Spring 2018

Assessment of Resistance to Botrytis Cinerea in Arabidopsis Expressing Grapevine STS Genes and Analysis of New Grapevine Vein Clearing Virus Isolates

Li Su
Missouri State University, Su16@live.missouristate.edu

As with any intellectual project, the content and views expressed in this thesis may be considered objectionable by some readers. However, this student-scholar’s work has been judged to have academic value by the student’s thesis committee members trained in the discipline. The content and views expressed in this thesis are those of the student-scholar and are not endorsed by Missouri State University, its Graduate College, or its employees.

Follow this and additional works at: https://bearworks.missouristate.edu/theses

Part of the Biotechnology Commons, and the Plant Pathology Commons

Recommended Citation

https://bearworks.missouristate.edu/theses/3248

This article or document was made available through BearWorks, the institutional repository of Missouri State University. The work contained in it may be protected by copyright and require permission of the copyright holder for reuse or redistribution.

For more information, please contact BearWorks@library.missouristate.edu.
ASSESSMENT OF RESISTANCE TO *BOTRYTIS CINEREA* IN *ARABIDOPSIS*

EXPRESSING GRAPEVINE *STS* GENES AND ANALYSIS OF NEW

*GRAPEVINE VEIN CLEARING VIRUS* ISOLATES

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

Li Su

May 2018
ASSESSMENT OF RESISTANCE TO BOTRYTIS CINEREA IN ARABIDOPSIS EXPRESSING GRAPEVINE STS GENES AND ANALYSIS OF NEW GRAPEVINE VEIN CLEARING VIRUS ISOLATES

William H. Darr College of Agriculture
Missouri State University, May 2018
Master of Science
Li Su

ABSTRACT

Grapevine stilbene synthase gene (STS) family is unusually large with more than 32 members of full-length genes. Resveratrol and piceid are synthesized in VaSTS7, VvSTS22-, VaSTS22-transgenic Arabidopsis. To study the function of these STS genes in the resistance against Botrytis cinerea in vivo, lesion sizes on these STS transgenic Arabidopsis were compared to that on non-transgenic control Col-0, at 96 hours post inoculation with B. cinerea spores at 0, 200, or 1,000 spores/μl. In each experiment there were three biological replicates. Three lesion sizes from each replicate were averaged. The experiment was repeated three times. The results showed that the STS-overexpressing Arabidopsis did not show resistance to B. cinerea, indicating that single STS gene might not be fully functional in defending against B. cinerea. In a second project, two new Grapevine vein clearing virus (GVCV) isolates were sequenced and analyzed. The two GVCV isolates have 99.8% identical nucleotide sequence, thus were considered as same strain although they were isolated from two plants in different genera of the Vitaceae family. The conserved domains, secondary structure of leader RNA, and phylogenetic relationship were predicted among seven isolates. The results provide new evidence that GVCV spreads across plants in different genera. Thus, it is urgent to take measures for preventing virus spread.

KEYWORDS:  stilbene synthase, Botrytis cinerea, GVCV, genomic feature, virus

This abstract is approved as to form and content

Wenping Qiu, PhD
Chairperson, Advisory Committee
Missouri State University
ASSESSMENT OF RESISTANCE TO *BOTRYTIS CINEREA* IN ARABIDOPSIS
EXPRESSING GRAPEVINE *STS* GENES AND ANALYSIS OF NEW
*GRAPEVINE VEIN CLEARING VIRUS* ISOLATES

By

Li Su

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

May 2018

Approved:

_________________________________________________________________
Wenping Qiu, PhD

_________________________________________________________________
Chin-Feng Hwang, PhD

_________________________________________________________________
Laszlo G. Kovacs, PhD

_________________________________________________________________
Julie Masterson, PhD: Dean, Graduate College

In the interest of academic freedom and the principle of free speech, approval of this thesis indicates the format is acceptable and meets the academic criteria for the discipline as determined by the faculty that constitute the thesis committee. The content and views expressed in this thesis are those of the student-scholar and are not endorsed by Missouri State University, its Graduate College, or its employees.
ACKNOWLEDGEMENTS

I would like to thank my mom, dad, and aunt, who do not usually understand me but always respect and believe my decisions. Also I would like to thank my boyfriend Yi. You always urge me doing the things I should do although that is a little nagging. Thank my dear friend Biying for sharing the happiness and sadness with me for ten years.

Thank my advisor Dr. Qiu for your passion towards research, which greatly affect my attitude to study and research. Thanks again for your encouragement and detailed revising during my writing. Under your instruction, I began to be more honest about myself. Thank Dr. Hwang for teaching me critical thinking and always giving me great suggestions. Thank Dr. Kovacs for your suggestion to my research and your detailed revising. You all are so kind to me! Thank all of you for being my committee members and giving me such great helps!

I truly appreciate every teacher I have ever had or met in Missouri State University for being so patient and kind to me. You always give me much more help than I expected. Every class or presentation I took in the campus makes me feel more confident about what I am studying.

Lastly, I am very grateful to my lab mates Sylvia Petersen and Cory Keith. We learn, compete, and collaborate with each other. Staying with them in the lab is the most joyful thing in MSU. Also, thank Ru, Kaylie, Susanne, Dr. Li, Joshua, Yanning and Trystan for sharing their lab knowledge and experience. Thank all of you for your help and being so kind to me.
# TABLE OF CONTENTS

## CHAPTER 1: Assessment of Resistance to *Botrytis Cinerea* in Arabidopsis Expressing Grapevine STS Genes

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Grapevine</td>
<td>1</td>
</tr>
<tr>
<td>Plant Resistance</td>
<td>2</td>
</tr>
<tr>
<td>Grapevine Stilbenes</td>
<td>4</td>
</tr>
<tr>
<td>Stilbene Synthase Genes in Grapevine</td>
<td>5</td>
</tr>
<tr>
<td>Objectives</td>
<td>6</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>7</td>
</tr>
<tr>
<td>Overall Experimental Design</td>
<td>7</td>
</tr>
<tr>
<td>Plants</td>
<td>7</td>
</tr>
<tr>
<td>Culture and Maintenance of Grey Mold (<em>B. cinerea</em>)</td>
<td>8</td>
</tr>
<tr>
<td>Infection with Grey Mold</td>
<td>8</td>
</tr>
<tr>
<td>Statistical Methods</td>
<td>8</td>
</tr>
<tr>
<td>Results</td>
<td>9</td>
</tr>
<tr>
<td>Assessment of <em>VaSTS22</em>-transgenic Arabidopsis against <em>B. cinerea</em></td>
<td>9</td>
</tr>
<tr>
<td>Assessment of <em>VvSTS22</em>-transgenic Arabidopsis against <em>B. cinerea</em></td>
<td>9</td>
</tr>
<tr>
<td>Assessment of <em>VaSTS7</em>-transgenic Arabidopsis against <em>B. cinerea</em></td>
<td>10</td>
</tr>
<tr>
<td>Discussion</td>
<td>10</td>
</tr>
</tbody>
</table>

