nucleotides. Master mix for the High-Fidelity polymerase reaction is recorded in Table 3 and thermocycler settings are shown in Table 4. The forward and reverse primers used were different for each of the three fragments so they are simply represented in Table 3 as F and R respectively but a detailed definition of each can be found in Table 5 using their

numerical names as shown in Figure 2. Thermocycler settings (Table 4) were also varied from the initial assay because of the increased length of the amplicons and to accommodate the optimal temperature for the High-Fidelity polymerase.

Purification and Isolation of GVCV Fragments

Once the large fragments had been amplified using PCR, 15 µL of each reaction were mixed with 5 µL of loading dye and loaded into a 1% agarose gel that had been stained with GelRed nucleic acid stain. Electrophoresis was carried out at 100 volts for 40 minutes in a 1X Tris-Borate EDTA buffer. Electrophoresis was used, not only to verify the correct fragments had been amplified, but also to isolate the desired fragments of viral DNA (91). The large fragments migrating to the correct location on the gel in comparison with the DNA ladder ensures that only the DNA of the correct size is present. The gel was placed under UV light so that the DNA was visible and then the fragments to be purified were cut from the gel and placed in tubes for separation from the agarose. For this purpose the Qiagen MinElute© Gel Extraction kit was used in accordance with Qiagen protocol. The pure DNA was eluted in 15 µl of autoclaved distilled water and then concentration was measured using the NanoDrop spectrophotometer.

Cloning the Three Overlapping Fragments

Recombinant, or chimeric DNA was pioneered by scientists in the 1960s and made a mature appearance in lab processes with the work of Cohen and Chang in the 1970s (92). The technique allowed combination of the purified viral DNA of the three overlapping fragments with a predesigned vector, which together formed a recombinant

plasmid. The 2817 bp predesigned vector included three important features for this research; a spectinomycin resistance gene, GW1 and GW2 priming sites that flanked the inserted viral DNA, and a TOPO® cloning site. Following the protocol of Invitrogen's pCR®8/GW/TOPO® TA Cloning Kit, the viral DNA was inserted into the vector with ease due to the complementarity of *Taq* polymerase-amplified PCR product (ends with nucleotides AA) and the TOPO® cloning site (with overhanging TT nucleotides). The recombinant DNA was then transformed into competent One Shot® TOP 10 *E. coli* bacteria by the heat shock method.

The transformed *E. coli* were plated on sterile Luria-Bertani (LB) agar plates that contained 100 µg/m of spectinomycin and encouraged to form colonies by being placed in a 37°C incubator. Two plates were inoculated for each fragment, one with 90 µl of the transformation and one with 40 µl. The inclusion of spectinomycin in the growing medium for the transformed bacteria was complementary to the spectinomycin resistance gene engineered into the pCR®8/GW/TOPO® vector, as this allowed only bacteria containing the recombinant plasmid to grow on the plates. Because plasmids replicate autonomously as bacteria replicate, many thousands of copies of the recombinant plasmid were made over the 18 hours of incubation.

After incubation, three clearly defined colonies were chosen that represented each plasmid created from one viral DNA fragment plus the predesigned vector. Colony PCR and visualization on agarose gel was performed to verify the presence of the inserted DNA fragments in the recombinant plasmids. This was achieved using the initial primers that had been employed in amplifying the viral DNA shown in Figure 2. Each of these confirmed colonies were placed in a 15 ml tube with 5 ml LB broth and 100 µg/ml of

spectinomycin, again to ensure the growth of only transformed bacteria. The tubes were placed in an incubator at 37°C with shaking at 260 rpm for a period of 14 hours. As the bacteria replicated, so did the inserted plasmid, manufacturing thousands of clones of the recombinant DNA.

Following the final incubation period, plasmid DNA was purified from 3 ml of the 5 ml solution. The rationale in not using the entire 5 ml was to save verified plasmid-containing bacteria as an inoculant for subsequent replication if necessary and to reserve for long term storage the bacteria containing the DNA fragments of interest. For purification of molecular biology grade DNA from the plasmids, Qiagen's QIAprep® Spin Miniprep Kit was used according to the protocol, except 50 µl of autoclaved distilled water was the elution medium as opposed to 50 µl of Buffer EB.

Sequencing the Genomes

Once the bacteria contained the fragments of interest and the plasmid DNA had been extracted from the bacteria, the DNA samples were sent for sequencing. The samples were sent to Nevada Genomics Center, University of Nevada, Reno. Sanger sequencing (93) was the chosen method of DNA sequencing because of its high reliability. Protocol for sending samples to the Nevada Genomics center requires at least a 5 µl reaction in a labeled tube. The reaction components for plasmids greater than 5000 bp must include at least 250 ng of DNA along with 1 µl of 10 µM primer in each tube. Tubes were created for each fragment of interest with GW1 primer along with DNA, and tubes were created with each fragment of interest with GW2 primer included with DNA. As mentioned earlier, TOPO® vector contains GW1 and GW2 priming sites that flank

our inserted DNA allowing sequencing of only that portion of the plasmid. This was the first step in the sequencing process, as typically only around 500 quality nucleotides are read from each sequencing event, and the fragments of interest were approximately 3000 bp each.

After sequencing, Nevada Genomics Center reports the nucleotides read along with a chromatograph that includes a phred quality score for each nucleotide. Phred quality scores are defined as a property that is logarithmically related to the base-calling error probabilities, and was designed in large part for use in the Human Genome Project (94). In this study, only nucleotides with a phred score equal to or greater than 20 were accepted. A phred score of 20 indicates a probability of 99% that the base called is accurate. After examining the sequenced nucleotides along with their corresponding phred quality scores in the CodonCode Aligner software package, that information was used to discriminate between reliable nucleotides to begin building contiguous sequences, and unreliable nucleotide regions that should be re-subjected to sequencing. The GVCV-CHA sequence was used as a reference in aligning the newly obtained sequence fragments. This process was conducted for both AMP1 and AMP2 viral genomes.

Primers and the Primer Walking Technique

Because DNA polymerase cannot begin transcribing *de novo,* a short stretch of oligonucleotides were used to "prime" the transcription process. By designing a primer of complementary bases to the DNA desired, the primer will anneal to the targeted DNA and the polymerase will begin adding nucleotides to the primer. Fortunately, many primers had already been designed that were specific to the viral genome, GVCV-CHA,

and many of these available primers were suitable for sequencing the AMP1 and AMP2 isolates, as would be expected. However, there were areas in the AMP isolates that were dissimilar enough from the reference genome that the primers designed for the reference genome did not anneal to the viral DNA found in the *Ampelopsis* plants. Primers were designed for these stretches of DNA sequence that could not be accessed by the available primers. Once a stretch of DNA was sequenced and the results were viewed, it was possible to design a primer toward the end of that sequence because those nucleotides were now known. The new primer would then yield an additional 500 or so quality base reads. In this way, little by little, unknown sequences of bases were revealed until the entire genome was represented with bases having phred quality scores of 20 or greater.

A detailed list of primers used in sequencing AMP1 and AMP2 may be found in Table 5 and 6, along with a table (Table 7) showing the previously designed GVCV-specific primers that did not anneal to the viral DNA found in Ampelopsis. Excluding GW1 and GW2; in total 24 primers were effectively used in sequencing AMP1, and 28 primers were effectively used in sequencing AMP2. GVCV-CHA primers that were not effective in sequencing the AMP genomes numbered seven in total.

Sequence and Phylogenetic Analysis

Sequence analysis was conducted by examining all five GVCV genomes at the nucleotide level, and comparing open reading frames at both nucleotide and amino acid levels. Identity percentages at both nucleotide and amino acid level were calculated using the EMBOSS Needle program. The entire nucleotide sequence of each genome

was uploaded to ORF Finder (NCBI) to determine ORF lengths, and also to MOTIF (GenomeNet) to investigate protein motifs in the sequence.

Phylogenetics is a scientific method of grouping organisms based on evolutionary divergence. The first modern phylogenetic tree is credited to Darwin (95), but with the advent of computer driven bioinformatics phylogenetic trees have become much more detail oriented and complex. For comparison purposes an unrooted phylogenetic tree was constructed at the nucleotide level with the five genomes of GVCV isolates that have been sequenced. Each sequence was then converted to its amino acid equivalent by EMBOSS Transeq software. Each isolate's open reading frame II (ORF II) was compared at the amino acid level, and an unrooted tree was constructed using ClustalW.

Designing Definitive Primers

It has been proposed that criteria for determining a new isolate of GVCV be based on the most variable region of the genome, ORF II. Anything with a less than 90% identity at the nucleotide level to any other GVCV isolate's ORF II would be considered a novel isolate of the virus (Qiu, unpublished). To that end, all five known genomes were aligned in CodonCode Aligner to explore conserved areas common to all five that flank ORF II, so that universal primers could be designed to capture this area in potentially all viral isolates of GVCV. This would allow this area to be quickly sequenced in order to determine if it were necessary to sequence the entire genome, or if this were an isolate that had already been described.

Local Survey of Vitaceae Species

During the spring of 2016, just as native Vitaceae plants were flowering, a localized survey was conducted to determine the status of *Ampelopsis* and *Vitis* plants in regard to GVCV infection. Six areas in and around Springfield, MO (Figure 3 and Table 8) were chosen and were documented by GPS coordinates as noted in Table 8. A total of 17 *Ampelopsis* and 15 *Vitis* samples were subjected to the protocol detailed above under the 'Collection and Preservation of Samples' and 'Extraction and Visualization of DNA' headings, except a triplex of primers was used for PCR. The PCR assay used for the local survey included the 835 bp and 16S primers sets, and also a 442 bp primer set. The former primer sets have already been defined, the latter consists of the GVCV specific primers 4363F (5'-ATCTGCTCAATTTCTGAAGGAGAAG-3') and 4804R (5'-GGAATGCATTGTGCTCGTAG-3'). The components for the triplex master mix are listed in Table 9 and the thermocycler program for PCR on these samples is listed in Table 10. It should be noted that primer 4804R is a primer listed in Table 7 as one that does not anneal to the AMP1 or AMP2-GVCV isolates. This was intentional to include this primer in the assay, and the rationale will be discussed with the results.

After completion of PCR, 10 µl of each reaction were loaded into a 1% agarose gel that had been stained with GelRed and subjected to electrophoresis for 30 minutes in a 1X Tris-Borate EDTA buffer. The gel was viewed under UV light with the GelDoc-It Imaging System.

RESULTS

Symptoms on Two Accessions of *Ampelopsis cordata*

Symptoms associated with GVCV have been documented in both accessions of *A. cordata* that were the subjects of this thesis. The plant sample referred to as AMP1P was collected from Linn Creek, MO and sample AMP2P was taken from Close Memorial Park in Springfield, MO (Figure 1).

The initial interest in pursuing the research was based on symptoms seen in Ampelopsis plants in their native habitat; such as mild vein clearing, chlorotic spots, necrosis of leaf tissue, and slight curling of leaf margins, all of which may indicate the presence of GVCV (2). As shown in Figure 4, the vegetatively propagated clones of both plants used in this study exhibited mild to moderate GVCV associated symptoms. These clones remain in the greenhouse at the Missouri State Fruit Experiment Station.

Discovery of GVCV in *A. cordata*

The initial assay on the *A. cordata* DNA was a standard test for detection of GVCV, consisting of primers designed to yield an 835 bp amplicon if GVCV is present and a quality control band of 105 bp representing 16S ribosomal RNA. The results indicated that GVCV was present in both AMP1P and AMP2P samples (Figure 5).

Further testing with GVCV specific primers was done to verify the initial results, and a 246 bp amplicon was shown to be present as well. Because two PCR assays showed positive results, the decision was made to pursue obtaining the entire genome of the suspected viral isolate present in each *A. cordata* plant. The strategy of using

overlapping fragments to determine an unknown nucleotide sequence is a traditional biotechnological method (96). This technique serves more than one purpose. For this research it not only ensured the entire genome was represented, it also ensured sequencing of only one isolate of GVCV by using three PCR reactions cloned by three discrete bacterial colonies. The technique of using overlapping fragments based on the GVCV-CHA reference sequence was employed to divide the putative viral genomes into three pieces of approximately 3000 bp each. Figure 6 shows three lanes of each of the three fragments found in AMP1P, and it was from this gel that the DNA was extracted, purified and cloned. Because the GVCV isolates in the two *A. cordata* differed from the reference genome and each other, obtaining three fragments from AMP2 DNA proved to be more complicated.

Fragments 1 and 2 of AMP2 were acquired by the same methods as described above in Materials and Methods; however, to obtain Fragment 3 of AMP2 alternate primers had to be utilized because of the slightly differing sequences. After two failed attempts to capture Fragment 3 of AMP2 with the previously designed primer set, two alternative primer sets were attempted that would yield a new Fragment 3 of a larger size. Both attempted alternate fragments performed well in PCR, gel visualization, and DNA purification from the gel. The longer 3,877 bp fragment was chosen for cloning and sequencing. The primers used to obtain this alternate Fragment 3 were 6004F (5'-AGTCTGCCTGGAATCACCTC-3') and 2128R (5'-TACCGTATCCTCTGGTGGTA-3'), and the previously detailed High-Fidelity polymerase PCR protocol (Tables 3 and 4) was likewise used in obtaining this fragment for sequencing. All fragments were cloned

and plasmid DNA was extracted and sent for Sanger sequencing as described in Materials and Methods.

Sequencing and Sequence Analysis

Over a period of seven months, from initial purification of the AMP1 and AMP2 DNA to generation of overlapping fragments, through cloning and sequencing, the two AMP viral isolate genomes were assembled in their entireties. The entire nucleotide sequences may be seen in Appendix A (AMP1) and B (AMP2). This marks the third and fourth GVCV isolates discovered in native, non-domesticated plants, and the first time GVCV has been found in a genus other than *Vitis*. There are now five GVCV isolates that have been fully sequenced, including the GVCV-CHA reference genome, and comparisons have been made between the known sequences in Table 11.

Both AMP1 and AMP2 were found to have an ORFII that was less than 90% identical at the nucleotide level to the GVCV-CHA reference genome, demonstrating that they would both be considered novel isolates. Table 12 shows the percent identity matrix for ORFII at the nucleotide and amino acid levels among all known GVCV isolates, while Figure 7 specifically demonstrates similarities and differences at each nucleotide location. A full comparison of identity percentages of the five known isolates at the nucleotide level is presented in Table 13.

Phylogenomic and Phylogenetic Analysis

Phylogenomic analysis has become essential in uncovering evolutionary relationships between organisms. Multiple sequence analysis is the crucial first step in

phylogenomic comparisons (97). Changes in DNA sequences over time help determine evolutionary rates and are the basis for building phylogenetic trees. MUSCLE was used to align the nucleotide sequences of the five known GVCV genomes and the results were sent to ClustalW2 Phylogeny where an Unweighted Pair Group Method with Arithmetic Mean (UPGMA) tree (98) was built as can be seen in Figure 8A. When the same process was used, but a neighbor joining tree was built, the results were slightly different as can be seen in Figure 8B. By comparing the two methods of phylogenomic tree building it can be observed that AMP1 is considered its own clade when neighbor joining is used, but shares a more recent common ancestor with CHA when the UPGMA method is employed. Two related badnaviruses, *Gooseberry vein banding virus* and *Fig badnavirus* were then included in the UPGMA tree and the results are shown in Figure 8C.

Phylogenetic relationships between the five available GVCV isolates were determined based on the most variable region, ORFII. Both the nucleotide and amino acid levels were used to build phylogenetic trees and the results shown in Figures 8D and 8E indicate that GVCV-CHA has evolved independently for a period of time in relation to the four GVCV isolates that infect wild Vitaceae.