## CHAPTER 2: Analysis of New *Grapevine Vein Clearing Virus* Isolates

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>18</td>
</tr>
<tr>
<td>Badnaviruses</td>
<td>18</td>
</tr>
<tr>
<td><em>Grapevine Vein Clearing Virus</em></td>
<td>20</td>
</tr>
<tr>
<td>Overall Objectives and Major Discoveries of This Research</td>
<td>23</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>24</td>
</tr>
<tr>
<td>Materials</td>
<td>24</td>
</tr>
<tr>
<td>DNA Extraction and Triplex PCR of Detecting GVCV</td>
<td>24</td>
</tr>
<tr>
<td>Cloning of the Viral Genome and Sequencing Strategy</td>
<td>25</td>
</tr>
<tr>
<td>Genome Sequence Analysis and Genomic Feature Prediction</td>
<td>26</td>
</tr>
<tr>
<td>Results</td>
<td>27</td>
</tr>
<tr>
<td>Symptoms on Two Infected Vines</td>
<td>27</td>
</tr>
<tr>
<td>Comparison of the Two New GVCV Isolates</td>
<td>27</td>
</tr>
<tr>
<td>Genome Analysis among Seven GVCV Isolates</td>
<td>28</td>
</tr>
<tr>
<td>Annotation and Prediction of Genomic Features</td>
<td>29</td>
</tr>
<tr>
<td>Phylogenetic Relationship among Seven GVCV Isolates</td>
<td>29</td>
</tr>
<tr>
<td>Prediction of the Leader Sequence in the Pre-genomic RNA of Seven GVCV Isolates</td>
<td>30</td>
</tr>
<tr>
<td>Discussion</td>
<td>30</td>
</tr>
<tr>
<td>References</td>
<td>51</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1. A list of primers used for amplifying .............................................................. 35
Table 2. Components of PCR for amplifying ............................................................... 36
Table 3. Thermocycler protocol for amplifying ........................................................... 37
Table 4. Comparative analysis of seven GVCV isolates .............................................. 38
Table 5. Nucleotide identity of intergenic region (IGR) ............................................ 39
Table 6. Nucleotide identity and amino acid identity of ORF I .................................. 40
Table 7. Nucleotide identity and amino acid identity of ORF II ............................... 41
Table 8. Nucleotide identity and amino acid identity of ORF III ............................. 42
LIST OF FIGURES

Figure 1. Mean lesion size on leaves of Col-0................................................................. 14
Figure 2. Mean lesion size on leaves of VaSTS22-transgenic ........................................ 15
Figure 3. Mean lesion size on leaves of VvSTS22-transgenic........................................... 16
Figure 4. Mean lesion size on leaves of VaSTS7-transgenic ............................................ 17
Figure 5. Amplification and assembling strategy of GVCV .............................................. 43
Figure 6. GVCV associated symptoms.............................................................................. 44
Figure 7. Pairwise nucleotide identity analysis................................................................. 45
Figure 8. Pairwise amino acid identity analysis................................................................. 46
Figure 9. GVCV genome structure.................................................................................... 47
Figure 10. Graphical representation of the conserved domains......................................... 48
Figure 11. Phylogenetic relationship among seven GVCV isolates.................................... 49
Figure 12. Predicted structure of leader RNA. ................................................................. 50
CHAPTER 1: ASSESSMENT OF RESISTANCE TO *BOTRYTIS CINEREA* IN ARABIDOPSIS EXPRESSING GRAPEVINE STS GENES

Introduction

**Grapevine.** Grapevine (*Vitis* spp.) is among the most widely cultivated fruit crops in the world with important economic value (Dai, 2016). Around 71% of the grape products are used to make wine (Papademetriou & Dent, 2001). Most wine grapes are cultivated varieties of the species *Vitis vinifera* (Dai, 2016). The cultivar ‘Cabernet Sauvignon’ is the most renowned grape variety and is widely grown for producing a fine red wine (Bowers & Meredith, 1997; Dai, 2016). However, this Eurasian cultivar is highly susceptible to most grapevine pathogens, including grapevine grey mold (Mullins *et al.*, 1992). The causative agent of grey mold is fungal pathogen *Botrytis cinerea* which attacks more than 200 different plant species (Elad, 1997). The *Botrytis* spores prefer entering and germinating from wounded sites on mature tissues, such as mature flower petals, dead bulb scales, and dying foliage in cool and humid conditions, causing a notable greyish cover with fungal spores on the infected tissues (Agrios, 2005). Different parts of the plant could be infected by *B. cinerea*. Symptoms are various, including damping-off on seedling, blight on blossomed flowers, leaf spot on foliage, fruits and stem rot (Agrios, 2005).

The American hybrid cultivar ‘Norton’ is a fortuitous cross between *Vitis aestivalis* and ‘Bland’, *labrusca × vinifera* hybrid. It was discovered by Dr. Daniel Norborne Norton (Ambers & Ambers, 2004). Unlike Cabernet Sauvignon, it is cold-hardy and resistant to most fungal diseases including grey mold (Sapkota *et al.*, 2015).
When applied with the same concentration of *B. cinerea* spore suspension, Cabernet Sauvignon showed an average disease incidence and disease severity greater than 90% while less than 10% Norton berries were infected after inoculation (Sapkota *et al.*, 2015). The ‘Norton’ grape also produces high-quality red wine favored by Southern and Midwest consumers (Ambers & Ambers, 2004). These superb properties make Norton an elite red grape cultivar. This cultivar is Missouri’s official state grape and accounts for 16% of all grapes grown in Missouri (MissouriWines, 2015).

**Plant Resistance.** During their lifetime, perennial plants are exposed to various types of plant pathogens. In the co-evolution with diverse pathogens, they have developed different levels of immunity. The basal defense is PAMP/MAMP-triggered immunity (PTI), a result of interaction of pathogen/microbe-associated molecular patterns (PAMP) and plasma membrane-localized pattern recognition receptors (PRRs) (Eulgem & Somssich, 2007). For example, flg22, a highly conserved 22 amino acid peptide in the bacterial flagellin N-terminus, is recognized by FLS2, an intensively studied Arabidopsis PRRs, to induce PTI (Chisholm *et al.*, 2006). It is expressed in response to influx of H⁺ and Ca⁺, activation of calcium-dependent protein kinases and mitogen-activated protein kinase (MAP kinase), and transcriptional reprogramming of defense genes (Asai *et al.*, 2002). Faced with the plant PTI, pathogens employ effectors to hamper the pathogen-responsive signaling transduction pathway in plants (Baker *et al.*, 1997). Co-evolved R proteins in plants recognize these pathogen effectors to trigger a highly effective defense reaction, termed as effector-triggered immunity (ETI) (Jones & Dangl, 2006). Both PTI and ETI are the result of massive defense-oriented transcriptional reprogramming within the infected cells (Schön *et al.*, 2013; Hahlbrock *et al.*, 2003). The
induced response includes the hypersensitive response, the production of reactive oxygen species, the reinforcement of host cell walls, the synthesis of proteins toxic to the invading pathogens and enzymes to detoxify pathogen toxins, and the accumulation of phenolics and phytoalexins (Briggs, 1995; Levine et al., 1996).

Phytoalexins are a group of small-molecular-weight secondary metabolites produced by plant cells in response to pathogen attack (Hammerschmidt, 1999). Phytoalexins possess antimicrobial activity (Ahuja et al., 2012; Pedras et al., 2011). The concept was introduced when it was found that the potato plants (Solanum tuberosum) would induce resistance to the compatible race of Phytophthora infestans after the plants were previously infected with incompatible race of P. infestans (Ahuja et al., 2012). It is speculated that the plant was induced by the incompatible pathogen to produce substance (phytoalexins) that could be used to defend the later attack of a compatible pathogen and limit the growth of the pathogen to protect the plant cells (Ahuja et al., 2012). Different plants produce different phytoalexins. Apigeninidin and luteolinidin, primary phytoalexins in sorghum, belong to the 3-deoxyanthocyanidin chemical group. In maize, the members of the terpenoid group are believed to function as phytoalexins. The biosynthesis of the phytoalexin in these plants are well understood (Poloni & Schirawski, 2014; Jeandet, 2015). The specific phytoalexin in Arabidopsis, is camalexin (Browne et al., 1991). Stilbenes are a family of phenolic phytoalexins belonging to the non-flavonoid polyphenol group and are produced in the phenylalanine/polymalonate pathway (Langcake & Pryce, 1976). Stilbenes are thought to be important defense molecules in grapevine (Chalal et al., 2014).
Grapevine Stilbenes. Stilbenes naturally occur in several plant families such as Cyperaceae and Vitaceae (Hart, 1981; Sotheeswaran & Pasupathy, 1993). Grapes and grape products, especially red wine, are thought to be the major source of stilbene compounds in diet (Waterhouse et al., 1997; Waffo-Teguo et al., 2008). The concentration of stilbenes could reach 50 mg/l in red wine (Waffo-Teguo et al., 2008).

The basic unit of the stilbene family is trans-resveratrol (3,5,4’-trihydroxy-trans-stilbene). Also, it is the most active and studied stilbene in grapevine (Jeandet et al., 2002; Chong et al., 2009). From this simple structure, several derivative compounds are synthesized through the substitution of hydroxyl groups with other residues such as sugars, methyl and methoxy or through alternative steric configuration of the molecules (Jeandet et al., 2010). The resveratrol could be oxidatively condensed into oligomers such as dimers, trimers and tetramers (Waffo-Teguo et al., 2008). These modifications are necessary for the function of each compound (Pont & Pezet, 1990; Orsini et al., 1997).

In grapevine, stilbenes are maintained in high concentration in the hardwood to prevent the development of wood decay (Hart, 1981; Jeandet et al., 2010). Stilbenes in berry skins accumulated after grapevine were infected by several major pathogens, such as powdery mildew (Erysiphe necator) (Romero-Pérez et al., 2001) and downy mildew (Plasmopara viticola) (Van Leeuwen et al., 2012).

Trans-resveratrol concentration increases after véraison and reaches highest levels at harvest stage (Dai et al., 2012). The trans-resveratrol accumulates in the infected site, showing moderate antimicrobial properties (Jeandet et al., 2010). Resveratrol was also observed to inhibit the germination of B. cinerea spore in liquid culture by deforming the
B. cinerea conidia in vitro (Adrian et al., 1997). Current research focuses on the identification of more active stilbene derivatives involved in plant defense. When 13 trans-resveratrol derivatives in grapevine were assessed for antimicrobial ability, only 3 of them were effective against grey mold and downy mildew. They are pterostilbene, 2-hydroxy, 4-methoxystilbene and 3-methoxy-4,4'-hydroxystilbene (Chalal et al., 2014).

**Stilbene Synthase Genes in Grapevine.** Stilbene synthase (STS) is the last enzyme in the biosynthesis pathway of stilbenes. It is a member of the type III polyketide synthase family and catalyzes three molecules of malonyl-CoA and one CoA-ester of a cinnamic acid derivative (p-coumaroyl-CoA) to form trans-resveratrol (Austin & Noel, 2003). STS and chalcone synthase (CHS) are closely related. The two enzymes share 75-90% identical amino acid sequence (Austin et al., 2004). They use the same substrates in the phenylalanine or flavonoid pathway, producing chalcone, a complicated C_{15} compound, and simple stilbene unit, respectively (Jeandet et al., 2010).

After stilbene synthase was firstly purified from Arachis hypogea (peanut) (Schöppner & Kindl, 1984), stilbene synthase genes have been found in several plants. They form small families consisting of two to five closely related paralogs (Parage et al., 2012). Five STS genes have been found in pine (Preisig-Müller et al., 1999). Three STS genes have been isolated from Pinus densiflora (Kodan et al., 2002). Only one STS gene, the ShSTSI gene, was isolated from Sorghum, a monocotyledonous plant (Christine et al., 2005). In grapevine, a large family of STS genes have been identified in the 12x genome sequence of the grapevine ‘Pinot Noir’ cultivar (Parage et al., 2012; Vannozzi et al., 2012). Forty-eight STS genes are annotated, with thirty-two complete genes, five partial genes and eleven probable pseudogenes (Parage et al., 2012). Six of the STS genes are
located on chromosome 10. The other forty-two STS genes are clustered on chromosome 16 (Parage et al., 2012). Grapevine is among the plant species with the highest number of stilbene synthase genes (Jeandet et al., 2010). The large number of STS genes and the condensed distribution on single chromosome suggest the importance of stilbene in the lifespan of grapevine.

The expression of STS genes in grapevine was shown to be inducible in response to biotic and abiotic stresses. Plant hormones regulate the accumulation of stilbene compound in grapevine. Application of salicylic acid (SA) enhanced the expression of STS genes in *V. amurensis* (Kiselev et al., 2010). The expression of STS in grapevine increased upon application of ethylene (Belhadj et al., 2008). It was shown that the STS mRNAs accumulated after the cell of *V. vinifera* was treated with the cell walls of *Phytophthora cambivora*. There were two peaks of the STS mRNAs transcriptional level, 6 and 20 h after treatment (Wiese et al., 1994). Early in 1991, researchers used fragments of *B. cinerea* cell wall and cultured grapevine cells to simulate interaction between grapevine and *B. cinerea*. They found that upon treatment with the fungal cell wall material, general protein synthesis shut off, except for the synthesis of stilbene synthase and a few other resistance related proteins. Interestingly there were nearly undetectable levels of stilbene compounds in the grapevine cell before the treatment, but in response to the treatment, grapevine cells produced and secreted resveratrol and its derivatives (Melchior et al., 1991).

**Objectives.** Transcripts of STS22 and STS7 increased differentially in ‘Norton’ and ‘Cabernet Sauvignon’ in response to the grapevine powdery mildew (Dai et al., 2012). The STS22 genes from both cultivars encode 98.2% identical amino acid
sequences. The mechanism underlying differential expression of STS22 and STS7 in two cultivars has not been studied before. It is still unknown if stilbene compound produced by the two STS gene-encoded enzymes have any antimicrobial activities.

In this project, the activities against a necrotrophic pathogen (B. cinerea) of the two STS genes were studied. Since Arabidopsis does not encode STS genes on its genome, and malonyl-CoA and p-coumaroyl-CoA, the substrates of the STS, exist in Arabidopsis (Kubasek et al., 1992), we used the model plant Arabidopsis to generate STS-transgenic lines and to investigate the functions of grapevine STS genes. In a previous project, stilbene compounds have already been assayed in transgenic Arabidopsis over-expressing VaSTS7, VvSTS22 and VaSTS22 in Col-0 background. In this study, evaluation of disease severity was quantified by measuring the lesion size on the transgenic Arabidopsis to study the antifungal activities of the two STS genes.

Materials and Methods

Overall Experimental Design. Five or six-week old T3 generation of Arabidopsis over-expressing the STS genes were inoculated with different concentration of B. cinerea spore suspension (0, 200 spores/μl and 1,000 spores/μl). Leaves from transgenic and untransformed Arabidopsis were collected and maintained on agar plates, and then inoculated with spores. Lesion size was recorded at 96 h after inoculation. Experiments with spore inoculation were repeated three times.

Plants. T3 generation of Arabidopsis over-expressing the STS gene in the ecotype Col-0 background have been generated in a previous research project (Dai, 2016). The transgenic Arabidopsis were grown under conditions of 25 °C, 12h light/12h dark.
**Culture and Maintenance of Grey Mold (B. cinerea).** *B. cinerea* strain was isolated from grape berries in a vineyard at the Missouri State Fruit Experiment Station and cultivated in the laboratory throughout this project. The fungus was maintained on potato dextrose agar (PDA) cultural media. All cultures were kept in the dark at room temperature. New spores were inoculated onto PDA cultural media every 10-14 days (Austin, 2017).

**Infection with Grey Mold.** To induce sporulation, *B. cinerea* cultures were incubated at 25 °C under 12h light/ 12h dark. Conidia of *B. cinerea* were collected from PDA culture media plate using water, and filtered through cheesecloth (Dai, 2016). *B. cinerea* spores were counted using a Marienfeld counting chamber (Austin, 2017). The conidial suspension was adjusted to 200 spores/μl or 1000 spores/μl. The 5 or 6-week old transgenic Arabidopsis and the Col-0 leaves were collected and wounded by small pipette tips, then treated with a 5 μl drop of spore suspension (Dai, 2016). Every leaf was inoculated with spores at 3 spots and maintained on 1% agar plates. After inoculation, the plates were placed in darkness for first 24 hours, followed by incubating at 26 °C, 12h light/ 12h dark (Dai, 2016). Three biological replicates were treated, each replicate consisted of one leaf with three inoculation sites. The experiment was repeated three times.

Symptoms of *B. cinerea* infection were inspected and recorded at 96 h after inoculation. The necrotic lesion sizes at the infection sites were measured by using image analysis software Assess 2.0.