Definitive Primer Set Design

Because ORFII has been found to be the most variable region of the GVCV genome, it has been used to characterize isolates of the virus. A universal primer set was needed that would flank ORFII, which is approximately 300 bp long, so that this region could be amplified and sequenced efficiently. All known sequences of GVCV were compared in CodonCode Aligner and a set of primers was designed that would

potentially capture ORFII in all known isolates and improve efficiency in determining novel isolates of GVCV. The primers designed for this purpose are as follows: 963F (5'-TCCATCACAGATCTAACGGCA-3') and 1634R (5'-CAAGGTAGCGGGCACGAG-3') and would yield an amplicon of approximately 672 bp from any given GVCV isolate. An initial test PCR was performed with the veriflex option to determine the optimal annealing temperature for the new primer set. Results showed that 60°C was the best choice, and the subsequent PCR was programmed in the same way as the initial assay (Table 1 and 2), with the exception of an annealing temperature of 60°C instead of 55°C. The results presented in Figure 9 show that this primer set was effective for all GVCV isolates except VRU2.

Local Vitaceae Survey

A sampling of Vitaceae in and around Springfield, Missouri was conducted in the spring of 2016 in six locations (Table 8). A total of 30 Vitaceae samples, both *Vitis* and *Ampelopsis*, were randomly collected with no regard for visual assessment as "symptomatic" or "asymptomatic". DNA was extracted from the samples and subjected to PCR assays for detection of GVCV. The results of the survey are shown in Table 14 and indicate an incidence of GVCV infection in *A. cordata* of 29% in the Springfield area.

The primer triplex that was used to test these Vitaceae samples for the presence of GVCV included primer 4804R, which does not anneal to the two known GVCV-AMP isolates, therefore the 442 bp fragment is not seen on the gel image even if the *A. cordata* plant tests positive for the 835 bp fragment. This set of primers can be used to

differentiate between GVCV isolates infecting vineyard grapes. In Figure 10 two positive controls were used for the 835, 442, and 16S triplex; one is AMP1 and the other CHA. By observation of the band, or lack thereof, at 442 bp an assumption may be made about the GVCV isolate infecting the plant. If the band at 442 bp is present, the isolate would not be either AMP1 or AMP2 since the 442 bp primer set does not anneal to the sequences of either of these isolates.

DISCUSSION

The results show that *A. cordata* in Missouri is infected with GVCV (Figure 5). The genomes of two viral GVCV isolates were sequenced from *A. cordata* plants that were separated by 155 miles, and shared 89.1% identity at the nucleotide level (Table 12), indicating that they are genetically divergent and that this plant as a GVCV host is not an issue focused solely near Springfield. However, a survey of wild *A. cordata* plants in the Springfield area gives a glimpse into the prevalence of GVCV in native populations of the plant. Approximately 30% (5 out of 17) of randomly sampled *A. cordata* plants in five locations tested positive for GVCV.

In every Missouri vineyard inspected this summer by scientists from the Center for Grapevine Biotechnology (Missouri State University) *A. cordata* has been found growing within 100 yards or less from the vineyard borders. To be sure, this alone may not provide ample evidence that GVCV originally spread to cultivated grapes from native *A. cordata,* admittedly, the opposite scenario could be true. The possibility exists that *A. cordata* plays only a subsidiary role and could be one of many hosts, though so far the only native plants that have been found to host GVCV are *V. rupestris* and *A. cordata*. However, several other clues must be considered. To begin with, an edge effect has been observed with GVCV infected vineyards, virus infected vines are concentrated on the outer rows of plants, those closer to native plant populations. Secondly, the known range of GVCV infection in cultivated grapes lies within the native range of *A. cordata* (Figures 11 and 12). Thirdly, GVCV has been detected in native *V. rupestris*, a plant that not only shares the Vitaceae family with *A. cordata*, but shares habitat in an extremely

proximate manner. Figure 13 shows one of many instances of the intertwined growth habit observed between *Vitis* and *Ampelopsis*. Finally, the prevalence of GVCV in *A. cordata* in the localized survey, where cultivated grapes are much less common than in other Missouri regions, indicates that the probability for GVCV to have spread from vineyard grapes to native Vitaceae populations is lower than the probability for GVCV to have spread from wild plants to cultivated grapes.

Consider also surveys conducted for GVCV in other regions of the United States and globally. In 2013, nine years after GVCV symptoms inspired research into the disease in Missouri, testing for a wide array of grapevine viruses, including GVCV, in 99 samples of grapevines from the United States and Europe found no indication of GVCV in any of the samples (99). This year a survey for GVCV in 384 *Vitis* germplasm samples was conducted in our Missouri State Plant Science laboratory. These samples were generously provided to Missouri State University by Jason Londo from the National Plant Germplasm Collection in Geneva, NY. The samples represented native and cultivated *Vitis* species from locations around the globe. GVCV was not detected in any of these samples (Wenschel et al., unpublished). These assays combined with the incidence of GVCV detected in the Midwest United States bolsters the plausibility of our hypothesis that GVCV is emergent in this region, the indigenous range of *A. cordata* and several wild *Vitis* species.

Decades of previous research have solidified the model of viral emergence in cultivated plants. Though viral emergence is extremely complex, the occurrence is simplified to three phases: 1) The virus that has been maintained in a plant population (typically the reservoir is a wild plant population) encounters the new host. Preexisting

genetic variation of the virus and fitness of the host determine the outcome of the encounter (i.e., infection of the new host plant occurs, or not). 2) The virus adapts to the new host in such a way that transmission of the virus between individuals of the new host is possible without any further spillover from the reservoir. 3) Epidemiology dynamics shift so that between-host transmissions in the new plant population are optimized (25, 28, 30, 100). It could be speculated that GVCV was transmitted to native *Vitis* species from *A. cordata*, adapted to native *Vitis*, made a facile transition to cultivated *Vitis*, and is now in the third phase of viral emergence.

The detection and sequencing of two GVCV isolates in *A. cordata* provides new and salient information on the genetic changes in the viral populations among native plants in two host genera of Vitaceae. Of particularly notable interest is the high variability of ORF II, in part due to the 9 bp insertion/deletion. Based on phylogenetic trees constructed using both the nucleotide and amino acid sequences of ORF II of all five GVCV genomes (Figure 8D and 8E), it is clear that the four isolates found in native plants are more closely related to each other than they are to GVCV-CHA, which infects grape cultivars in commercial vineyards. Phase 2 of the above model of viral emergence states that there are genetic changes in a virus as it adapts to a new host, changes that scientists can now map through sequencing. The sequencing of two isolates of GVCV found in a different plant genus provides the data for beginning to map these genetic changes and track the evolution of the virus. *A. cordata* DNA samples which tested positive for GVCV in the localized survey may now be used, along with the designed definitive primers for ORF II, for sequencing of this variable region so that comparisons may be made and inferences drawn based on genetic variations. This, along with

comparisons of homologous sequences of ORF II in cultivated grapes, should reveal much about the origin and evolution of GVCV.

Implications of this study also include vineyard pest and disease management practices. One of the most sustainable approaches for management of GVCV-associated disease is selection and deployment of host plant resistance. In light of the mild symptoms present on infected *A. cordata*, perhaps future work could uncover the genetic source of tolerance/resistance these plants possess for transfer to cultivated grapes. *A. cordata* is, at the least, a host plant or reservoir for GVCV, and possibly the source of the virus's spillover into vineyard grapes. Either way, consideration should be taken when selecting and planning a new vineyard site as to the native population of *A. cordata* and *V. rupestris*. Proximity and density of these wild plant populations should be taken into account. One pressing issue for effective management of GVCV is discovery of the transmission vector(s). There is now an additional plant to observe *in situ* as to which insects feed on this native host as well as cultivated grape hosts.

Grapevines are transported and traded at a global level. Since this is the case, it would behoove any distributor of stock from a GVCV infected region to create protocol for the detection of GVCV before sending out plant material. Standard tests for virus-free vines, especially in the Midwest, should include detection of GVCV to ensure that nurseries are not distributing the virus with grapevine sales. The virus appears to be endemic to the Midwest United States and necessary precautions should be taken to minimize the loss caused locally and to keep it from spreading to unaffected areas.

The economic impact of GVCV on grape production has not yet been quantified, but a comprehensive survey of vineyard damage in Missouri, and eventually in the

Midwest, would be an important step in assessing the severity of GVCV associated disease. Vineyard managers could be requested to supply number of vines infected and number of vines lost each year, while researchers might assess fruit set reduction, sugar concentration shifts, photosynthetic effects of the viral disease, etc. A quantitative study of profit loss for the grape industry that is attributable to GVCV could encourage urgency in research and funding. Most importantly, containing the spread of the virus to protect other major viticulture regions throughout the world should be the motivation of future research and education.

From a broad perspective, the addition of two more badnavirus genome sequences to GenBank will hopefully serve all involved in research and exploration of the *Caulimoviridae* family of plant viruses by contributing more information on the genetic nature of this family of viruses infecting wild and cultivated Vitaceae plants.

REFERENCES

1. **Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P.** 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends in Ecology & Evolution **19:**535-544.

2. **Zhang Y, Singh K, Kaur R, Qiu W.** 2011. Association of a Novel DNA Virus with the Grapevine Vein-Clearing and Vine Decline Syndrome. Phytopathology **101:**1081-1090.

3. **Guo Q, Honesty S, Xu ML, Zhang Y, Schoelz J, Qiu W.** 2014. Genetic Diversity and Tissue and Host Specificity of Grapevine vein clearing virus. Phytopathology **104:**539-547.

4. **Müller D, Stache B.** 1992. Worldwide occurrence of virus-infections in filamentous marine brown algae. Helgoländer Meeresuntersuchungen **46:**1-8.

5. **Hull R.** 1968. A virus disease of hart's-tongue fern. Virology **35:**333-335.

6. **Brunt AA, Crabtree K, Gibbs A.** 1990. Viruses of tropical plants. Descriptions and lists from the VIDE database. CAB International.

7. **Yahara T, Oyama K.** 1993. Effects of virus infection on demographic traits of an agamospermous population of Eupatorium chinense (Asteraceae). Oecologia **96:**310-315.

8. **Rodríguez-Cerezo E, Elena SF, Moya A, García-Arenal F.** 1991. High genetic stability in natural populations of the plant RNA virus tobacco mild green mosaic virus. Journal of Molecular Evolution **32:**328-332.

9. **Raybould AF, Maskell LC, Edwards ML, Cooper JI, Gray AJ.** 1999. The Prevalence and Spatial Distribution of Viruses in Natural Populations of Brassica oleracea. The New Phytologist **141:**265-275.

10. **MacClement WD, Richards MG.** 1956. Virus in Wild Plants. Canadian Journal of Botany **34:**793-799.

11. **Cooper I, Jones RA.** 2006. Wild Plants and Viruses: Under-Investigated Ecosystems. Advances in virus research **67:**1-47.

12. **Coutts BA, Jones RAC.** 2000. Viruses infecting canola (Bassica napus) in south-west Australia: incidence, distribution, spread, and infection reservoir in wild radish (Raphanus raphinistrum). Australian Journal of Agricultural Research **51:**925-936.

13. **Remold SK.** 2002. Unapparent Virus Infection and Host Fitness in Three Weedy Grass Species. Journal of Ecology **90:**967-977.

14. **Lindbo JA, Silva-Rosales L, Proebsting WM, Dougherty WG.** 1993. Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. The Plant Cell **5:**1749-1759.

15. **Wilson TM.** 1993. Strategies to protect crop plants against viruses: pathogen-derived resistance blossoms. Proceedings of the National Academy of Sciences **90:**3134-3141.

16. **Rodoni B.** 2009. The role of plant biosecurity in preventing and controlling emerging plant virus disease epidemics. Virus Research **141:**150-157.

17. **Lovisolo O, Hull R, Rösler O.** 2003. Coevolution of viruses with hosts and vectors and possible paleontology. Advances in virus research **62:**325-379.

18. **Jones JD, Dangl JL.** 2006. The plant immune system. Nature **444:**323-329.

19. **Harper J, Burdon J, Leather S.** Pests, pathogens and plant communities: an introduction, p 3-14. *In* (ed), Blackwell Scientific Publications,

20. **Attafuah A.** 1965. Occurrence of cocoa viruses in wild plant species in Ghana. Ghana journal of science.

21. **Pagán I, González-Jara P, Moreno-Letelier A, Rodelo-Urrego M, Fraile A, Piñero D, García-Arenal F.** 2012. Effect of biodiversity changes in disease risk: exploring disease emergence in a plant-virus system. PLoS Pathog **8:**e1002796.

22. **Prendeville HR, Ye X, Jack Morris T, Pilson D.** 2012. Virus infections in wild plant populations are both frequent and often unapparent. American Journal of Botany **99:**1033-1042.

23. **Hammond J.** 1982. Plantago As a Host of Economically Important Viruses, p 103-140. *In* Max A. Lauffer FBBKM, Kenneth MS (ed), Advances in Virus Research, vol Volume 27. Academic Press.

24. **Williams E.** 1991. Natural incidence of peanut stunt virus infection in hybrid populations of Trifolium ambiguum x T. repens. Plant disease **75**.

25. **Roossinck MJ, García-Arenal F.** 2015. Ecosystem simplification, biodiversity loss and plant virus emergence. Current opinion in virology **10:**56-62.

26. **Fargette D, Konate G, Fauquet C, Muller E, Peterschmitt M, Thresh J.** 2006. Molecular ecology and emergence of tropical plant viruses. Annu Rev Phytopathol **44:**235-260.

27. **Luo H, Wylie SJ, Coutts B, Jones RAC, Jones MGK.** 2011. A virus of an isolated indigenous flora spreads naturally to an introduced crop species. Annals of Applied Biology **159:**339-347.

28. **Elena SF, Fraile A, García-Arenal F.** 2014. Evolution and emergence of plant viruses.

29. **Hanssen IM, Thomma BPHJ.** 2010. Pepino mosaic virus: a successful pathogen that rapidly evolved from emerging to endemic in tomato crops. Molecular Plant Pathology **11:**179-189.

30. **Moreno-Pérez MG, Pagán I, Aragón-Caballero L, Cáceres F, Fraile A, García-Arenal F.** 2014. Ecological and genetic determinants of Pepino mosaic virus emergence. Journal of virology **88:**3359-3368.

31. **Stubbs L, Guy J, Stubbs K.** 1963. Control of lettuce necrotic yellows virus disease by the destruction of common sowthistle (<I>Sonchus oleraceus</I>). Australian Journal of Experimental Agriculture **3:**215-218.

32. **Cho J, Mau R, Mitchell W, Gonsalves D, Yudin L.** 1987. Host list of plants susceptible to tomato spotted wilt virus (TSWV).

33. **Cho J, Mau R, Gonsalves D, Mitchell W.** 1986. Reservoir weed hosts of tomato spotted wilt virus. Plant Disease **70:**1014-1017.

34. **D'Arcy CJ, Burnett PA.** 1995. Barley yellow dwarf: 40 years of progress. American Phytopathological Society (APS Press).

35. **Pavek DS, Lamboy WF, Garvey EJ.** 2003. Selecting in situ conservation sites for grape genetic resources in the USA. Genetic Resources and Crop Evolution **50:**165-173.

36. **USDA N.** 6-25-16. The PLANTS Database. http://plants.usda.gov. Accessed

37. **Klaassen VA, Sim ST, Dangl GS, Osman F, Rwahnih MA, Rowhani A, Golino DA.** 2011. Vitis californica and Vitis californica × Vitis vinifera Hybrids are Hosts for Grapevine leafroll-associated virus-2 and -3 and Grapevine virus A and B. Plant Disease **95:**657-665.

38. **Pacifico D, Stigliano E, Sposito L, Spinelli P, Garfì G, Gristina AS, Fontana I, Carimi F.** 2016. Survey of viral infections in spontaneous grapevines from natural environments in Sicily. European Journal of Plant Pathology **145:**189-197.

39. **Regner F, Hack R, Gangl H, Leitner G, Mandl K, Tiefenbrunner W.** 2015. Genetic variability and incidence of systemic diseases in wild vines (Vitis vinifera ssp. silvestris) along the Danube. VITIS-Journal of Grapevine Research **43:**123.

40. **Vivier M, Pretorius IS.** 2000. Genetic improvement of grapevine: Tailoring grape varieties for the third millennium-a review. South African Journal for Enology and Viticulture **21:**5-26.