**Statistical Methods.** The lesion sizes from three inoculated sites in one leaf were measured and averaged as one replicate. There were three biological replicates in each
experiment. The mean and standard deviation from three independent experiments were calculated and analyzed by Excel 2016 using Student’s $t$ test. To decide the type of the $t$ test, $F$ test was used to analyze if the variances of the two groups are equal. When variances between two groups were equal, type 2 $t$ test was used. When variances between two groups were non-equal, type 3 $t$ test was used. * and ** indicate differences between non-transgenic control and transgenic lines that are statistically significant at the $p<0.05$ and $p<0.01$ level, respectively. The data was displayed as mean ± standard error.

**Results**

**Assessment of $VaSTS22$-transgenic Arabidopsis against $B. cinerea$.** As shown in Figure 1, lesion size on Col-0 inoculated with 200 spores/μl or 1000 spores/μl $B. cinerea$ increased in comparison to those on Col-0 leaves inoculated with 0 spores/μl $B. cinerea$ ($p<0.05$), indicating Col-0 is susceptible to $B. cinerea$ and thus is used as a control in the study of $B. cinerea$ disease.

Three independent $VaSTS22$-transgenic lines were grown under the same condition for disease assessment. They were named as 15-1-8, 15-3-5, and 15-12-6. As shown in Figure 2, upon the inoculation of $B. cinerea$ (200 spores/μl or 1000 spores/μl), no significant difference in the lesion size was observed between three $VaSTS22$-transgenic lines and the non-transgenic control, Col-0. Infected with 0 spores/μl $B. cinerea$, 15-3-5 show larger lesion size in comparison to the non-transgenic line ($p<0.05$).

**Assessment of $VvSTS22$-transgenic Arabidopsis against $B. cinerea$.** Three independent $VvSTS22$-transgenic lines, 17-1-1, 17-2-4, and 17-6-6, were assessed. After inoculation with 200 spores/μl or 1000 spores/μl $B. cinerea$, lesion sizes induced by $B.$
cinerea spores on three VvSTS22-transgenic lines were not significantly different in comparison to the non-transgenic control Col-0 (Figure 3). The results showed that over-expression of VvSTS22 did not change the size of disease lesions upon the infection with B. cinerea. In the 0 spores/µl B. cinerea inoculation, the lesion sizes on two independent transgenic line, 17-1-1 and 17-2-4, increased relative to the non-transgenic control.

Assessment of VaSTS7-transgenic Arabidopsis against B. cinerea. Only one VaSTS7-transgenic line, 20-5-2, was generated and assessed for resistance against B. cinerea. There was no significant difference in the lesion size between non-transgenic control and the VaSTS7-transgenic line with the treatment of 0 spores/µl, 200 spores/µl, or 1000 spores/µl of B. cinerea (Figure 4).

Discussion

The 12× Grapevine reference genome (inbred V. vinifera cv. Pinot noir PN40024) was searched for presence and annotation of STS genes by two groups independently in 2012 (Parage et al., 2012; Vannozzi et al., 2012). As a result, 48 VvSTS genes were annotated. Grapevine contains the largest STS gene family ever known. VvSTS1-VvSTS6 are located on chr10, within 90 Kb region, whereas the left 42 STS gens are located on chr16, within 500 Kb region (Vannozzi et al., 2012). Considering the average length of the coding region of STS gene is 1 Kb, the STS genes in grapevine genome is highly condensed. More than 32 STS genes are predicted to have a complete ORF coding region (Vannozzi et al., 2012). The complete set of grapevine STS genes are clustered into three clades by phylogenetic analysis. Group A contains all the complete STS genes on chr10, while Group B and Group C contain 22 and 13 STS genes, respectively, from chr16. The
members of \( VvSTS \) Group A appear to be expressed constitutively and developmentally, while the members of the \( VvSTS \) Group B and Group C tend to be induced by stress such as mechanical wounding, UV-C, and downy mildew infection (Vannozzi et al., 2012). Although it was observed that members of both Group A and Group C respond to these stresses, it should be noted that the transcriptional activity of the members in Group B is so strong that even in the early stage of the stress, the transcriptional level of member in Group B is still greater than that observed for the members in Group A and C. Moreover, the transcription of members in Group B continues to increase up to the late stage of the stress (Vannozzi et al., 2012).

In this study, the \( STS7 \) belongs to Group B and \( STS22 \) belongs to Group C. They were assumed to be induced by stresses as reported by Vamzonni et al. (Vannozzi et al., 2012). Indeed, in the previous study, transcripts of \( VvSTS7 \) and \( VvSTS22 \) increased after the grapevine was infected with powdery mildew. The two genes were amplified from ‘Norton’ and ‘Cabernet sauvignon’, and were introduced into Arabidopsis to test their ability to inhibit fungal pathogens (Dai, 2016). \( B. \ cinerea \) was chosen since it causes damage to grape berries. Moreover, this pathogen is easy to propagate and maintain under laboratory conditions. In Figure 1, it is observed that the lesion size on non-transgenic Col-0 leaves increases as the concentration of \( B. \ cinerea \) spores increases \((p<0.05)\), which indicates that the Col-0 is susceptible to \( B. \ cinerea \) and thus is suitable as a control. In the \( STS \)-transgenic line, resveratrol and trans-piceid were detected by HPLC in a previous study (Dai, 2016). The resveratrol is generally considered as the direct product of \( STS \) enzymes (Jeandet et al., 2010), whilst the trans-piceid is derived from resveratrol by the endogenous glycosyltransferases in Arabidopsis (Parage et al., 2012;
Huang et al., 2016). As discussed by other researchers, sub-lethal or lethal concentrations of resveratrol cause cytological abnormalities in B. cinerea conidia (Adrian et al., 1997). One would expect that over-expression of the STS genes in Arabidopsis is able to generate resveratrol and its derivatives to enhance the defense against B. cinerea. However, in this study, the transgenic lines did not show any sign of resistance to B. cinerea.

One possible reason could be that the amount of stilbene compound produced in the transgenic line is insufficient to reach the effective level to limit the growth of B. cinerea and prevent the development of the disease. In this project, the STS-transgenic Arabidopsis synthesize 4.54 -38.66 μg/g trans-piceid in fresh weight (Dai, 2016), while the same compound produced in VqSTS21-transgenic Arabidopsis reaches 216.30-531.15 μg/g in fresh weight (Huang et al., 2016). However, with the high yield of stilbene compound, the VqSTS21-transgenic Arabidopsis does not show enhanced resistance to B. cinerea either (Huang et al., 2016). VaSTS19-transgenic Arabidopsis, on the other hand, show improved resistance to B. cinerea (Wang et al., 2017). In general, grapevine STS-transgenic Arabidopsis do not show consistent effect on the resistance to B. cinerea. Although stilbene compounds show antifungal activity against B. cinerea in vitro (Pezet & Pont, 1990; Adrian et al., 1997), studies on STS-transgenic plants reveal that the role of the stilbene compounds on the defense against B. cinerea is much more complicated in vivo.

The grapevine genome contains large numbers of highly conserved STS genes, which might be a result of large-scale genome duplication during angiosperm evolution (Moore & Purugganan, 2003; Tang et al., 2008; Vannozzi et al., 2012). By evaluating the
selection pressures acting on the grapevine STS family, researchers found that the evolution of grapevine STS family is dominated by purifying selection, which indicates the coded enzymes is globally strongly constrained with similar catalytic activity (Parage et al., 2012). This assumption is confirmed in grapevine STS-transgenic plants, since same compound, resveratrol and piceid, are detect in the STS-transgenic plants, such as the transient VaSTS7-transgenic grape cells (Kiselev & Aleynova, 2016), VqSTS21-transgenic Arabidopsis (Huang et al., 2016). Also, nine representative STS genes from grapevine were transiently transformed into N. benthamiana leaves. As a result, piceid was detected in all these transformed leaves, indicating all putative grapevine STS genes encode STS enzymes with functional redundancy (Parage et al., 2012). One hypothesis proposed to explain multiple-gene family expansion is selection for increasing dosage (Conant & Wolfe, 2008; Parage et al., 2012), which could be obtained by duplicating a gene multiple times (Parage et al., 2012). This hypothesis could also be used to explain that maybe one copy of STS gene in the genome of the transgenic Arabidopsis is unable to effectively elevate stilbene compounds in the transgenic Arabidopsis.