41. **Terral J-F, Tabard E, Bouby L, Ivorra S, Pastor T, Figueiral I, Picq S, Chevance J-B, Jung C, Fabre L, Tardy C, Compan M, Bacilieri R, Lacombe T, This P.** 2010. Evolution and history of grapevine (Vitis vinifera) under domestication: new morphometric perspectives to understand seed domestication syndrome and reveal origins of ancient European cultivars. Annals of Botany **105:**443-455.

42. **USDA.** 4-29-16 2016. National Agricultural Statistics Service. https://www.nass.usda.gov/. Accessed 6-25-16.

43. **Frank RC, LLP.** 2015. The Economic Impact of Missouri Wine and Wine Grapes — 2013.

44. **Martelli GP.** 2014. Directory of virus and virus-like diseases of the grapevine and their agents. Journal of Plant Pathology **96:**1-136.

45. **López-Fabuel I, Wetzel T, Bertolini E, Bassler A, Vidal E, Torres LB, Yuste A, Olmos A.** 2013. Real-time multiplex RT-PCR for the simultaneous detection of the five main grapevine viruses. Journal of virological methods **188:**21-24.

46. **Fuller KB, Alston JM, Golino DA.** 2013. The benefits from certified virus-free nursery stock: A case study of grapevine leafroll-3 in the North Coast region of California. Robert Mondavi Institute-Center for Wine Economics University of California**:**37.

47. **Binzen Fuller K, Alston JM, Golino DA.** 2015. The Economic Benefits from Virus-Screening: A Case Study of Grapevine Leafroll in the North Coast of California. American Journal of Enology and Viticulture doi:10.5344/ajev.2014.14055.

48. **Montero R, Mundy D, Albright A, Grose C, Trought M, Cohen D, Chooi K, MacDiarmid R, Flexas J, Bota J.** 2016. Effects of Grapevine Leafroll associated Virus 3 (GLRaV-3) and duration of infection on fruit composition and wine chemical profile of Vitis vinifera L. cv. Sauvignon blanc. Food chemistry **197:**1177-1183.

49. **Carlson PS.** 2013. The biology of crop productivity. Elsevier.

50. **Almeida RP, Daane KM, Bell VA, Blaisdell GK, Cooper ML, Herrbach E, Pietersen G.** 2013. Ecology and management of grapevine leafroll disease.

51. **Rowhani A, Uyemoto JK, Golino DA, Martelli GP.** 2005. Pathogen Testing and Certification of Vitis and Prunus Species*. Annu Rev Phytopathol **43:**261-278.

52. **Maliogka VI, Martelli GP, Fuchs M, Katis NI.** 2015. Chapter Six - Control of Viruses Infecting Grapevine, p 175-227. *In* Gad L, Nikolaos IK (ed), Advances in Virus Research, vol Volume 91. Academic Press.

53. **Bisztray GD, Civerolo EL, Dula T, Kolber M, Lazar J, Mugnai L, Szegedi E, Savka MA, Szabo P, Shojania J.** 2012. Grapevine pathogens spreading with propagating plant stock: Detection and methods for elimination. Grapevines: Varieties, cultivation, and management Nova Science Pub, Hauppauge, NY**:**1-86.

54. **Duffus JE.** 1971. Role of weeds in the incidence of virus diseases. Annual Review of Phytopathology **9:**319-340.

55. **Thresh J.** 2003. Control of plant virus diseases in sub-Saharan Africa: the possibility and feasibility of an integrated approach. African Crop Science Journal **11:**199-223.

56. **Bota J, Cretazzo E, Montero R, Rossello J, Cifre J.** 2014. Grapevine fleck virus (GFKV) elimination in a selected clone of Vitis vinifera L. CV. Manto Negro and its effects on photosynthesis. J Int Sci Vigne Vin **48**.

57. **Chooi KM, Cohen D, Pearson MN.** 2016. Differential distribution and titre of selected grapevine leafroll-associated virus 3 genetic variants within grapevine rootstocks. Archives of Virology **161:**1371-1375.

58. **Bragard C, Caciagli P, Lemaire O, Lopez-Moya J, MacFarlane S, Peters D, Susi P, Torrance L.** 2013. Status and prospects of plant virus control through interference with vector transmission. Annual review of phytopathology **51:**177-201.

59. **Pietersen G, Spreeth N, Oosthuizen T, van Rensburg A, van Rensburg M, Lottering D, Rossouw N, Tooth D.** 2013. Control of grapevine leafroll disease spread at a commercial wine estate in South Africa: a case study. American Journal of Enology and Viticulture**:**ajev. 2013.12089.

60. **Chiumenti M, Giampetruzzi A, Morelli M, Savino VN, Martelli GP, La Notte P, Palmisano F, Saldarelli P.** 2016. Detection and molecular characterization of a Grapevine Roditis leaf discoloration-associated virus (GRLDaV) variant in an autochthonous grape from Apulia (Italy). Virus genes **52:**428-431.

61. **Krenz B, Thompson JR, Fuchs M, Perry KL.** 2012. Complete genome sequence of a new circular DNA virus from grapevine. Journal of virology **86:**7715-7715.

62. **Rwahnih MA, Dave A, Anderson MM, Rowhani A, Uyemoto JK, Sudarshana MR.** 2013. Association of a DNA virus with grapevines affected by red blotch disease in California. Phytopathology **103:**1069-1076.

63. **Maliogka VI, Olmos A, Pappi PG, Lotos L, Efthimiou K, Grammatikaki G, Candresse T, Katis NI, Avgelis AD.** 2015. A novel grapevine badnavirus is associated with the Roditis leaf discoloration disease. Virus research **203:**47-55.

64. **Geering ADW, Maumus F, Copetti D, Choisne N, Zwickl DJ, Zytnicki M, McTaggart AR, Scalabrin S, Vezzulli S, Wing RA, Quesneville H, Teycheney P-Y.** 2014. Endogenous florendoviruses are major components of plant genomes and hallmarks of virus evolution. Nat Commun **5**.

65. **Shepherd R, Wakeman R, Romanko R.** 1968. DNA in cauliflower mosaic virus. Virology **36:**150-152.

66. **Bhat AI, Hohn T, Selvarajan R.** 2016. Badnaviruses: The Current Global Scenario. Viruses **8:**177.

67. **Schoelz JE.** 2008. Caulimoviruses: General features. Encyclopedia of Virology **1:**457-464.

68. **Hohn T, Rothnie H.** 2013. Plant pararetroviruses: replication and expression. Current opinion in virology **3:**621-628.

69. **Borah BK, Sharma S, Kant R, Johnson A, Saigopal DVR, Dasgupta I.** 2013. Bacilliform DNA-containing plant viruses in the tropics: commonalities within a genetically diverse group. Molecular plant pathology **14:**759-771.

70. **Schoelz J, Bourque J.** 2008. Caulimoviruses: general features. En: Mahy, BWJ, and MHV Van Regenmortel (eds) Desk encyclopedia of plant and fungal virology Academic Press, San Diego, CA**:**147-155.

71. **Zhang Y, Angel CA, Valdes S, Qiu W, Schoelz JE.** 2015. Characterization of the promoter of Grapevine vein clearing virus. Journal of General Virology **96:**165-169.

72. **Cheng C-P, Lockhart B, Olszewski NE.** 1996. The ORF I and II Proteins ofCommelinaYellow Mottle Virus Are Virion-Associated. Virology **223:**263-271.

73. **Jacquot E, Hagen LS, Jacquemond M, Yot P.** 1996. The open reading frame 2 product of cacao swollen shoot badnavirus is a nucleic acid-binding protein. Virology **225:**191-195.

74. **Hohn T, Fütterer J, Hull R.** 1997. The proteins and functions of plant pararetroviruses: knowns and unknowns. Critical reviews in plant sciences **16:**133-161.

75. **Knowles N, Hovi T, Hyypiä T, King A, Lindberg AM, Pallansch M, Palmenberg A, Simmonds P, Skern T, Stanway G.** 2012. Virus Taxonomy, Ninth Report of the International Committee on Taxonomy of Viruses: Picornaviridae.

76. **Bhat AI, Sasi S, Revathy KA, Deeshma KP, Saji KV.** 2014. Sequence diversity among badnavirus isolates infecting black pepper and related species in India. VirusDisease **25:**402-407.

77. **Lockhart BEL, Kiratiya-Angul K, Jones P, Eng L, De Silva P, Olszewski NE, Lockhart N, Deema N, Sangalang J.** 1997. Identification of Piper yellow mottle virus, a mealybug-transmitted badnavirus infecting Piper spp. in Southeast Asia. European Journal of Plant Pathology **103:**303-311.

78. **Lockhart BE, Menke J, Dahal G, Olszewski N.** 2000. Characterization and genomic analysis of tobacco vein clearing virus, a plant pararetrovirus that is transmitted vertically and related to sequences integrated in the host genome. Journal of General Virology **81:**1579-1585.

79. **Ahlawat Y, Pant R, Lockhart B, Srivastava M, Chakraborty N, Varma A.** 1996. Association of a badnavirus with citrus mosaic disease in India. Plant disease **80:**590-592.

80. **Kenyon L, Lebas B, Seal S.** 2008. Yams (Dioscorea spp.) from the South Pacific Islands contain many novel badnaviruses: implications for international movement of yam germplasm. Archives of virology **153:**877-889.

81. **Geering A, McMichael L, Dietzgen R, Thomas J.** 2000. Genetic diversity among Banana streak virus isolates from Australia. Phytopathology **90:**921-927.

82. **Eni A, Hughes JdA, Asiedu R, Rey M.** 2008. Sequence diversity among badnavirus isolates infecting yam (Dioscorea spp.) in Ghana, Togo, Benin and Nigeria. Archives of virology **153:**2263-2272.

83. **Muller E, Dupuy V, Blondin L, Bauffe F, Daugrois J-H, Nathalie L, Iskra-Caruana M-L.** 2011. High molecular variability of sugarcane bacilliform viruses in Guadeloupe implying the existence of at least three new species. Virus research **160:**414-419.

84. **Muller E, Sackey S.** 2005. Molecular variability analysis of five new complete cacao swollen shoot virus genomic sequences. Archives of virology **150:**53-66.

85. **Ortiz R.** 1996. The potential of AMMI analysis for field assessment of Musa genotypes to virus infection. HortScience **31:**829-832.

86. **Dahal, Ortiz, Tenkouano, Hughes, Thottappilly, Vuylsteke, Lockhart.** 2000. Relationship between natural occurrence of banana streak badnavirus and symptom expression, relative concentration of viral antigen, and yield characteristics of some micropropagated Musa spp. Plant Pathology **49:**68-79.

87. **Wambulwa M, Wachira F, Karanja L, Kiarie S, Muturi S.** 2013. The influence of host and pathogen genotypes on symptom severity in banana streak disease. African Journal of Biotechnology **12:**27.

88. **Karanja L, Wangai A, Pathak RS, Harper G.** 2013. Effect of environment and cultivar on the expression of banana streak disease symptoms in Kenya. African Journal of Biotechnology **12**.

89. **Lehrer AT, Komor E.** 2008. Symptom expression of yellow leaf disease in sugarcane cultivars with different degrees of infection by Sugarcane yellow leaf virus. Plant Pathology **57:**178-189.

90. **Morano LD, Walker MA.** 1995. Soils and Plant Communities Associated with Three Vitis Species. The American Midland Naturalist **134:**254-263.

91. **Vogelstein B, Gillespie D.** 1979. Preparative and analytical purification of DNA from agarose. Proceedings of the National Academy of Sciences **76:**615-619.

92. **Cohen SN, Chang AC, Boyer HW, Helling RB.** 1973. Construction of biologically functional bacterial plasmids in vitro. Proceedings of the National Academy of Sciences **70:**3240-3244.

93. **Sanger F, Coulson AR.** 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of molecular biology **94:**441-448.

94. **Ewing B, Green P.** 1998. Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. Genome Research **8:**186-194.

95. **Darwin C.** 1859. The origin of species by means of natural selection or the preservation of favoured races in the struggle for life.

96. **Chao S, Sederoff RR, Levings CS.** 1983. Partial Sequence Analysis of the 5s to 18s rRNA Gene Region of the Maize Mitochondrial Genome. Plant Physiology **71:**190-193.

97. **Kumar S, Filipski aA.** 2007. Multiple sequence alignment:
In pursuit of homologous DNA positions. Genome Research **17:**127–135.

98. **Sokal RR, Michener CD.** 1958. A statistical method for evaluating systematic relationships. University of Kansas Scientific Bulletin **28:**1409-1438.

99. **Thompson JR, Fuchs M, McLane H, Celebi-Toprak F, Fischer KF, Potter JL, Perry KL.** 2013. Profiling Viral Infections in Grapevine Using a Randomly Primed Reverse Transcription-Polymerase Chain Reaction/Macroarray Multiplex Platform. Phytopathology **104:**211-219.

100. **Jones RA.** 2009. Plant virus emergence and evolution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. Virus research **141:**113-130.

Table 1. PCR master mix components for initial assay.

Reagent	**Concentration**
Autoclaved dH_2O	To 25 µl
5X buffer	1X
1101F primer	0.3 µM
1835R primer	0.3 µM
16SF primer	0.08 µM
16SR primer	0.08 µM
dNTPs	0.2 mM
Taq polymerase	1.25 units
DNA	0.4 ng/µl

Table 2. Thermocycler PCR program for initial assay.

Step	**Cycle**	**Temperature**	**Time**
Initial denaturing	1	94°C	1 minute
Denaturing	35	94°C	30 seconds
Annealing	35	55°C	30 seconds
Extension	35	72°C	1 minute
Final extension	1	72°C	10 minutes

Table 3. High-Fidelity PCR master mix components.

Reagent	**Concentration**
Autoclaved dH_2O	To 25 µl
10X hifi buffer	1X
F primer	0.2 µM
R primer	0.2 µM
50mM $MgSO_4$	2 mM
dNTPs	0.2 mM
High-Fidelity polymerase	1 unit
DNA	0.4 ng/µl

Table 4. Thermocycler PCR program for high-fidelity polymerase.

Step	**Cycle**	**Temperature**	**Time**
Initial denaturing	1	94°C	2 minutes
Denaturing	35	94°C	30 seconds
Annealing	35	55°C	40 seconds
Extension	35	68°C	4 minutes
Final extension	1	68°C	10 minutes

Table 5. Primers used in sequencing AMP1 (excluding GW1 and GW2).

Primer	Sequence 5' to 3'	T_{m} °C	Location
1101F	CTGAAAGGTAGATGTCCACG	60.4	
1869F	TGGTACGAGAAGGTATGCAGC	62.6	
1915F	AGAATACAAGTGCTACACCGA	62.4	
2112R[1]	GCAGGTGGTGGTAGAAATCAT	60.6	
2128R	TACCGTATCCTCTGGTGGTA	60.4	Fragment 1
2507R	TTCGAAGGTTCCAACTAGGGC	62.6	988-4,387
2569F[1]	CGGAGGAGAATGGCTGGGTAA	64.5	
3122R	GCTAAAACTTTCGAGCTAAC	56.3	
3163F	AGGGTAAAAACTGCGACGGCTA	62.7	
3468F	ATCCTCCCTCCTGAAGTAGC	62.4	
3615R	TTCTCTTTCCCTTGGTCC	57.6	
4363F	ATCTGCTCAATTTCTGAAGGAGAAG	61.3	
4632F	ACTATACTAGGTCGACGTGC	60.4	Fragment 2
4828F	AAACAGGAACTCCAAGCTGC	60.4	4,142-6,795
5405F	CAGCCTTCGAAATGAACATGC	60.6	
6004F	AGTCTGCCTGGAATCACCTC	62.4	
6158R	GCCATTCATATAGTCCTGCG	60.4	
6407F	GAAATAGCAGAGTGAGTCTG	58.4	
6690R	GATAACTGCGTGGGGTGGAG	64.5	
7635F	CCAGTTCCAGTTCCAGTGTTCTTAATGC	66.1	Fragment 3
217R	TCTCACAACGGGCTACTACC	62.4	6,666-1,935
1179R	GCCACGTGGACATCTACCTT	62.4	
1615R	GTATTCGTTGCTCCGCAAG	60.2	
1827R	CTGCCGGTCTATGACATGGG	64.5	

[1] Primer designed specifically for AMP genomes.