In conclusion, after inoculation of 0, 200, or 1 000 spores/μl B. cinerea, lesion sizes on the VaSTS7-, VaSTS22-, and VvSTS22-transgenic Arabidopsis were compared to that of the non-transgenic control Col-0. These STS-overexpressing Arabidopsis have not shown any sign of improved resistance to B. cinerea after three independent experiments although they were able to synthesize resveratrol and piceid.
Figure 1. Mean lesion size on leaves of non-transgenic *A. thaliana* Col-0 lines after they were inoculated with *B. cinerea* at concentration of 0, 200, and 1000 spores/μl suspension. Lesion size was measured at 96 hpi. Data were presented as the means ± SE from three independent experiments (n=3). Asterisk * and ** indicate statistically significant level between *B. cinerea* spore-inoculated leaves and mock-treated (0 spores/μl) leaves at the p<0.05 and p<0.01, respectively.
Figure 2. Mean lesion size on leaves of VaSTS22-transgenic and non-transgenic *A. thaliana* Col-0 lines after they were inoculated with *B. cinerea* spore suspension at 0, 200, and 1000 spores/μl. The lesion size was measured at 96 hours post inoculation (hpi). Data were presented as the means ± SE from three independent experiments (n=3). Asterisk * and ** indicate statistically significant level between transgenic and non-transgenic lines at the *p*<0.05 and *p*<0.01, respectively.
Figure 3. Mean lesion size on leaves of *VvSTS22*-transgenic and non-transgenic *A. thaliana* Col-0 lines after they were inoculated with *B. cinerea* spore suspension at 0, 200, and 1000 spores/μl. The lesion size was measured at 96 hpi. Data were presented as the means ± SE from three independent experiments (n=3). Asterisk * and ** indicate statistically significant level between transgenic and non-transgenic lines at the p<0.05 and p<0.01, respectively.
Figure 4. Mean lesion size on leaves of *VaSTS7*-transgenic and non-transgenic *A. thaliana* Col-0 lines after they were inoculated with *B. cinerea* spore suspension at 0, 200, and 1000 spores/μl. The lesion size was measured at 96 hpi. Data were presented as the means ± SE from three independent experiments (n=3). Asterisk * and ** indicate statistically significant level between transgenic and non-transgenic lines at the *p*<0.05 and *p*<0.01, respectively.
CHAPTER 2: ANALYSIS OF NEW GRAPEVINE VEIN CLEARING VIRUS ISOLATES

Introduction

**Badnaviruses.** Plant viruses in the genus *Badnavirus* of the family *Caulimoviridae* are pararetroviruses (Chiumenti *et al*., 2016). So far, 32 species of *Badnavirus* have been discovered. They infect a broad range of economically important crops such as cocoa, banana, sugarcane, citrus and yams (Umber *et al*., 2017). Badnaviruses cause 10% to 90% economic losses in various crops, especially in the tropical and temperate climate regions of Asia, Africa, Australia, Europe, South and North America (Bhat *et al*., 2016). *Commelina yellow mottle virus*, the type species of the *Badnavirus* genus, was discovered in 1978 and was the first badnavirus identified (Migliori & Lastra, 1978; Medberry *et al*., 1990).

The genomes of badnaviruses are circular double-stranded DNA varying from 7.2 to 9.2kb in length (Bhat *et al*., 2016), with one discontinuity in each strand (Medberry *et al*., 1990; Chiumenti *et al*., 2016). The genomic DNAs of viruses in the *Badnavirus* and *Tungrovirus* genus are contained in a bacilliform virion, whereas viruses in other genera in the family of *Caulimoviridae* have isometric virions (Bhat *et al*., 2016). The genomic dsDNA is released from the capsid after the pararetroviruses penetrate the host plant cell. Subsequently, the discontinuous sites are repaired by the polymerase and ligated by ligase in the plant nucleus and then a minichromosome is formed (Hohn & Rothnie, 2013; Chiumenti *et al*., 2016). A pre-genomic RNA with terminal redundancies is transcribed and transferred into the cytoplasm for further translation of viral proteins or
served as the template for genome replication (Chiumenti et al., 2016). Using primer extension method, researchers mapped the 5’-ends of both plus- and minus-strand of *Commelina yellow mottle virus* (CoYMV) and found that the CoYMV uses the host cytosolic initiator methionine tRNA and polypurine-rich region in the pre-genomic RNA as the primer for synthesizing the minus and the plus strand, respectively (Medberry et al., 1990). Although plant pararetroviruses (Family *Caulimoviridae*), similarly to retroviruses use an intermediate RNA as template to make the dsDNA genome, not all pararetroviruses genomes integrate into the host genome. The presence of integration has no direct relationship with the viral infection. In the *Badnavirus* genus, *Banana streak OL virus* (BSOLV), *Banana streak GF virus* (BSGFV), *Banana streak IM virus* (BSIMV), and *Tobacco vein clearing virus* (TVCV) are known to be integrated into the host genome. They generate episomal forms and cause infections in certain hybrid plants under specific abiotic stresses (Lockhart et al., 2000; Gayral et al., 2008; Chabannes et al., 2013; Bhat et al., 2016).

Most badnaviruses, encode three main ORFs on the positive strand (Chiumenti et al., 2016). ORF I encodes a protein with a conserved domain (pfam07028: DUF1319), but its function remains elusive (Sether et al., 2012; Chiumenti et al., 2016). In ComYMV, researchers found that ORF I co-localized with immature virions (Cheng et al., 1996). ORF II is the shortest coding region. In *Cacao swollen shoot virus* (CSSV), it was shown to be a nucleic acid-binding protein (Jacquot et al., 1996). The ORF III has the longest sequence, ranging from 5,100 to 6,000 bp in length (Bhat et al., 2016). It encodes four proteins in the following order: a putative cell-to-cell movement protein,

**Grapevine Vein Clearing Virus.** In 2004, severe symptoms, including short internodes, crinkled and mosaic leaves were observed in *Vitis vinifera* ‘Chardonnay’ vines in a Missouri vineyard (Qiu *et al*., 2007). More than 90% of the vines were affected in this commercial vineyard and had to be removed and destroyed in 2007 (Lunden *et al*., 2008). To exclude the possibility that the symptoms were caused by non-biological factors, such as adverse nutrition condition, hardwood cuttings of the symptomatic vine were collected and propagated in the greenhouse. Similar symptoms appeared on the new leaves and shoots on the propagated vines under greenhouse conditions, which indicated that the symptoms were possibly caused by virus-like pathogens rather than environmental factors (Qiu *et al*., 2007).

To identify the causative agent, enzyme-linked immunosorbent assay (ELISA) or reverse transcription polymerase chain reaction (RT-PCR) were used to detect six common viruses in this symptomatic Chardonnay vines (Qiu *et al*., 2007). Although *Grapevine fanleaf virus* (GFLV), *Tomato ringspot virus* (ToRSV) and *Grapevine rupestris stem pitting-associated virus* (GRSPaV) were detected in 11 affected vines by RT-PCR, these viruses were not closely associated with the vein clearing symptom (Zhang *et al*., 2011).