Table 6. Primers used in sequencing AMP2 (excluding GW1 and GW2).

Primer	Sequence 5' to 3'	T_{m} °C	Location
988F	ACCTAAGCCGATTGAAGCAG	60.4	
1101F	CTGAAAGGTAGATGTCCACG	60.4	Fragment 1
1869F	TGGTACGAGAAGGTATGCAGC	62.6	988-4,387
2109F	TACCACCAGAGGATACGGTA	60.4	
2112R[1]	GCAGGTGGTGGTAGAAATCAT	60.6	
2460F	AGACACAGGAGAAAGGGTAACT	60.8	
2569F[1]	CGGAGGAGAATGGCTGGGTAA	64.5	
2589R	TCCTTCCACATGTTTCACCC	60.4	
3163F	AGGGTAAAAACTGCGACGGCTA	62.7	
3468F	ATCCTCCCTCCTGAAGTAGC	62.4	
4162R	CATGAGAGTCATGAGGTTTAC	60.8	
4387R	CTTCTCCTTCAGAAATTGAGCAGAT	61.3	
4363F	ATCTGCTCAATTTCTGAAGGAGAAG	61.3	Fragment 2
4632F	ACTATACTAGGTCGACGTGC	60.4	4,142-6,795
4828F	AAACAGGAACTCCAAGCTGC	60.4	
5117F	CGTATACTGAAATGCCTCG	58.0	
5558R	GCTATCTTTTGAGGTTCCAGGG	62.7	
5755F	GATATCACCATTGAGGCAAAGC	60.8	
5880R	GGACCTATGTCTGCTCTTGC	62.4	
6158R	GCCATTCATATAGTCCTGCG	60.4	
6690R	GATAACTGCGTGGGGTGGAG	64.5	Fragment 3
6004F	AGTCTGCCTGGAATCACCTC	62.4	6,004-2,128
6606F	GATAACTGCGTGGGGTGGAG	55.1	
81F	AATCGTGTAGGGAATCGTTA	56.3	
217R	TCTCACAACGGGCTACTACC	62.4	
697F	GCTGCTGAATACACTGTACG	60.4	
1535R	GCACGAGGCAACTCAGTGGTCT	66.4	
2128R	TACCGTATCCTCTGGTGGTA	60.4	

[1] Primer designed specifically for AMP genomes

Table 7. Primers designed for GVCV-CHA that did not work for GVCV-AMP1 and GVCV-AMP2.

Primer	Sequence 5' to 3'	T_m °C	Location
697F[1]	GCTGCTGAATACACTGTACG	60.4	Fragment 3
922R	TGACTGATTAGCCTTGAT	53.7	Fragment 3
1915F[2]	AGAATACAAGTGCTACACCGA	58.7	Fragment 1
3074R	GCTGGTAGTGTCGAAGATAGG	62.6	Fragment 1
3615R	TTCTCTTTCCCTTGGTCC	57.6	Fragment 1
4804R	GGAATGCATTGTGCTCGTAG	60.4	Fragment 2
7068F	AAGGCTTGCCCAGAATGT	57.6	Fragment 3

[1] Primer did not anneal to AMP1
[2] Primer did not anneal to AMP2

Table 8. Geographic locations of Vitaceae plants in a local survey.

Map Number	Plant Samples	Location	Latitude	Longitude
1	6 *Ampelopsis* 4 *Vitis*	Sunset and Kansas	37°10’4”	93°19’3”
2	2 *Ampelopsis* 3 *Vitis*	Close Memorial Park	37°10’13”	93°19’31”
3	2 *Ampelopsis* 3 *Vitis*	Confluence of Wilson and South Creeks	37°8’54”	93°22’15”
4	5 *Ampelopsis* 3 *Vitis*	Lake Springfield Park	37°6’52”	93°15’50”
5	1 *Ampelopsis* 1 *Vitis*	Ritter Springs Park	37°16’50”	93°20’36”
6	1 *Ampelopsis* 1 *Vitis*	Dr. Qiu’s yard		

Table 9. PCR master mix components for local Vitaceae survey.

Reagent	**Concentration**
Autoclaved dH_2O	To 25 µl
5X buffer	1X
1101F primer	0.4 µM
1835R primer	0.4 µM
4363F primer	0.24 µM
4804R primer	0.24 µM
16SF primer	0.14 µM
16SR primer	0.14 µM
dNTPs	0.28 mM
Taq polymerase	1.25 units
DNA	0.4 ng/µl

Table 10. Thermocycler PCR program for local Vitaceae survey.

Cycle	**Repetitions**	**Temperature**	**Time**
Initial denaturing	1	94°C	1 minute
Denaturing	35	94°C	30 seconds
Annealing	35	55°C	30 seconds
Extension	35	72°C	1 minute
Final extension	1	72°C	10 minutes

Table 11. Comparative analysis of genome and regional lengths in five GVCV isolates.

GVCV isolate	Genome length in nucleotides (nt)	IGR	ORFI	ORFII	ORFIII
CHA	7,753 nt	7,321-7,753; 1-484 917 nt	485-1,111 627 nt	1,112-1,495 384 nt	1,495-7,320 5,826 nt
VRU1	7,755 nt	7,332-7,755; 1-483 907 nt	484-1,110 627 nt	1,111-1,503 393 nt	1,503-7,331 5,829 nt
VRU2	7,726 nt	7,317-7,726; 1-474 884 nt	475-1,104 630 nt	1,105-1,488 384 nt	1,488-7,316 5,829 nt
AMP1	7,749 nt	7,336-7,749; 1-481 895 nt	482-1,108 627 nt	1,109-1,501 393 nt	1,501-7,335 5,835 nt
AMP2	7,765 nt	7,341-7,765; 1-486 911 nt	487-1,116 630 nt	1,117-1,509 393 nt	1,509-7,340 5,832 nt

Table 12. Percent Identity matrix of five GVCV isolates' ORF II. Italicized numbers represent amino acid level, and non-italicized numbers represent nucleotide level.

	CHA	VRU1	VRU2	AMP1	AMP2
CHA	100	*85.5*	*93.0*	*88.5*	*90.8*
VRU1	83.0	100	*90.8*	*92.4*	*93.9*
VRU2	88.1	88.3	100	*91.6*	*90.8*
AMP1	83.2	88.5	89.3	100	*93.9*
AMP2	88.0	92.9	88.0	89.1	100

Table 13. Percent identity matrix of five GVCV genomes at the nucleotide level.

	CHA	VRU1	VRU2	AMP1	AMP2
CHA	100	91.6	91.7	92.3	92.8
VRU1	91.6	100	93.2	91.9	92.2
VRU2	91.7	93.2	100	91.8	91.6
AMP1	92.3	91.9	91.8	100	92.4
AMP2	92.8	92.2	91.6	92.4	100

Table 14. Results of local Vitaceae Survey.

Map Number	Location	Genus	Results
1	Sunset and Kansas	6 *Ampelopsis* 4 *Vitis*	1 positive, 5 negative 4 negative
2	Close Memorial Park	2 *Ampelopsis* 3 *Vitis*	1 positive, 1 negative 3 negative
3	Confluence of Wilson and South Creeks	2 *Ampelopsis* 3 *Vitis*	1 positive, 1 negative 3 negative
4	Lake Springfield Park	5 *Ampelopsis* 3 *Vitis*	5 negative 3 negative
5	Ritter Springs Park	1 *Ampelopsis* 1 *Vitis*	1 positive 1 negative
6	Dr. Qiu's yard	1 *Ampelopsis* 1 *Vitis*	1 positive 1 positive

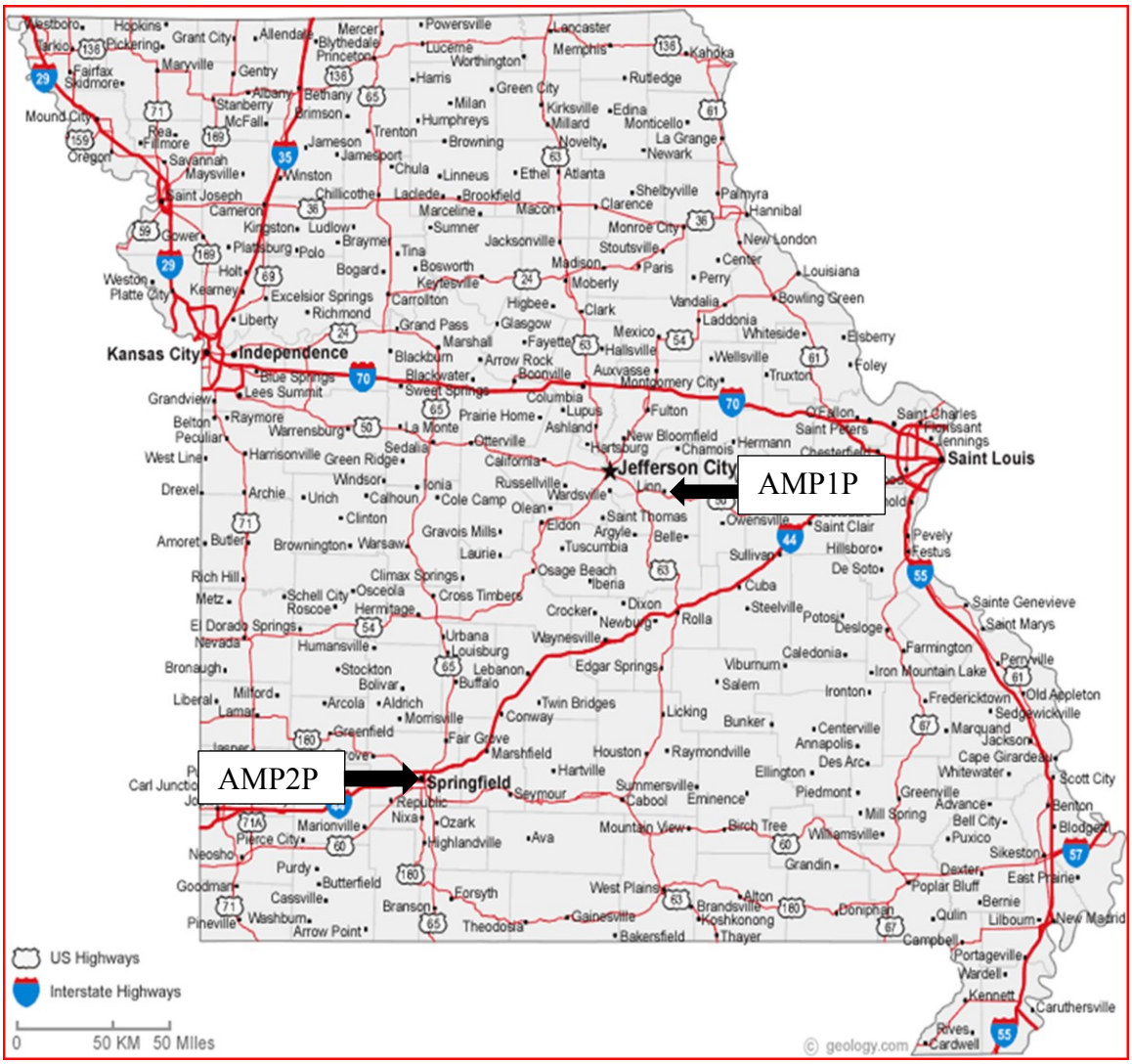

Figure 1. Mapped locations of two *A. cordata* accessions used in this research, AMP1P and AMP2P, native wild plants that were collected and tested for the presence of GVCV.

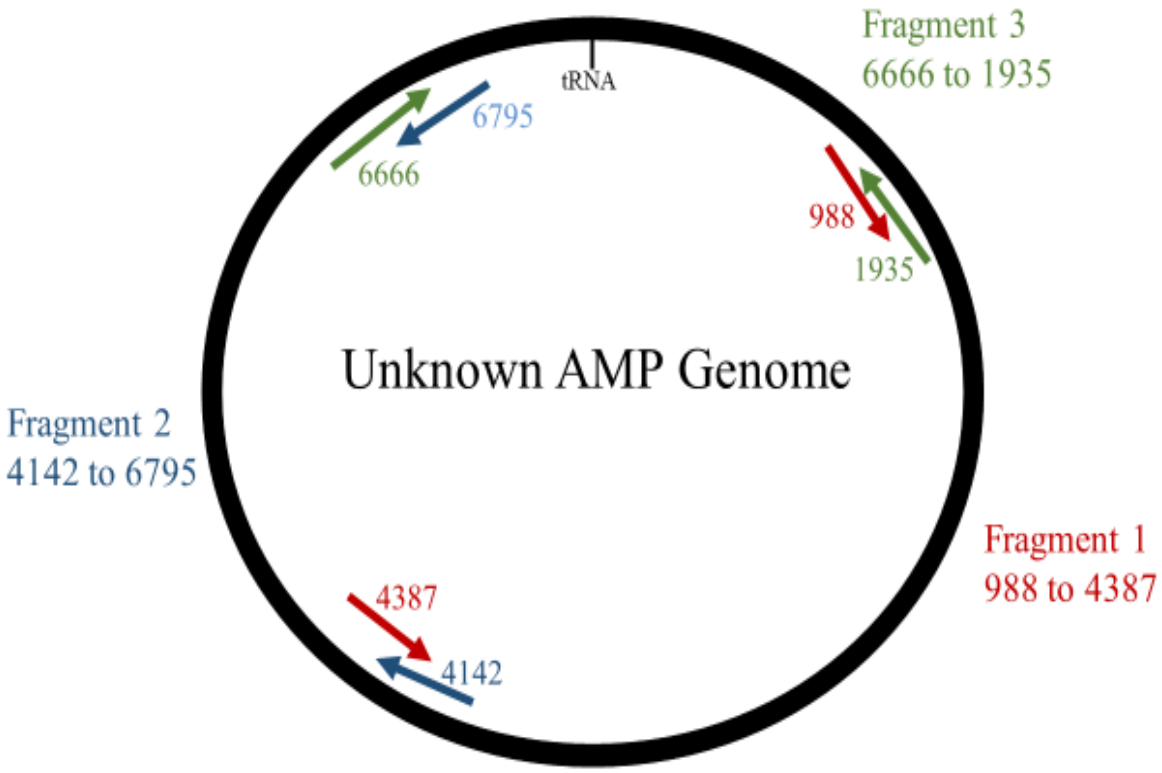

Figure 2. Genome map of overlapping fragments. Locations of three sets of primers that were used in PCR for acquiring the unknown AMP viral genomes. Fragment 1 ≈3,400 bp. Fragment 2 ≈2,654 bp. Fragment 3 ≈3,023 bp.