To discover whether new viruses were associated with the disease, small RNAs (sRNAs) from the symptomatic Chardonel grapevine LBC0903 were sequenced. Sequences with length between 16 nt and 28 nt were selected for further analysis. These sRNA sequences were then assembled into contigs (Zhang *et al*., 2011). After they had
been searched against all available genome sequences of viruses and viroids downloaded from national center for biotechnology information (NCBI) databases, the contigs that share high homology with viral sequences were selected for further analysis (Zhang et al., 2011). Primers were designed from the highly conserved region shared by multiple reads to amplify the virus genome. Three overlapping DNA fragments were cloned, sequenced, and then assembled into an entire viral genome of 7,753 base pairs in length. The assembled virus was given a provisional name *Grapevine vein clearing virus* (GVCV). This isolate was named as GVCV-CHA. The GVCV genome starts from 5’-TGGTATCAGAGCTCCAG on the positive strand, which is complementary to the 3’-end 12 nt of the plant tRNA²met sequence (3’-ACCAUAGUCGUGGCUCAAA), following the numbering system conventional for other genomes of plant pararetroviruses. The three ORFs predicted on the plus-strand of the GVCV genome were similar to those found on the genome of ComYMV (Medberry et al., 1990). The assembled genome sequence shows highest homology with members of the genus *Badnavirus* in the family *Caulimoviridae* (Zhang et al., 2011). The nucleotide sequences which encode RT and RNase H differ from other badnaviruses by more than 20%, indicating that it is a new virus in the genus *Badnavirus* (Zhang et al., 2011). Based on the genome sequence, sets of primers were designed and were able to amplify identical fragments by using PCR with genomic DNA as template, suggesting that the GVCV is present in the infected cells as a DNA form (Zhang et al., 2011). The GVCV-specific fragments were then amplified in symptomatic vines but not in asymptomatic vine. In addition, a previous study found that healthy Chardonnay and Cabernet Franc showed similar vein clearing symptoms after they were grafted with two buds of GVCV-infected Chardonnay vines. The GVCV
specific fragments were detected both in the original symptomatic Chardonnay vine and
the grafted vines, suggesting that the presence of the GVCV genome in vines is
associated with vein clearing symptoms (Zhang et al., 2011).

To confirm the causative agent of the disease, Koch’s postulates needed to be
fulfilled. However, since the virus is too small to be isolated and can’t be cultured on the
culture media, grafting infected vines onto the healthy indicator vines was used as an
alternative to infer that the virus is the causative agent of the disease. In addition, by
using graft-transmission, we could also investigate whether certain cultivars are
susceptible to GVCV. By now, there are 13 cultivars or hybrid grapevines, mostly
possessing major genetic background of V. vinifera, were found susceptible to GVCV.
They are ‘Chardonnay’ (Qiu et al., 2007), ‘Cabernet Franc’, ‘Baco Blanc’, ‘LN-33’
(Lunden et al., 2008), ‘Chardonel’, ‘Cabernet Sauvignon’, ‘Riesling’, ‘Corot noir’
(Zhang et al., 2011), ‘Vidal Blanc’, ‘Cayuga White’, ‘Traminette’ (Guo et al., 2014),
‘Valvin Muscat’, ‘Vignette’ (Beach et al., 2016). Only cultivars ‘Chambourcin’ and
‘Norton’ were found resistant to GVCV in grafting assay (Guo et al., 2014; Wenping
Qiu, Missouri State University, USA, personal communication). The reason why these
two cultivars are resistant to this virus is not clear yet.

Information on the pathogenesis of GVCV or the function of the proteins coded
by the virus is scarce. In a study published in 2013, the RT and ZF regions amplified
from 13 vines collected from five grape varieties were used to analyze the selection
pressure on these regions. The results show that these regions were subjected to purifying
selection which might be constrained to retain function since these regions are essential
for virus replication (Guo et al., 2014). The same study found that the petiole contains the
highest titer of GVCV, indicating that phloem tissues are the main sites for the virus multiplication. Later in 2016, Zhang et al. found that the GVCV DNA fragment between nt 7,332 and nt 7,672 directs the transcription of RNA molecules. The promoter activity is stronger than the activity of CaMV 35S promoter. The transcription initiated from nt 7,571 and terminated at nt 7,676, generating a terminally redundant RNA, just as in other caulimoviruses (Zhang et al., 2015).

Since GVCV was detected mostly in the cultivated grapevines along edges in commercial vineyard, the survey of this virus was expanded to wild plants to discover where GVCV originates. In 2014, GVCV was discovered in two wild V. rupestris vines, which show translucent vein clearing symptoms on the young leaves and vein necrosis on the mature leaves. The isolates obtained from the two infected vines were named as GVCV-VRU1 and GVCV-VRU2, with genome of 7,755 and 7,725 bp in length, respectively (Beach et al., 2016). Later in 2015, mild GVCV-associated symptoms were observed on Ampelopsis cordata vines. GVCV was detected in A cordata, which is also in the Vitaceae family, as cultivated and native Vitis species. New isolates were named as GVCV-AMP1 and GVCV-AMP2.

**Overall Objectives and Major Discoveries of This Research.** The genetic diversity of badnaviruses is high in some species such as Cacao swollen shoot virus (CSSV) and BSV (Muller & Sackey, 2005; James et al., 2011), which complicates the diagnostics of these diseases. Therefore, the more isolates being sequenced, the more conserved region could be analyzed to improve the diagnosis of these viral disease, including GVCV. In this study, two new isolates of GVCV, one identified from Vitis genus and the other from Ampelopsis genus, were sequenced. The genetic diversity was
analyzed among these two and the five additional known isolates. The conserved domains and the structure of the leader RNA were predicted to infer the function of different regions of GVCV genome. The phylogenetic tree was predicted to analyze the evolutionary relationship among different isolates. In addition, it was also the first time that two GVCV isolates were found to be identical and that infected vines of two genera growing so closely, shedding light to the spread pattern of this virus.

Materials and Methods

Materials. Leaves of plant sample ‘16-30Cha’ were collected from a symptomatic Chardonel vine at a vineyard in Coffman, MO. Leaves of sample ‘16-32Amp’ were collected from a wild Ampelopsis cordata vine, which was located 10 feet away from the 16-30Cha vine. The ‘16-32Amp’ plant was propagated and kept in the greenhouse at the Missouri State Fruit Experiment Station, Mountain Grove, Missouri, USA.

DNA Extraction and Triplex PCR of Detecting GVCV. Plant leaves were weighed at 20-30 mg and processed following the protocol of SYNERGY™ 2.0 plant DNA Extraction kit (OPS Diagnostics, Lebanon, NJ.) to extract the total DNA from the samples. The DNA was eluted in 35 μl autoclaved deionized water. For large output of DNA, the CTAB method (Porebski et al., 1997) was used to extract the plant DNA from 130 mg grapevine leaves tissue, and then the QIAGEN QIAprep Spin Miniprep Kit was used to purify the circular viral DNA. The DNA was adjusted to 10 ng/μl as template for subsequent PCR assay (Austin, 2017).
The presence of viral DNA in the plant was tested via Triplex PCR as described in previous work (Austin, 2017). The plant 16S rDNA was amplified to check the quality of the plant total DNA with the forward primer (5’-TGCTTAACACATGCAAGTCGGA-3’) and reverse primer (5’-AGCCGTTTCCAGCTGTTGTTC-3’) (Krenz et al., 2014). The reagents mixture for triplex PCR assay and the PCR protocol were used as described by Austin (Austin, 2017). After the triplex PCR, the PCR products were electrophoresed in 1X Tris-borate EDTA buffer for 40 minutes. The gel was then visualized under UV light (Austin, 2017).

**Cloning of the Viral Genome and Sequencing Strategy.** Based on the reference genome, GVCV-CHA, published on the GeneBank (Accession number NC_015784), three pairs of primers were used to amplify three overlapping fragments of genomic DNA by PCR to ensure that the complete genome was covered and sequenced (Beach et al., 2016), as shown in Figure 5. The primers used for amplifying these fragments were shown in Table 1 as described in two previous projects of our group (Beach et al., 2016; Austin, 2017). Plantinum Taq DNA Polymerase with high fidelity was used for amplification. The PCR reagent mixture was shown in Table 2 following the protocol. The PCR procedure, which is similar to the one described in the previous section (Beach et al., 2016), is shown in Table 3.