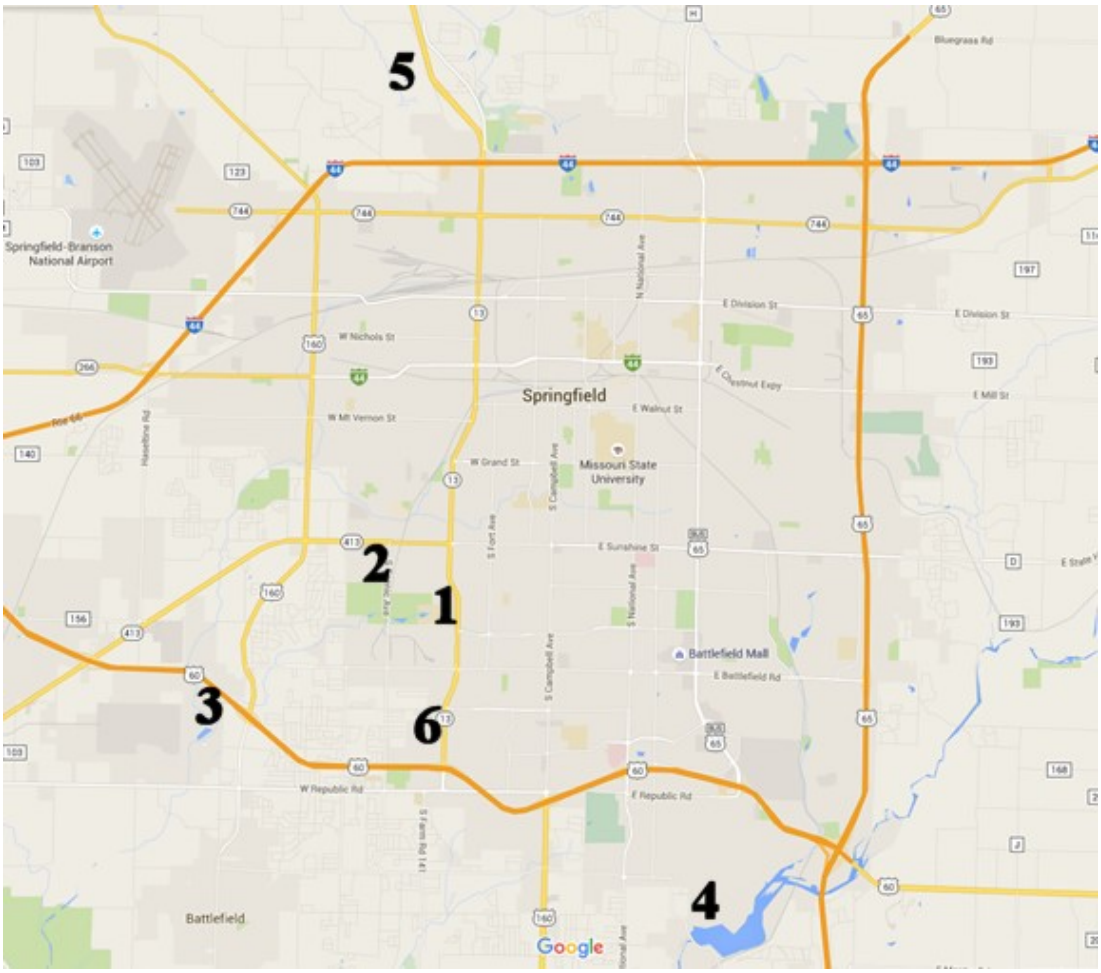

Figure 3. Six locations of local Vitaceae survey. 1; Sunset and Kansas, 2; Close Memorial Park, 3; Confluence of Wilson and South Creeks, 4; Lake Springfield Park, 5; Ritter Springs Park, 6; Dr. Qiu's yard.

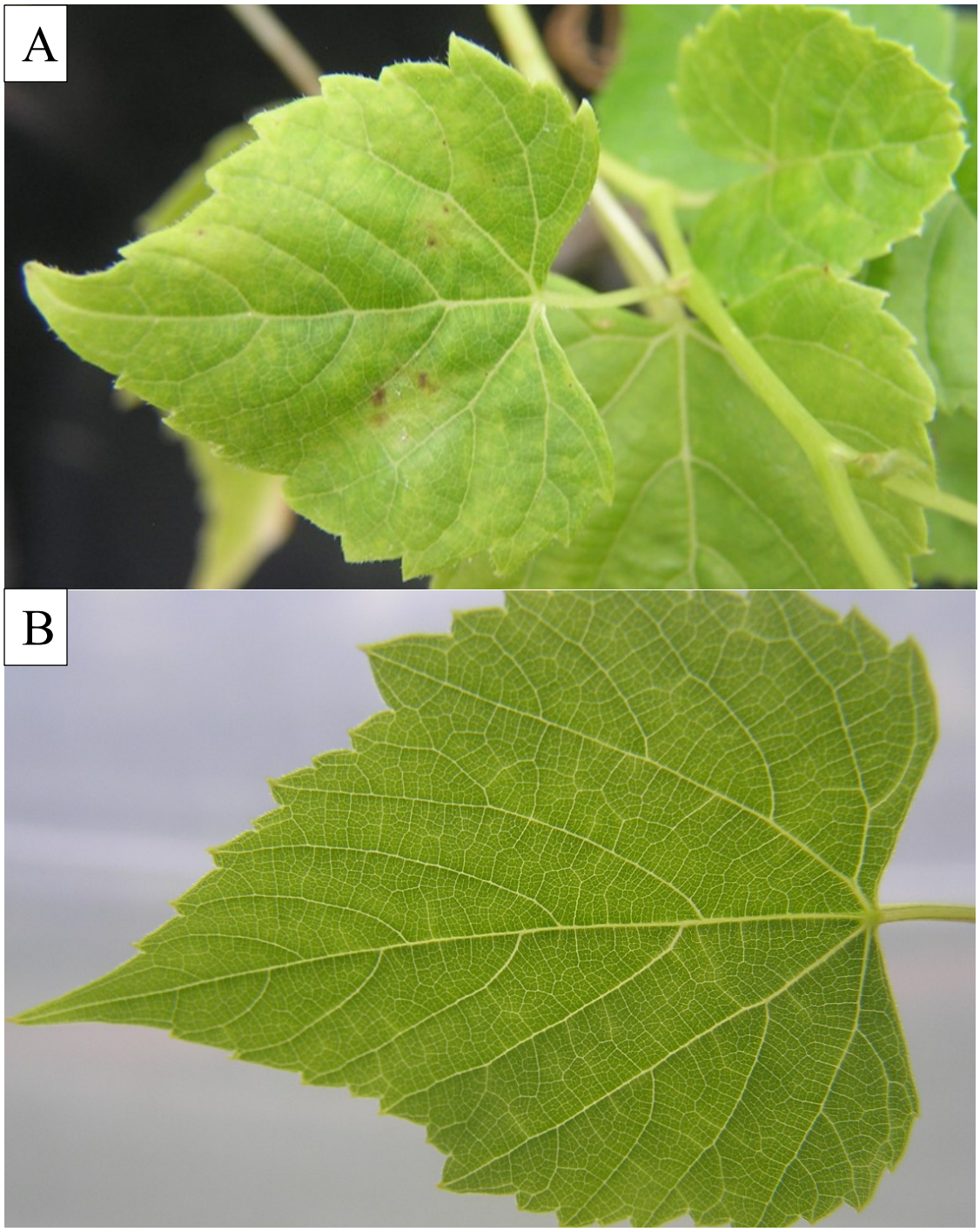

Figure 4. GVCV associated symptoms in two *A. cordata* accessions. (A) AMP1Pexhibiting mild vein clearing and chlorotic spots associated with GVCV. (B) AMP2P exhibiting moderate GVCV associated symptoms. Chlorotic and necrotic spots, along with slight leaf curling are present.

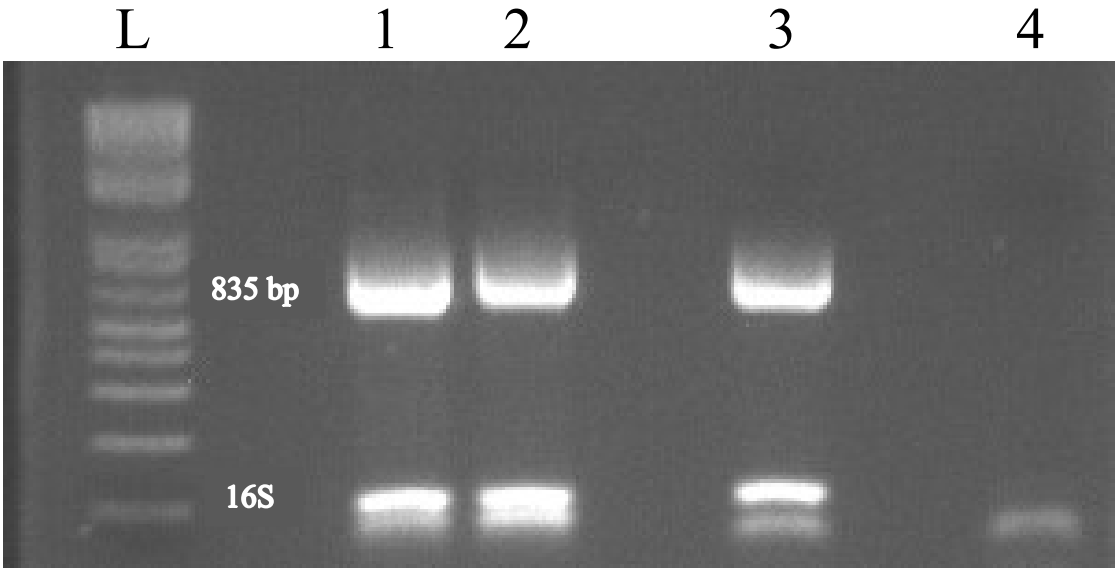

Figure 5. Initial GVCV assay results of *A. cordata*. Lanes 1; AMP1, 2; AMP2, 3; positive control, 4; negative control. The size is indicated to the right of the ladder.

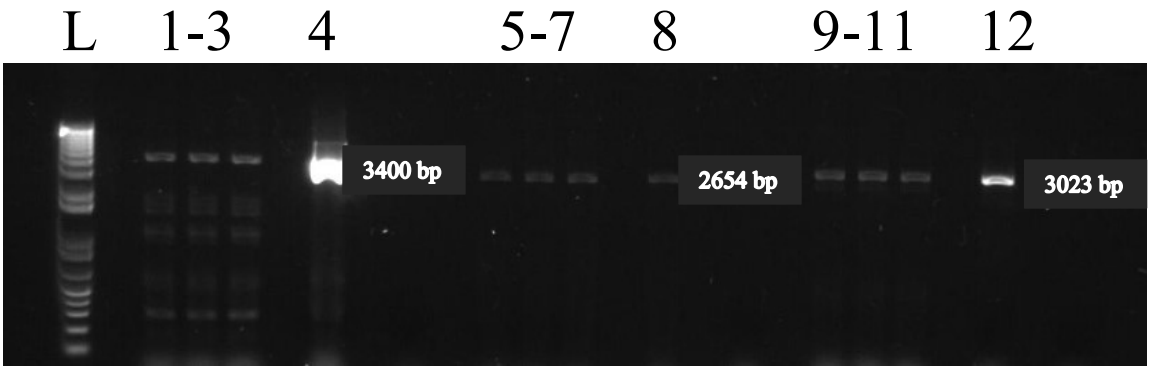

Figure 6. AMP1 DNA separated on an agarose gel. Lanes 1-3; Fragment 1, 4; positive control, 5-7; Fragment 2, 8; positive control, 9-11; Fragment 3, 12; positive control. The size is indicated to the right of each positive control.

```
CHA_OFRII       ATGTCCACGTGGCAAATTGCTGCTGCCGCAGAGGAATACAAAAACGCCATAAAGGCAACT
VRU1            ATGACCACGTGGCAAATCGCTGCTGCCACAGAAGAATATCAGAAAGCCATAAACGCAACT
AMP2ORFII       ATGTCCACGTGGCAAATCGCTGCTGCCACAGAAGAATATCAGAAAGCCATAAACGCAACT
AMP1ORFII       ATGTCCACGTGGCAAATTGCTGCTGCCACAGAAGAATACAAGAAAGCCATAGAAGCAACT
VRU2ORFII       ATGTCCACGTGGCAAATTGCTGCTGCCACAGAAGAATACAAAACCGCCATAAAAGCAACT
                *** ************* ********* **** *****  * *  ****** * ******

CHA_OFRII       GCAACCCTCACCAAGGACGAAAGAGCTGTTGGCTTTGTTAAGCCCCACGAGTTCGAGCCA
VRU1            GCAACCCTCACCAAGGACGAACGAGCAGTTGGCTTCGTTAAGCCCCACGAGTTCGAACCA
AMP2ORFII       GCAACCCTCACCAAGGACGAAAGAGCAGTTGGCTTCGTTAAGCCCCACGAGTTCGAACCA
AMP1ORFII       AAAACCCTCACTAAAGACGAAAGAGCAGTTGGCTTCGTCAAGCCCCACGAGTTCGAACCA
VRU2ORFII       GCAACCCTTACCAAGGACGAAAGAGCAGTTGGCTTTGTTAAGCCTCACGAGTTCGAGCCA
                  ****** ** ** ****** **** ******** ** ***** *********** ***

CHA_OFRII       AATTTCAGTGACACCAACATTCAAAGGCAAAACAATACTTTGATCCACTTGCTAATTCAA
VRU1            AATTACAGTGACACCAACATTCAGAGGCAAAATAATACTCTGATCCACCTGTTGATCCAG
AMP2ORFII       AATTACAGTGACACCAACATTCAGAGGCAAAACAATACTCTGATCCACCTGTTGATCCAG
AMP1ORFII       AACTACAGCGACACCAACATCCAAAGGCAAAACAATACCTTGATCCACCTATTGATCCAA
VRU2ORFII       AACTTCAGTGACACCAACATCCAAAGGCAAAACAATACCCTGATCCACCTGTTGATCCAG
                ** * *** *********** ** ******** *****  ******** *  * ** **

CHA_OFRII       AGCCTTGAAGAAATCAAAGAGCTCCGCGCTCAGGTTCAGACCCTCAACGATCGTATTATA
VRU1            AATCTTGAGGAAATCAAAGAGCTTCGTTCTCAAGCCCAAACCCTCAACGATCGTATTGCA
AMP2ORFII       AATCTTGAGGAAATCAAAGAGCTCCGTGCTCAGGTCCAGACCCTCAACGATCGTATTGTG
AMP1ORFII       AATCTTGAGGAAATCAAAGAGCTCCGTGCTCAGGTCCAGACCCTCAACGATCGAATTGTG
VRU2ORFII       AATCTTGAGGAAATCAAAGAACTCCGTGCTCAAGTCCAGACCCTCAACGATCGTATTGTG
                *  ***** *********** ** **  **** *  ** ************** ***

CHA_OFRII       ACTTTGGAAAAGGGAAAGGCTC---------CAGTCACTCTTCCTGATAACGTGGTAGAA
VRU1            GCCCTAGAAAAGGGAAAGTCAAAAGCAACAGCAGTCACCCTTCCTGATAACGTGGTAGAA
AMP2ORFII       GCCCTAGAAAAGGGAAAAGCTAAAGCAACAGCCGTCACTCTTCCTGATAACGTGGTAGAA
AMP1ORFII       GCCTTAGAAAAAGGGAAGTCCAAAGCGACCCCTGTCACTCTTCCTGACAACGTGGTAGAA
VRU2ORFII       ACCTTGGAAAAGGGGAAGTCAG---------CCGTCACTCTTCCTGACAACGTGGTAGAA
                 *  * ***** ** **  *           * ***** ******** ************

CHA_OFRII       CAAATATCCACTCAACTCAAGGAAGCCAAGTTTGGACAACCAAAAGAAGGTTTGGTCAAA
VRU1            CAAATCTCCACTCAATTAAAGGAAGCAAAATTCGGAACTCAGAAAGAAGGCTTGGTAAAA
AMP2ORFII       CAGATCTCCACTCAACTCAAGGAAGCCAAGTTCGGACAACCAAAGGAAGGTTTGGTAAAA
AMP1ORFII       CAGATCTCCACTCAATTAAAGGAAGCAAAGTTCGGAACTCAGAAAGAAGGCTCAGTAAAA
VRU2ORFII       CAAATCTCCACACAACTAAAGGAAGCAAAGTTTGGAACTCAGAAAGAAGGCTTGGTAAAA
                ** ** ***** *** * ******** ** ** ***   *  ** ***** *  ** ***

CHA_OFRII       GGGACAAAAGGCACCTTCCGGGTCTGGAAGTGA
VRU1            GGGACAAAAGGCACCTTCCGTGTCTGGAAGTG-
AMP2ORFII       GGGACAAAAGGCACCTTCCGGGTCTGGAAGTGA
AMP1ORFII       GGGACAAAAGGCACTTTCCGGGTCTGGAAGTGA
VRU2ORFII       GGGACAAAAGGCACCTTCCGTGTCTGGAAGTG-
                ************** ***** ************
```

Figure 7. CLUSTAL multiple sequence alignment by MUSCLE showing all five GVCV isolate's ORFII. The (*) denotes all sequences are identical at this position.

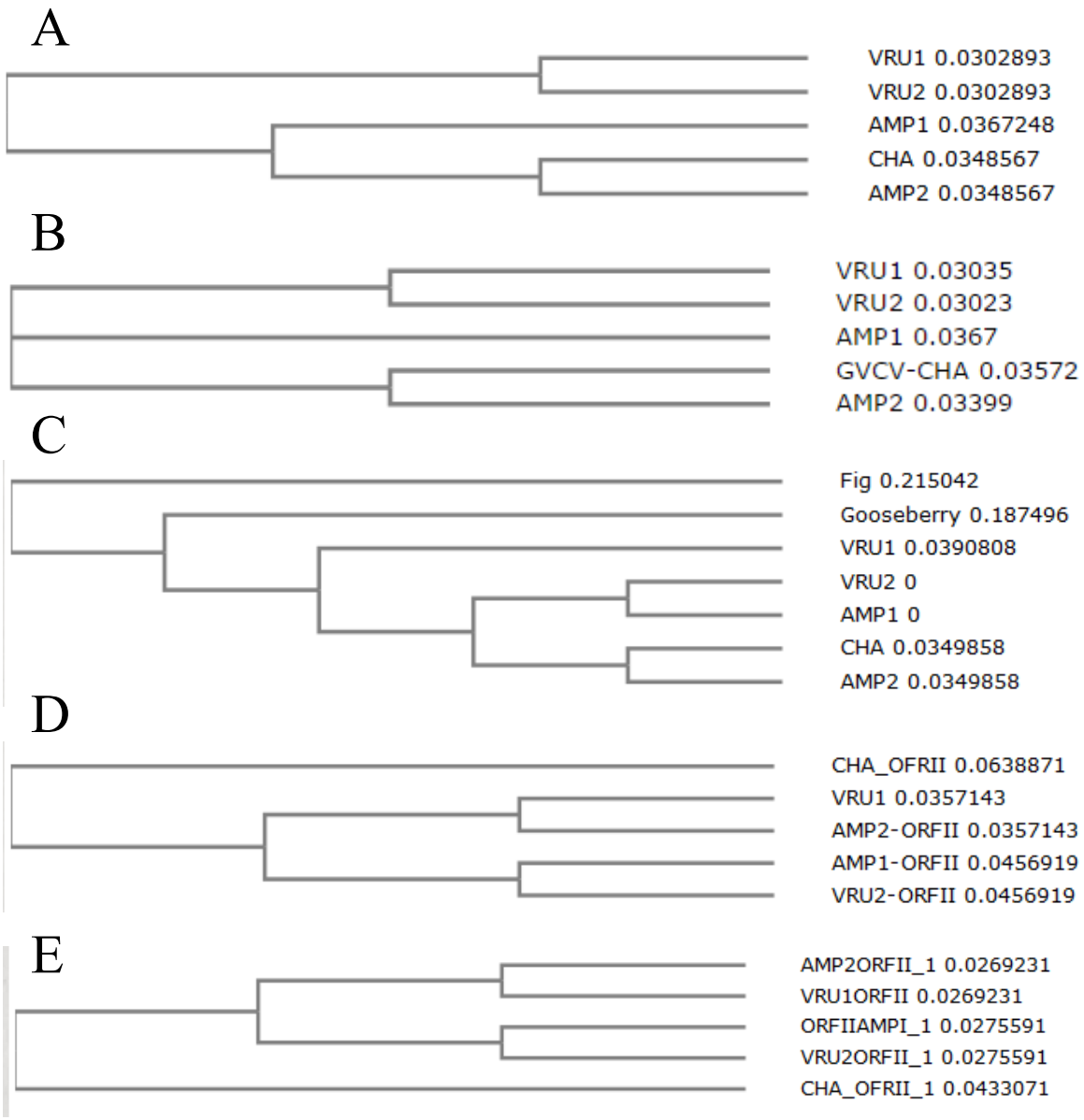

Figure 8. Phylogenetic trees of five GVCV isolates. (A) UPGMA rooted tree built with ClustalW2 from the five available GVCV nucleotide sequences. (B) Neighbor joining tree built with ClustalW2 from the five available GVCV nucleotide sequences. (C) UPGMA rooted tree built with ClustalW2 showing the five GVCV isolates in relation to two other badnaviruses. (D) UPGMA rooted tree built on the nucleotide sequences of ORFII in all five GVCV genomes. (E) UPGMA rooted tree built on the amino acid sequences of ORFII in all five GVCV genomes.

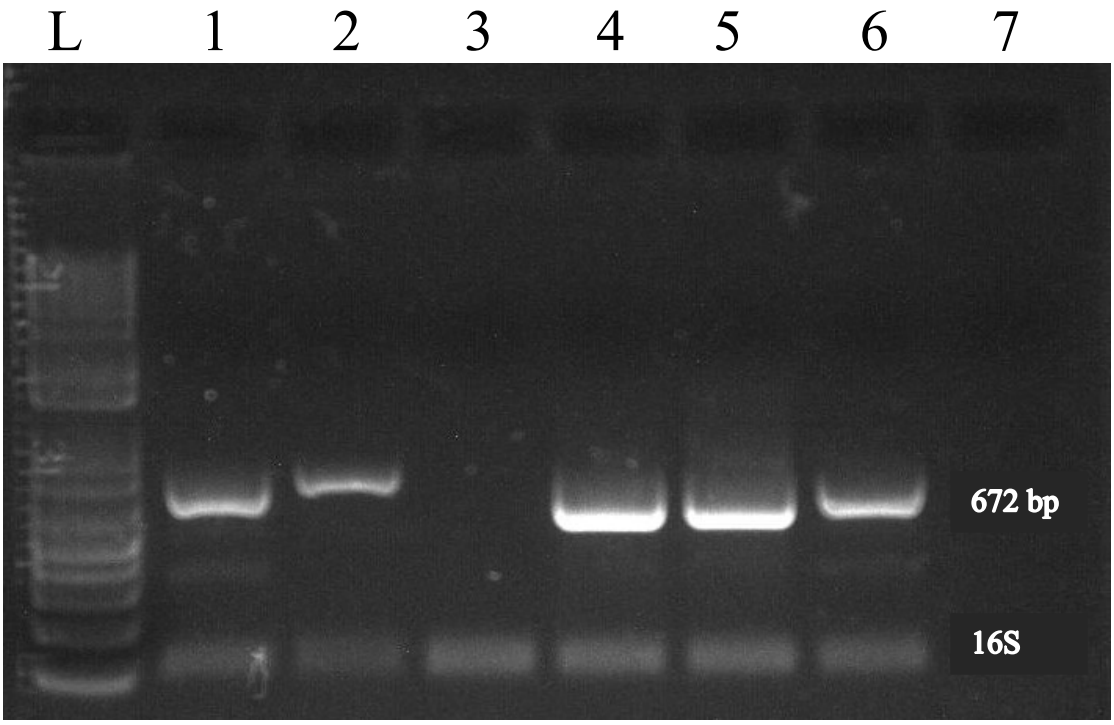

Figure 9. Agarose gel image of DNA fragments acquired in PCR using a primer set to isolate all of ORF II in five known GVCV isolates. Lane 1; CHA, 2; VRU1, 3; VRU2, 4; AMP1, 5; AMP2, 6; positive control, 7; negative control. The size is indicated to the right of positive control.

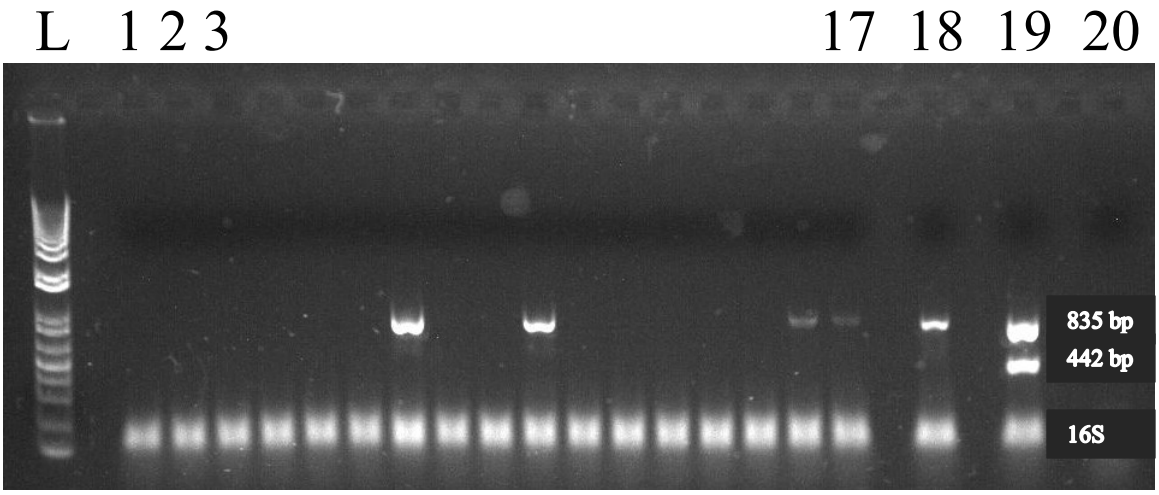

Figure 10. Agarose gel image of a portion of local Vitaceae survey results on 2 *Vitis* and 15 *Ampelopsis* samples. Lanes 1 and 2; *Vitis*, 3-17; *Ampelopsis*, 18; AMP1 positive control, 19; CHA positive control, 20; negative control. The size is indicated to the right of GVCV-CHA positive control.

Figure 11. Present known range of GVCV in cultivated grapes.

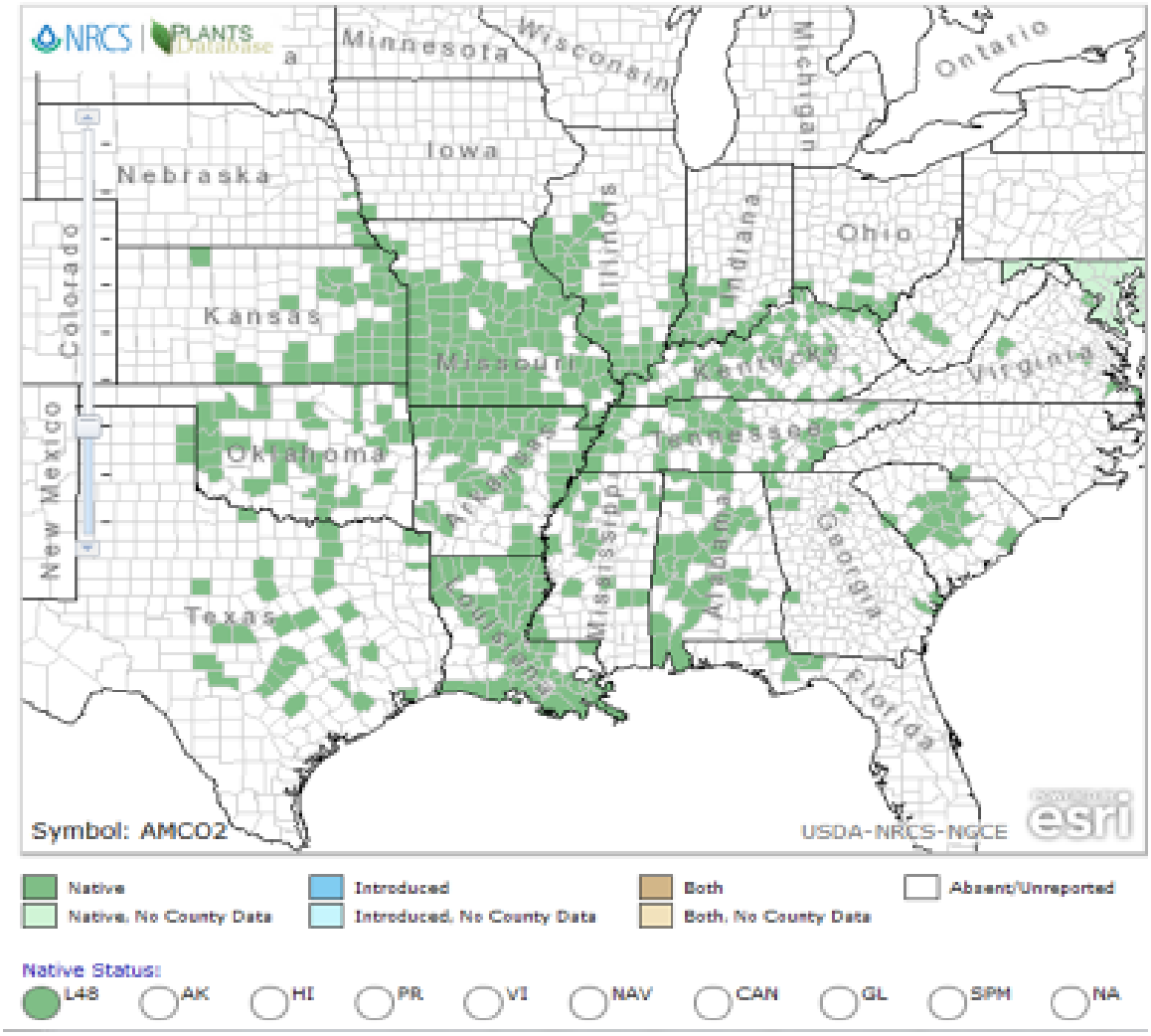

Figure 12. USDA, NRCS map of native range of *A. cordata* in the United States.

Figure 13. *Ampelopsis* and *Vitis* sharing habitat.