After the DNA fragments were visualized under the UV light, bands containing PCR products were sliced by a clean blade from the gel. The DNA was extracted using the Qiagen MinElute Gel Extraction kit by following the protocol. The DNA fragments were ligated onto the pCR™8/GW/TOPO® vector via TA Cloning by using the Gateway TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) (Beach et al., 2016). The recombinant
vectors were transformed into One Shot™ TOP10 Chemically Competent *E. coli* cell (Invitrogen) by heat-shock method. The transformed competent *E. coli* were recovered in the S.O.C. medium (Invitrogen) for one hour and then were spread onto the Luria-Bertani (LB) agar medium plates with 100 μl/ml spectinomycin in a 37°C incubator for 15-17 hours (Beach *et al*., 2016). Three individual colonies were selected and cultured in liquid LB medium supplemented with the same concentration of spectinomycin in a 37°C shaker overnight at 250 rpm (Beach *et al*., 2016). Plasmid DNA was extracted using QIAprep® Spin Miniprep Kit (QIAGEN).

After confirmed by PCR using vector- or GVCV- specific primers or restriction enzyme digestion, the recombinant vectors were sequenced at Nevada Genomics Center, University of Nevada in Reno, Nevada. The vector specific primers, GW1 and GW2, were used for sequencing the terminals of the inserted viral genome fragments (Beach *et al*., 2016). Then based on the returned sequences and the reference genome, GVCV specific primers were used for sequencing the remaining sequence between the inserted sites until the entire inserts were sequenced (Beach *et al*., 2016). The sequences were analyzed and assembled using the codon code aligner software (CodonCode Corporation Centerville, MA). The three overlapping fragments were then assembled into completed circular GVCV genome sequences (Petersen, 2016).

**Genome Sequence Analysis and Genomic Feature Prediction.** Pairwise sequence alignment and identity were analyzed using global sequence alignment on the EMBOSS Needle web server(https://www.ebi.ac.uk/Tools/psa/emboss_needle/). Multiple sequences alignments of entire genome sequences were performed by ClustalW in MEGA7 package (Kumar *et al*., 2016). Phylogenetic relationship among the
7 GVCV isolates was analyzed using Neighbor-Joining (NJ) method, with 1000 bootstrap replicates, following p-distance method calculating percentages of different nucleotide sites between two isolates to infer the evolutionary distances (Nei & Kumar, 2000).

Seven complete GVCV genome sequences were submitted to search for open reading frames (ORF) in NCBI ORF finder web server (https://www.ncbi.nlm.nih.gov/orffinder/). Conserved domains in each isolate were predicted with the NCBI CD-search (Marchler-Bauer et al., 2016). Genomic features were visualized using the SnapGene Viewer program (GSL Biotech, Chicago, IL). Leader sequence in the pre-genomic RNA of GVCV isolates were predicted and trimmed from the homologous region of GVCV-CHA, which is between the transcription initiation site and the fifth codon of ORFI (Zhang, 2016). Secondary structure of RNA was predicted by the mfold program (Zuker, 2003).

Results

**Symptoms on Two Infected Vines.** The symptoms observed on the cultivated ‘16-30Cha’ included severe vein clearing, curling on the leaf margins, sparse berry sets, and dwarfism. All these symptoms have been previously linked to the presence of GVCV (Zhang et al., 2011). On the other hand, the *Ampelopsis cordata* vine ‘16-32Amp’, which was collected in its native habitat, 10 feet away from the location of ‘16-30Cha’, show mild vein-clearing symptom only (Figure 6).

**Comparison of the Two New GVCV Isolates.** After the three fragments from each isolate were sequenced separately, these sequences were assembled into two complete genome sequences. The isolate from the ‘16-30Cha’ vine was named GVCV-
CHA2. The isolate from *Ampelopsis cordata* vine ‘16-32Amp’ was named GVCV-AMP3. The two isolates have the same length of genome, which is 7,742 base pair (bp). Basic information about the two isolates together with known 5 isolates, such as the start and stop sites of each ORF, and NCBI accession numbers were shown in the Table 4.

As shown in the Table 4, ORFs of the two new isolates start and end at the same position. In the alignment of the entire genome sequences, only 17 out of the 7,742 bp are non-homologous, indicating that the two isolates are 99.8% identical at nucleotide sequences. After further alignment between every pair of ORF sequences, two of the divergent nucleotides fell within the ORF I region and did not result in changes at the amino acid sequence. Twelve of the divergent nucleotides were within the ORF III region, but caused only two amino acid substitutions, neither of which was within the conserved domains. In addition, there were 3 divergent nucleotides in the intergenic region between ORF I and ORF III.

**Genome Analysis among Seven GVCV Isolates.** Pairwise nucleotide identity of intergenic region (IGR) and amino acid identity (in bold) of three ORFs among seven GVCV isolates are shown in Table 5 to Table 8, respectively. The ranges of these values were furtherly analyzed in Figure 7 and Figure 8. In Figure 7, it is shown that the pairwise nucleotide identity varied from 89.4 to 99.7% (IGR region), from 92.2 to 99.7% (ORF I), from 83.4% to 100% (ORF II), from 91.8 to 99.8% (ORF III). ORF II is the most divergent region with the lowest identity.

The ranges of amino acid identity values in ORFs region were analyzed and were shown in Figure 8. It is shown that the ORF III is the most conserved region with 96% to
99.9% identity and ORF II is the most divergent region with identity ranging from 86.2 to 100%.

**Annotation and Prediction of Genomic Features.** As shown in Figure 9, the reference genome GVCV-CHA was annotated with three ORFs and promoter region (Zhang et al., 2015). The conserved domains were searched and annotated in the Figure 10. Except for GVCV-CHA, other six isolates contain the same domains in same order. The representative isolate GVCV-AMP3 was annotated in Figure 10, together with the GVCV-CHA. As shown in this figure, all the isolates contain Reverse transcriptase (RTs), RNase H, Zinc Knuckle, AIR1, and DUF1319 domain. According to the annotation by Conserved Domain Database, the RT catalyzes the DNA replication from an RNA template; The RNase H is responsible for the degradation of the RNA strand of an RNA/DNA hybrid; Zinc Knuckle is a zinc binding motif, which is mostly present in retroviral gag proteins (nucleocapsid); the AIR1 superfamily refers to methyltransferase-interacting protein, which plays a role in posttranslational modification, protein turnover, chaperons, or intercellular trafficking and secretion; The motif DUF1319 is restricted to badnaviruses with unknown function; The SMC_N domain is predicted to be in GVCV-CHA exclusively and is thought to be involved in chromatin and DNA dynamics, DNA metabolism and recombination (Marchler-Bauer & Bryant, 2004; Marchler-Bauer et al., 2010; Marchler-Bauer et al., 2014; Marchler-Bauer et al., 2016).

**Phylogenetic Relationship among Seven GVCV Isolates.** The phylogenetic relationship among seven GVCV isolates were analyzed and shown in Figure 11. The *Gooseberry vein banding virus* (GVBaV, GenBank accession number NC_018105.1) was set as outgroup for constructing the rooted tree. The predicted evolutionary distance
between GVCV-CHA2 and GVCV-AMP3 is the shortest. The isolate GVCV-AMP1 was shown closest to the two new GVCV isolates. The two isolates identified in wild Vitis vine were grouped in the same branch.

**Prediction of the Leader Sequence in the Pre-genomic RNA of Seven GVCV Isolates.** The leader sequence in the pre-genomic RNA of CaMV was studied to form an intricate stem-loop structure, which facilitates the ribosomal to switch and reinitiate the translation of the first true long open reading frame (Ryabova & Hohn, 2000). GVCV-CHA was thought to be able to form such a stem-loop structure, thus the ribosome would terminate the translation of the sORFI and be shunted to the start codon of true ORF I (Zhang, 2016). To verify whether the leader sequence of each GVCV isolate would fold into a stem-loop structure, the homologous region of the leader sequence in GVCV (from the transcription initiation site to the first five codons of ORF I) was trimmed and submitted into mfold program for prediction of secondary mRNA structure. As shown in Figure 12, all the leader sequences from each isolate were predicted to form a similar secondary structure except for the GVCV-VRU1. Generally, it seems likely that the ribosome would be blocked by the several stem structures in the downstream of the transcription initiation site and would switch and reinitiate the translation of ORF I a few nucleotides away.