APPENDICES

Appendix A. Nucleotide Sequence of GVCV-AMP1 Genome in FASTA Format

>AMP1
TGGTATCAGAGCAAGTTTCAAATCTGGGAATTTCTACAATTATTCCTTCAAGA
TTATGATGAGGAACTAACTCTCATAATCGTGTAGGGAATCGTTAGTAGGATCT
CAGAACAAGGTTCTTATCCCCTCAGACTACTGATTCTGGTATATAGGCTGGAA
ACACGACACTGTTACGATCCCACTTCTGTTGGAGTGGTAGTAGCCCGTTGTGA
GACAACGCCACGTACCATTTTCAGTCCTCCTAGCCCAAATCCCCATGAACAG
AACTCCCACAGTCAATAAGCTTCAACAGGATCCCTAGCCCAACAATACTGAA
AGTCCTAGGACAGGCTGCGACGCGAAGTATCACTAGTTCAGGCGATGCTGTT
CCGCCGACTATTTGTGAGAAAACTGCAGTAGGAGAGGACGGACAACTATTCA
AGGGAACGGAACCTGGAGACACTGGCCGAGTTCTTAGTAAGCGGTTCAAGGA
AGGAGACTGATGCAAAATATAGAACAACAACAGTTTGAGGCGGAGATAGAA
TCTTGGGAGAGATCTGAACGCACACCCCTACACGGTTACCGTGATCTTGTGG
AATACCCCCGTTACGAAAGGAACCAGCACTTCCCATCTGCAAAGTTCCCCTG
TTACCACTTTGTTGCTGAGAAAGACAACGTTCACGCCACTTACACTAAGGGA
GATAGAATCCCTCAGTTGCTAAATACACTGTACGACCTACAGGTCAACCAGT
GTCACAACCAGGCAGTGATCTACGATCGGATCCAACTCCTTTCGAGGTATAC
GGTCCGAAAGGATAAGCCTTTACCGGCTATCCCTGAGGAATCTGTCCTCAAA
GAGCCAGAAGAAAGCTCAACTGAGCTTAAGCACCAGATCGAGCTCCTTCGAG
CAGATCTAAGGGAAATCAAGGCCAACCAGTCAGGTCTTCGCCTTGCCATCTC
TGAGATCCGTGACTCCATCACAGATCTAACGGCAAGAGAATCGGCACCCAAG
CCGATTGAAGCAGAGACAGCCTACCTGACCGCCCAGCTGAAGGTCCAGGTTC
AGGAGATCAAAACAGCTTTAGCAGAGATTAAAACCTTTGCCAGATCCCTGGT
CCCTGAAAGGTAGATGTCCACGTGGCAAATTGCTGCTGCCACAGAAGAATAC
AAGAAAGCCATAGAAGCAACTAAAACCCTCACTAAAGACGAAAGAGCAGTT
GGCTTCGTCAAGCCCCACGAGTTCGAACCAAACTACAGCGACACCAACATCC
AAAGGCAAAACAATACCTTGATCCACCTATTGATCCAAAATCTTGAGGAAAT
CAAAGAGCTCCGTGCTCAGGTCCAGACCCTCAACGATCGAATTGTGGCCTTA
GAAAAAGGGAAGTCCAAAGCGACCCCTGTCACTCTTCCTGACAACGTGGTAG
AACAGATCTCCACTCAATTAAAGGAAGCAAAGTTCGGAACTCAGAAAGAAG
GCTCAGTAAAAGGGACAAAAGGCACTTTCCGGGTCTGGAAGTGATGTCTCGG
TCCAGAACTCAGACCACTGAGTTGCCTCGCGCAACCAGGAGATCCACTAGCC
CAGTTGAGAGGCTAGATGATCAGATCCGCGGTTACCGGCGGATGGCTCGTGC
CCGCTACCTTGCGGAGCAACGGATACGTAGGTCCTTCTCAAGGAACTACAGG
GAAACTCTGGAAAGACGCCTAGATCCAGAGGCAGAATTGCAGCTCAGTCGAA
GACGAAGAGCTAACCTAGTACCAGCAGAAGTACTATACTCCCTCAACTACAA
TGAACCCCAGAATAGGGTTTATCAACACTATGAAGAGGTGAGATCCCACGTC
ATAGACCGGCAGCAAGATTTCCGGTTTATTGAAGAACAGTCCTACCGCACCC
TGGTGCAAGAAGGCATGCAGCATATCCACCCTGGAATGCTGATGGTAAGAAT
ACAGATGCTACACCGAGTAGACGCAGGAATCAGTGCCATGATCGTGTTCCGA
GACACGAGGTGGAATGATGAAAGGCAAATCATCAGTGCTATGACTGTAGATA

TGGCCAGAGGCGCACAGCTGGTCTACGCAATCCCAGATCTCATGATGTCAAT
TCATGATTTCTACCACCACCTGCAAGTCAGCATCACCACTAGAGGATACGGT
ACCGGTTGGGCAGGAGGAGAAAGTAACCTCATAGTAACTCGGTCACTCACCG
GAAGAATCACCAATACTAGTCAAGCCAACTTCAATTATCAGATTGAAGGAGT
AGCTGACTACCTGGCGAGCCATGGCGTGCAGAGTATACCAGGACAGCCATGG
AGAGACATAAACCAGGAAGGTTCCTGGAACCTAAGGCCTTCGTCAATACAGG
CCCCTACTCAGGTCCCTACAGGTCTTGTCTCAAGACAGACTGCCACCGGCAAT
ATAAGCCTAAGGTTCACTAGTTTTCAGGACCAGGTCCAGACTGTTGATACAGa
AGAAGAAGCAGGGACTACTGACACAGAAAGGGTAACTCACTATGCCCTTGTT
GGAACCTTCGAATGGTTGGAGGAATGTCCTTCATATCAACAAAGAAGGAGTC
AAGAAACGGAGGAGAATGGCTGGGTAAACCATGTGGAAGGAGATAAGGGGT
TCAACTTCAAAGTCCGCATGACCCCTCCAGCATGGAGCCATAATCCGCAGCC
TATTACGGCTACAGGATGGGGAGATGATTTTGATGACTCTCCACCACCTCCTA
AACCTCCTAAGACTGAAGAGGAGGAAATACTGGAATTATACCCAGTAAGAA
GACAACCCGATCCAGTACAGATAGCAAGGAAAGAAAAGGCAGTAGTTTTCTC
TCAAGCTGTCAACGATATCTTTGAGCAAGAAGGGAAGGGTGTCTCTAGAATG
CAACCATCAGGGGAAGCCCCTGACTCAGATCCAGATAGCCCAGTCTGGAAGA
TAAAGAAAAGCCCATATCCGCAACAACCTCTAAAGCTGAAGGATGAGAAGG
GTAAAAGTCCTTTTGAGGACTTAGAGCTGAAACAAGACCTAGTTCAAAGCTG
GATAGCTCAACTAGGAAGTGGCTCAGGAAGCAGAACGGAGAAACCAGTCTT
CGATACCACTAGCAGCGACTCAGATTCTGATTTATCTGATGTCAGCTCAAAAG
TACTAGCCTACGCTGGAGTTGAAGAAGCGGTAATGGAATACCCAAGAAGGGT
AAAAACTGCGACGGCTAAGCTAGCAGACATGGAAAAGGCTTTTGCCGGAGA
AACAACTGCAGCAGTGGGAGGAGACTCGGAGATGACGACTGGTCAGTCTTCC
AGATCTACCCTCATACCACCAAATGAAGGAGGAGGACCTATACGGTACCCAC
CAGCAGAAAGACCGTCCACATCGGCCTCTACGTATAACGCCACAGCCCCACC
TCTTTTTGAAGGGACTGTTAGGCCTGGAAGATATGGTCGCCCCCTGGCACCAT
GGTCTCTGCCATCAGCACAGCACTCCCAAGGAGCTTTGCTGATCCTCCCTCCT
GAAGTAGCAAGTCACGCTGACGCCATCACTACATGGGAAACAATCACCCTAA
ATCATTTGATGAATATATCATTTGATTCCCTCCAAGATAGGGTTGATTACATT
GAAAACCTCCTTGGACCAAGAGAAAGAGAAGCTTGGGTCACTTGGAGAATG
GCGTATGATACGGAGTACAGACAGCTGGTCGAGCTCTCTGGAGAACCAAGAA
ATGTGACCAGCACCATTAAAAGAGTTCTAGGGATTAATGACCCCTATACAGG
AACTACTCATATCCAGAATCAGGCTTATGCTGATCTTGAGCGCCTGCAGTGCA
AAAATCTGGAATCAGTAATGCCGTTCCTAAACTCTTATTTCCAACTCGCTGCA
AAGAGTGGAAAGATGTGGAGTAGCCCTGAACTCTCAGAAAAGCTTTTTAGAA
AGCTCCCACCAGAAATCGGTCCCACCATAGCAAAGGAGTATGCTGAGCGATA
CCCTGGCATGTTGATTGGAGTAAATGCCAGAATACAGTTCGTCTCTGAGTATC
TCCAGGACCTCTGTAAACAAGCAGATCTTCAAAGAAAATTGAAGAATTTGAA
TTTCTGCAAGGCAATTCCCATTCCTGGTTACTATGACCAAGGAGTGAAGAAG
AAGTACGGCCTGCGCAAATCCAAGACATACAAAGGGAAACCTCATGACTCTC
ATGTCAAGGTTATCAAAAACAAGTACAAAGGTGCCCAAGGTCGTAAGTGCAA
ATGCTACTTATGTGGCATTGAAGGACATTACGCTCGCGAATGCCCAAAGAAG
CATGTTAGGCCTGAAAGAGCAGCATACTTCGAAGGCATGGGCTTGGATGTCA
ACTGGGATGTAATAAGTGTTGACCCAGGAGATCAAGATGGATCTGACATCTG
CTCAATCTCTGAAGGAGAAGCCCAACATGGGATGGAGGAATTAGCTGCATTC

AAAGCCCAACTTCCATATCCAGTGGAAGCCCAATATGAGCAGCACCAGGCCT
TTGTGGTTATCCAAACAACTTTCAAGAAGGAGGATAAGCCCCAAGGTTCCTG
GCGCATGTCCAAGCCCATCCCCGAAGCCCAACAACAATGCCAGCACACATGG
GATGATATGTACGCCCTAGCAGAAGGACAGCAAGCGTGCAGCACTTGCCAGA
CCATCACTGTACTTGGTCGCCGTGCCACATGCACCCTTTGCCTACTCAACCTC
TGTTCACTCTGCGCTGGCCTAGACTTCGGTCTCAAAATAGTTCCTAAAACTGC
CACACGTGCTGACTGGAAATTTCAGGATCGTGATACCCTCATCGCCTCCTTAT
ATGAGCACAATGCATTCCTTCTTCGCCAAGTCGAAGGGCTGAAACAGGAACT
CCAAGCTGCCAAAGAACAGCTTCAACTGCTACACTCGGTTGATATGATCAAC
CTCTCTGATGATGGATTAGAGAATTTTTCCGTTGAGGAAAAATCCTTTTTAAG
AGGGGGAGGGGGTACCAGTAGTAGTTCAATCAAAATTTCATCAACCACAACA
CCCCCTGGTTTTCCTACAACACCCAATAGATTCCAGCCTCTTGCGCAGGAAAA
ACTTAAAGGAATACAGGAAGACCTATCTCTAGTGGTACAGTTTGATAATGCT
AGACAACAAGAACAGGCGTATACTGAAATGCCTCGAGGAGCCCACAACAAG
TTATACCACGTGGTGGTAACTTTCAGAATCCCTGATGTTAAGGGACAGCTCCT
AGAGTTTGATATCAACGCCATTATAGACACTGGCTGTACCTGTTGCTGCATCA
ACCTCACAAAGGTGCCTGATGGAGCAATCGAAAACGCCTCCATAATCCAAGA
AGTCTCTGGGATTAACAGCAAAACAGTAGTCACCAAGAAACTCAGGCAAGGC
AAGATGATCCTCGCAGGGAATGATTTCTACATTCCTTATGTATCAGCCTTCGA
GATGAACATGCCCGGGATTGACATGCTGATAGGCTGCAACTTCATTAGAGCA
ATGAAGGGAGGAATACGGTTGGAAGGAACTGAGGTCACCTTCTACAAAACC
ATCACCAGGATTCAAACTACCCTGGAACCTCAGAAGATAGCGTACTTGGAAG
AGCTAGTAGAAGCAGAAGATCTACACTATGAGCTCGCAGCTGCAAGTATGCC
TGAGCCCACTGCTGAAGGACTCAGAAACACCAAGCTCCTAGCCGAGTTAAAA
GAACAAGGCTACATAGGAGAAGAGCCTCTCAAACACTGGTCAAAGAATAGG
GTACGATGCAAGCTTGACATCATCAACCCTGACATCACCATTGAGGCAAAGC
CACCTGGTCACCTAACACTGGAGGACAAGGTCAAATATCAGAAGCACATTGA
CGCCCTCCTAGATCTTGGAGTCATAAGACCCAGCAAGAGCAGACACAGGTCC
GCAGCTTTCATAGTTGCCTCTGGGACCTCTGTAGATCCTAAAACTGGCAAAGA
AACACGCGGTAAAGAAAGAATGGTGATCGACTACCGCATGCTTAATGACAAC
TGCTACAAGGATCAATACAGTCTGCCTGGTATCACCTCCATCATCAAATCCCT
TGGGCAAGCCAAAATATTCAGCAAATTTGACCTGAAGTCTGGCTTCCACCAA
GTCATGATGGAAGAAGAAAGCATCCCCTGGACTGCTTTTATCAGCCCCGCAG
GATTATATGAATGGCTAGTTATGCCATTTGGGATCCAAAATGCTCCTGCAATA
TTCCAAAGAAAGATGGATGAGTGCTTCAAAGGAACTGAGGATTTCATCGCTG
TTTATATTGATGATATTCTAGTTTTCTCCAACTCCATCAAAGAGCATGAAAAG
CACCTACAGAGAATGCTGAGTATCTGCAAGGAACATGGGCTCGTCCTTAGCC
CAACAAAAATGAAGATCGCTGTCCCAGGAATTGATTTCCTTGGTGCCCATATC
AGAAACAGCAGAGTGAGTCTGCAACCACACATCATCAAGAAGATTGCTGACA
AGAAAGATGATGAGCTGATGACCCTAAAAGGCCTCAGAAGCTGGCTAGGGG
TAATCAACTATGTCAGGCAATACATTCCTAAGTGCGGAACCCTTCTCGGTCCC
CTCTATGCTAAAACTTCAGAGCATGGTGATCGAAGATGACACCCCAAAGACT
GGGAAATAGTAAGACAGATCAAGAAGATGGTTCAATCCCTTCCTGATCTTGA
ACTTCCTCCACCCCACGCAGTTATTATCATCGAATCTGATGGATGCATGGAAG
GATGGGGAGGAATCTGCAAATGGAAGAACTCAAAGGGGGAATCTAAAGGCA
AAGAGCGAATCTGTGCTTATGCCAGTGGAAAATTCCCAACAGTTAAATCCAC

CATAGATGCTGAAATCTATGCAGTCATGGCATCCCTGGAGAACTTCAAGATTT
ACTATCTTGATAAAAGGGAAATTACCATCAGAACAGACTGCCAAGCCATAAT
CAGTTTCTATGATAAAATGGCTGTTAAAAAACCTAGCAGAGTCCGCTGGATA
AATTTCTGTGACTACATCACTAACACAGGAATCAAAGTCCAGTTCGAGCACA
TAAAGGGTCAAGATAACCAGCTAGCAGACCAGCTCTCAAGGCTAGCCCAAG
GACTTTGCAGCATTCAAGTCATCCCTGAAGCAGCCCACGAAGCTCTCAATAT
CATCCTTGAACAGGATTGTACAGCCCAAGAGCTCATGGCCCAATTCAACTCT
ATGCTGCAAGCTAACCTCAGGGTTAACCAAGGAAGGCCCAATACAACTTGGT
ATTCTAGGACCAAGCCCAAGAAATCCAAAGCCCGCAAACCAGCCCATGTCCA
GCCATGCTTTGACGTAAGCAATGACGACGCGGGATAATAATGGAGGAATCTT
ACAAGGACAGCACATGGTCCTTTTCCTCTTTATTTTAGTTTGTTTTCTTGTGTC
GGCAACCTCTCCTTTTGTAAAGAGGAATCTGCTTTTGAGCTGTCGATGGGGCC
CAATGAGCACCCGAGCTCTAAAAGTAACTTACCTCTGGTTGCTTTTGTTAACC
TTAGTTTGGTTTGTTTGCTTTTCTCCCCTATATAAGGGAGCTTCTCATTTGTTA
GAAGGCATCGAACAGAGTAATACCTCTGAGCGCTCCTTCTCTCTAGTTTTCTT
ATGTTCTTGTATCTTTCCAGTTCCAGTGTTCTTAATGCAATTTGAAGTTTTCCT
ACTCTATGTTATTCTGTTCATAGTTCTTTTCCGCTACTTATACTCTGTGATCCA
AATTTTTAATTTGTGATCTGTTTCATC

Appendix B. Nucleotide Sequence of GVCV-AMP2 Genome in FASTA Format

>AMP2
TGGTATCAGAGCTCCAGTTTCAAATCTGGGAATTTCTGCAATTATTCCTTCAA
GATTATGATGAGGAACTAACTCTCATAATCGTGTAGGGAATCGTTAGTAGGA
TCTCAGAACAAGGTTCTTATCTCCAACTCCACTACTGATTTTTGGTATATAGG
CTGGAAACACGACACTGTTACGATCCCATTTCTGTTGGAATGGTAGTAGCCCG
TTGTGAGACAACGCCACGTACCATTTTCAGTCTTCCTAACCCAAATCCCCATG
AACAGAACTCCTACGGTCAATAAGCTTCAACAGGATCCCTAGCCCAACAATA
CTGAAAGTCCTAGGACAGGCTGCGACGCGAAGTACCACCAGTTCAGGCGATG
CTGTTCCGCCGATTGTTTGTGGGAAAACTGCAGTAGGAGAGGACGGACAACT
ATTCAAGGGAACGGAACCTAGAGACACCGGCCGAGTTCTTAGTAAGCGGTTC
AAGGAAGGAGACTGATGCAAAGAATCACAGAACAACAACAGTTTGAGGCGG
AGATAGAATCTTGGGAGAGATCTGAACGCACACCCCTACACGGTTACCGTGA
TCTTGTGGAATACCCCCGTTACGAAAGAAATCAGCATTTCCCATCAGCAAAG
TTCCCCTGCTACCACTTTGTTGCTGAGAAAGATAACGTTCACGCCACCTATAC
TAAGGGAGACCGAATCCCTCAGTTGCTGAATACACTGTACGACCTACAGGTC
AACCAGTGTCATAACCAGGCAGTGATCTACGAACGGATCCAACTCCTTGCGA
GATATACGGTCCGAAAGGGTACGCCGTTACCGGCTATCCCTGAGGAATCTGT
CCTCAAAGAACCAGAAGAAAGCTCAACTGAGCTTAAACACCAGATCGAGCTC
CTTCGAGCTGATCTAAGGGAAATTAAGGCCAATCAGACAGGCCTCAAGCTTG
CCATCTCTGAGATCCGTGATTCCATCACAGATCTAACGGCAAGAGAATCAGC
ACCCAAGCCGATTGAAGCAGAAACAGCCTACCTGACCGCCCAGCTAAAGGTT
CAGGTTCAAGAAATCAAAACGGCTTTGGCAGAGATTAAGACCTTTGCCAGGA
CTCTTGTTCCTGAAAGGTAGATGTCCACGTGGCAAATCGCTGCTGCCACAGA
AGAATATCAGAAAGCCATAAACGCAACTGCAACCCTCACCAAGGACGAAAG
AGCAGTTGGCTTCGTTAAGCCCCACGAGTTCGAACCAAATTACAGTGACACC
AACATTCAGAGGCAAAACAATACTCTGATCCACCTGTTGATCCAGAATCTTG
AGGAAATCAAAGAGCTCCGTGCTCAGGTCCAGACCCTCAACGATCGTATTGT
GGCCCTAGAAAAGGGAAAAGCTAAAGCAACAGCCGTCACTCTTCCTGATAAC
GTGGTAGAACAGATCTCCACTCAACTCAAGGAAGCCAAGTTCGGACAACCAA
AGGAAGGTTTGGTAAAAGGGACAAAAGGCACCTTCCGGGTCTGGAAGTGAT
GTCTAGGTCCAGGACTCAGACCACTGAGTTGCCTCGTGCAACCAGGAGATCT
ACTAGCCCAGTTGAAAGGCTAGATGACCAGATCCGCGGCTACAGGCGGATGG
CTCGTGCCCGCTACCTTGCGGAGCAGCAAATGCGTAGGACCTTTTCAAGGAA
CTACAGAGAAACCCTGGAAAGACGCCTTGATCCAGACGCAGAATTACAGCTA
AGCAGAAGGCGAAGGGCCAACCTAGTACCAGCAGAGGTACTATACTCCCTCA
ACTACAATGAACCCCAGAATAGGGTTTATCAGCACTATGAAGAGGTGAGATC
CCATGTCATAGACCGGCAGCAAGATTACCGGTTTATCGAAGAACAGTCCTAC
CGCACCCTAGTGCAAGAAGGCATGCAGCATATCCACCCAGGAATGCTGATGG
TAAGAATACAGATGCTGCACCGAGTTGATGCAGGGATCAGTGCCATGATTGT
GTTCCGAGACACAAGGTGGAATGATGAAAGGCAAATTATCAGTGCCATGACT
GTTGATATGGCTAGAGGAGCACAACTGGTCTATGCTATCCCAGATCTCATGAT
GTCAATTCATGATTTCTATCACCACCTACAAGTTAGCATCACCACCAGAGGAT
ACGGTACCGGATGGGAAGGAGGTGAAAGTAACCTCATAATAACTCGGTCACT
AACCGGCAGAATCACCAACACCAGTCAGGCCAACTTCAATTATCAGATTGAA

GGAGTAGCTGACTACCTGGCAAGCCATGGCGTGCAGAGTATACCAGGACAGC
CATGGAGAACCATAAACCAGGAAGGTTCTTGGAACTTAAGGCCTTCATCAAT
ACAGGCCCCTACTCAGGTCCCCACAGGCCTTATCTCAAGACAATCTGCCACC
GGTAATATCAGTCTTCGATTCACTGGTTTTCAGGACCAGGTCCAGACCATTGA
CACAGAAGAAGAATCCGGTATGACAGATACAGAGGAAAGGGTAACTCACTA
TGCCCTTGTTGGAACCTTCGAATGGTTGGAGGAATGTCCTTCATATCAACACA
GAAGGAGTCAAGAAACAGAGGAAAATGGCTGGGTGAACCATGTGGAAGGAG
ATAAGGGGTTCAACTTCAAAGTCCGTATGACCCCTCCAGCATGGAGCCATGA
TCCACAACCCATTATAGCCACGGGATGGGGAGATGATCTTGATAATCCTCCA
CCACCTCCACCTCCTAAGATTGAAGAGGAGGAGATACTGGAATTATACCCAG
TAAGAAGACGACCCGACCCTTTACAGATAGCCAGGAAGGAAAAGGCAGCAG
TCTTCTCTCAAGCTGTCAATACTATCTTCGAGCACGAAAGGGAAGGTACCTCA
AGGATGCAACCATCAGGGGAAGCCCCTGAATCAGATCCAGAAAGCCCAGTCT
GGAAGATAAAGAAAAGCCCCTATCCTCAGAAGTCAATGAAACTAAAGGATG
AGAAGGGTAAAAGTCCTTTTGAGGACTTAGAGTTGAAACAAGACCTAGTTCA
AAGCTGGATAGCTCAACTAGGAAGTGGCTCAGGAAGCAGAACGGAGAAACC
AATCTTCGACACCACCAGCAGCGACTCAGATTCTGATCTATCTGATGTCAGCT
CAAAGGTTCTAGCCTATGCTGGAGTTGAAGAAGTGGTAATGGAATACCCAAG
AAGGGTAAAAACTGCGACGGCCAAGCTAGCAGACATGGAAAAGGCTTTTGC
CGGAGAAACAACCGCAGCAGTAGGAGGAGACTCGGAGATGACAACTGGTCA
ATCTTCAAGATCTACCCTCATACCACCAAACGAAGGAGGAGGACCTATACGG
TATCCACCAGCAGAAAGACCGTCCACATCGGCCTCTACATACAACAATACAG
CCCCACCTCTCTTTGAAGGGACTGTCAGGCCTGGAAGATATGGTCGCCCTTTG
GCACCATGGTCCCTACCATCAGCACAGCACTCTCAAGGAGCCCTGCTGATCC
TCCCTCCTGAAGTAGCAAGTCACGCTGACGCCATCACCACATGGGAAACAAT
CACCCTGAATCATCTGATGAATATATCATTTGATTCCCTCCAAGACAGGGTTG
ATTATATCGAAAATCTCCTTGGACCAAGAGAACGTGAAGCATGGGTCACATG
GAGAATGGCGTATGATACGGAGTATAGACAGCTGGTTGAGCTCTCTGGGGAG
CCAAGAAATGTCACCAGCACTATCAAAAGAGTTTTAGGGATCAATGACCCCT
ACACAGGAACAACTCACATCCAGAACCAGGCTTATGCTGATCTTGAACGCCT
GCAGTGCAAAAATCTGGAGTCAGTAATGCCGTTCCTGAACTCTTATTTCCAAC
TCGCAGCAAAGAGTGGAAAAATGTGGAGCAGCCCTGAACTATCAGAAAAGC
TTTTTAGAAAGCTTCCCCCAGAAATCGGTCCTACTATAGCAAAGGAGTATGCT
GAGCGATACCCTGGTATGTTAATCGGAGTTAATGCCAGAATACAGTTCGTCTC
TGAGTATCTCCAGGACCTCTGTAAGCAAGCAGATCTTCAAAGAAAATTGAAG
AATTTGAATTTCTGCAAGGCAATTCCCATTCCTGGTTACTATGACCAAGGAGT
GAAGAAGAAGTACGGCCTACGCAAATCCAAGACATATAAAGGAAAACCTCA
TGACTCTCATGTCAAGGTTATCAAAAATAAGTATAAAGGAGCTCAAGGTCGT
AAATGCAAATGCTACCTCTGTGGTATTGAAGGCCACTATGCTCGTGAATGCCC
AAAGAAGCATGTCAGGCCTGAAAGAGCAGCCTACTTTGAAGGCATGGGCCTA
GATGTCAACTGGGATGTGATAAGTGTGGACCCAGGAGACCAAGACGGCTCAG
ACATCTGCTCAATTTCTGAAGGAGAAGAAGATGGGATGGAGGACCTAGCTGC
ATTCAAAGCCCAACTTCCATACCCAGTGGAAGCCCAATATGAACAGCACCAG
GCCTTTGTGGTTATCCAGACAACCTTTAAAAAGGAGGATAAGCCCCAAGGTT
CTTGGCGCATGTCAAAGCCCATCCTTGAAACCCAACAGCAATGCCAGCACAC
ATGGGATGACATGTATGCCCTAGCAGAAGGACAGCAAGCGTGCAGCACTTGC

CAGACCATCACTGTACTTGGTCGACGTGCTACCTGCACCCTCTGCCTACTCAA
CCTCTGCTCACTATGCGCTGGCCTAGACTTCGGTCTCAAAATAGTTCCTAAAA
CTGCAACTCGTGCTGACTGGAAATTCCAGGATCGCGATTCTCTCATCGCCTCC
TTATATGAGCACAACGCATTCCTTCTTCGACAGGTCGAAGGACTGAAACAGG
AACTCCAAGCTGCCAAGGAACAGCTTCAACTGCTACACTCGGTTGATATGAT
CAACCTCTCTGATGATGGATTAGAGAATTTTTCCGTTGAGGAAAAATCCTTTT
TAAGAGGGGGAGGGGGTACCAGTAGCAGTTCAATCAAAATCTCATCAACAAC
AACACCCCCTGGTTTTCCTACAACACCCAACAAATTCCAGCCTCTTGCGCAGG
AAAAACTAAAAGGAATACAGGAAGACTTATCTCTGGCAGTACAGTTTGATGA
TGTTAGACAACAAGAACAGGCGTATACTGAAATGCCTCGAGGAGCCCACAAC
AAGCTATACCACGTAGTGGTAACTTTCAGAATCCCTGATGTTAAGGGACAAC
TCCTTGAATTTGACATCAACGCCATCATAGACACCGGCTGTACATGCTGCTGC
ATCAACCTCACAAAGGTGCCTGATGGAGCAATCGAGAACGCCTCCATAATCC
AGGAAGTCTCTGGAATCAATAGCAAAACAGTAGTCACTAAGAAACTCAGGCA
AGGCAAGATGATCCTCGCAGGGAATGATTTCTACATTCCTTATGTTTCAGCCT
TTGAGATGAACATGCCTGGGATTGACATGCTGATAGGCTGTAACTTCATCAG
AGCAATGAAGGGAGGAATACGGTTGGAAGGAACTGAGGTCACCTTCTACAA
AACCATCACCAGGATCCAAACTACCCTGGAACCTCAAAAGATAGCGTACTTG
GAAGAGCTAGTAGAAGCAGAAGATCTACACTATGAGCTCGCAGCTGCAAGTA
TGCCTGAGCCCACTGCTGAAGGACTCAGAAACACAAAACTCCTAGCCGAGTT
AAAAGATCAAGGCTACATAGGAGAAGAGCCTCTTAAGCACTGGTCAAAGAA
TAGGGTAAGATGTAAGCTTGATATCATTAACCCTGACATCACCATTGAGGCT
AAACCACCTGGACACCTGACTCTGGAGGATAAGGTCAAATATCAGAAGCACA
TTGACGCCCTCCTAGATCTTGGAGTCATCAGACCTAGCAAGAGCAGACATAG
GTCCGCAGCTTTCATAGTTGCTTCTGGAACCTCTGTAGATCCTAAAACTGGTA
AAGAAACACGCGGTAAAGAAAGAATGGTGATCGACTACCGCATGCTAAATG
ACAATTGCTACAAGGATCAGTACAGTCTGCCTGGAATCACCTCCATCATCAA
ATCTCTTGGGCAAGCCAAAATATTCAGTAAGTTCGACTTAAAATCAGGCTTCC
ACCAAGTCATGATGGAAGAAGAAAGCATCCCCTGGACTGCTTTTATCAGCCC
CGCAGGATTATATGAATGGCTAGTTATGCCATTTGGAATTCAAAATGCACCTG
CAATCTTTCAAAGAAAGATGGATGAATGCTTCAAAGGAACTGAGGATTTCAT
CGCTGTCTATATCGATGATATTCTGGTTTTCTCTAACTCCATCAAAGAACATG
AAAAGCACCTGCAGAGAATGCTGAGTATCTGCAAGGAACATGGGCTTGTCCT
CAGCCCAACAAAAATGAAAATCGCTGTCCCAGGAATTGATTTTCTTGGTGCC
CATATCAGAAATAGCAGAGTAAGCCTGCAACCGCACATCATCAAGAAAATTG
CTGACAAGAAAGATGATGAGCTGATGACCCTCAAAGGCCTCAGAAGCTGGCT
AGGGGTAATCAACTATGTCAGGCAATACATTCCCAAGTGCGGAACACTTCTC
GGTCCCCTCTATGCTAAGACCTCTGAGCATGGTGATCGAAGATGGCACCCCA
AAGATTGGGAAATAGTGAGACAGATCAAGAAGATGGTCCAATCCCTTCCTGA
TCTTGAACTTCCTCCACCCCACGCAGTTATTATCATTGAATCTGATGGATGCA
TGGAAGGATGGGGAGGAATTTGCAAATGGAAGAACTCAAAAGGGGAATCCA
AAGGTAAAGAGCGAATCTGTGCTTACGCCAGTGGGGAATTCCCAACAGTCAA
ATCCACCATAGATGCTGAAATCTATGCAGTCATGGCATCCCTGGAGAATTTCA
AAATTTACTATCTTGATAAACGGGAAATCACCATCAGAACAGATTGCCAAGC
TATAATCAGCTTCTATGATAAGATGGCTGTCAAGAAACCCAGCAGAGTCCGC
TGGATTAATTTCTGTGATTATATCACTAACACAGGGATTAAAGTCCAGTTTGA

GCACATAAAGGGCCAAGATAATCAGCTCGCAGACCAACTCTCCAGATTAGCC
CAGAATGTCTGCGCAGTCCAAGTCATCCCTGAAGCAGCCCATGAAGCCCTCA
CCATTATCCTTGAACAAGATTGTACAGCCCAAGAGCTCATGGCCCAGTTCAA
CTCCATGCTGCAGGCTAACCTCAGGCTAAATCAAGGAAGGCCCAATACAACT
TGGTATTCTAAGACCAAGCCCAAGAAATCCAAAGCCCGCAAACCAGCCCAAG
TCCAGCCACGCTTTGACGTAAGCAATGACGATGAGGGATAATAATGGAGGAA
TCTTACAAGGACAGCACATGGTCCTTCCTTCCTCTTTTCTTTTGTAAAATTTTT
GTCTTCTTGTGTCGGCAACCGCTCCTTTTGTAAAGAGGAATCTACTTTTGAGC
TGTCGATGGGGCCCATTGAGTACCCGAGCTCCAAAAGTAACTTACCTCTGGTT
GCTTTTGTCAACCTTAGTTAGGTTTGTTTGCTTTTCTCCCCTATATAAGGGAGC
CTCTTATTTTGTAAGAAGGCATCGAACAGAGCAATACCTCTGAGCGTTCCTTC
TCTCTAGTTTTCTTGTGTTCTTGTATCTTTCCAGTTCCAGTGTTCTTAATGCAAT
TTGAAGTTTTTCTACTCTATGTCATTCTGTTCATAATTCTTTTCCGCTATTTTTA
CTCTGTGATCCAAATATTTAATTTGTGATCTGTTTCTATC