**Discussion**

As discussed in one review article, the viral infections in wild plants are frequent but usually undocumented (Prendeville et al., 2012). These features indicate that the genetic diversity and environmental complexity make it possible that wild plants could
evolve to tolerate or resist to virus. Observations from this study showed that the infected *Ampelopsis* usually exhibits mild symptoms. Only slight vein clearing symptoms were observed. The infected wild *V. rupestris* vines show distinct vein clearing in the young leaf and necrotic spots on the mature leaves (Beach *et al*., 2016). As described in the previous section, the infected cultivated vines show severe translucent vein clearing, diminished fruit sets and weak growth. However, cultivar ‘Chambourcin’ is an exception to this proposition, since ‘Chambourcin’ was found to be resistant to this virus (Guo *et al*., 2014). It is not clear why this *Vitis* interspecific cross from France is resistant to GVCV. It is also unclear why cultivar ‘Norton’, which is native to North America (http://ngr.ucdavis.edu/) is also resistant (Wenping Qiu, Missouri State University, USA, personal communication). Guo suggested that resistance of ‘Chambourcin’ might be caused by the barriers to transmission of GVCV or inhibition to the replication of the viral genome. A further genetic analysis with the two resistant grapes might provide valuable insights into the virus-resistance mechanism in grapevine (Guo *et al*., 2014).

This study showed that two nearly identical isolates were discovered in plants of different genera within a ten-foot distance. Out of 7,742 nt, only 17 nucleotide sites were found non-homologous between these two isolates. Among seven GVCV isolates, these two are separated by the least evolutionary divergence. Moreover, in a single GVCV infected vine, more than 60 single nucleotide polymorphisms (SNPs) were found, which is larger than the non-homologous sites we found between the two isolates. We think they were derived likely from a single source, as a result of generating viral quasispecies for a virus to adapt to a new host (Domingo *et al*., 2012; Howard & Qiu, 2017). Therefore, we consider them to be the different isolates of the same strain. Furthermore, it is the first
time we found that this GVCV strain can infect plants in different genera. Although, conclusive evidence has yet to be provided, GVCV virus is likely to spread from the wild grapevine relatives to the cultivated grape in our vineyard (Petersen, 2016). Therefore, it is reasonable to recommend that grape growers remove the wild vines near their vineyards.

Studying the pattern of GVCV infection depends on accurate detection method (Prendeville et al., 2012). Traditionally, ELISA, RT-PCR or RT-qPCR were used to detect viruses in grapevine (Fajardo et al., 2017). Since GVCV is a circular double-stranded DNA virus, duplex or triplex PCR primers were designed to detect the presence of the GVCV (Beach et al., 2016; Austin, 2017). Based on the seven GVCV genomes to date, the conserved domains were analyzed. In addition, the conserved regions could be used to design primers for more accurate virus diagnosis.

Generally, the seven isolates were clustered into two groups and they were not grouped into clades according to their geographic locations. These findings are consistent with the results in Guo's study, which suggested that RT and ZF region from 13 isolates collected in three states were not clustered by location or host variety (Guo et al., 2014).

Construction of an infectious clone is essential for verifying the cause of a viral disease. Also, it is an important tool for studying the function of a protein encoded by a virus (Zhang, 2016). Infectious clones have been created for Tomato yellow leaf curl virus (Salati et al., 2002), CaMV (Gardner et al., 1981), Citrus yellow mosaic virus (CYMBV) (Huang & Hartung, 2001), and ComYMV (Medberry et al., 1990). To prove that the GVCV is the causal agent of the grapevine vein clearing disease, the terminally redundant infectious clone was constructed by Zhang (Zhang, 2016). The infectious clone
was infiltrated into *Nicotiana benthamiana* leaves. Consequently, the vein clearing symptom together with flexuous rod virion were observed on the infected *N. benthamiana* leaves (Zhang, 2016).

Thus far, no experimental evidence has been generated for studying the function of proteins coded by GVCV. In the previous section, the conserved domain and three long ORFs were annotated. ORF I encodes a putative protein with 23 kDa. This protein was searched in the CDD and was identified to be a badnavirus-specific DUF1319 (CDD Accession: cl06184) (Sether *et al*., 2012; Chiumenti *et al*., 2016). In the middle part of ORF III, the CCHC-type “zinc finger” and “AIR1” were predicted to be in the same site. The former domains are thought to be from retroviral gag proteins, the prototype of which are gag proteins of human immunodeficiency virus (HIV), located in the inner membrane of the HIV nucleocapsid (Wehrly & Chesebro, 1997). It was also suggested to be involved in binding with single-stranded nucleic acids (Summers, 1991). The “AIR1” domains are part of arginine methyltransferase-interacting protein, which contain ring Zn-finger and are suggested to be involved in posttranslational modification (Marchler-Bauer *et al*., 2016). The remaining RT and RNase_H domain were predicted with low E-value, suggesting the promising role in synthesis of viral genome DNA from a pre-genomic RNA template (Guo *et al*., 2014).

Although the secondary structure of the leader RNA of GVCV-VRU1 differs from that of other isolates, all these isolates were predicted to form a stable large stem-loop and several small hairpin structures. Considering that the intergenic region is also very conserved with more than 90% identity in nucleotide sequence, leader RNA in different isolates might still contribute the same function in translation, such as leading
the ribosome reinitiate to translate true ORFI (Ryabova & Hohn, 2000), or serving as an internal ribosome entry site (Basso et al., 1994).

In conclusion, two new isolates of GVCV were identified and sequenced for obtaining the entire genome sequences. Results show that they belong to same strain, which indicates same GVCV strain spreads cross plants in different genera. Genetic diversity among total seven isolates was analyzed. Furthermore, conserved domains, the secondary structure of leader RNA, and the phylogenetic relationship among seven isolates were predicted to infer the function of different regions of the GVCV genome and to improve the diagnosis of GVCV.
Table 1. A list of primers used for amplifying three overlapping fragments of GVCV genome.

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Primer</th>
<th>Sequence (5' to 3' direction)</th>
<th>Tm/ ℃</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>988F</td>
<td>acctaagccgattgaacagcag</td>
<td>60.4</td>
</tr>
<tr>
<td></td>
<td>4387R</td>
<td>ctctctctcagaaattgagcagat</td>
<td>61.3</td>
</tr>
<tr>
<td>II</td>
<td>4142F</td>
<td>gtaaacctcatgactctcatg</td>
<td>58.7</td>
</tr>
<tr>
<td></td>
<td>6795R</td>
<td>gctggcgtaagcagactctcagatc</td>
<td>59.4</td>
</tr>
<tr>
<td>III</td>
<td>6666F</td>
<td>acctctccacccacgcagttctc</td>
<td>66.3</td>
</tr>
<tr>
<td></td>
<td>1935R</td>
<td>tcggtgtacactgtttact</td>
<td>58.7</td>
</tr>
</tbody>
</table>
Table 2. Components of PCR for amplifying GVCV fragments with Platinum Taq DNA Polymerase.

<table>
<thead>
<tr>
<th>Component</th>
<th>25-μl rxn</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autoclaved, distilled water</td>
<td>18.9 μl</td>
<td>-</td>
</tr>
<tr>
<td>10× High Fidelity PCR Buffer</td>
<td>2.5 μl</td>
<td>1×</td>
</tr>
<tr>
<td>50 mM MgSO4</td>
<td>1 μl</td>
<td>2.0 mM</td>
</tr>
<tr>
<td>10 mM dNTP Mix</td>
<td>0.5 μl</td>
<td>0.2 mM each</td>
</tr>
<tr>
<td>10 μM forward primer</td>
<td>0.5 μl</td>
<td>0.2 μM</td>
</tr>
<tr>
<td>10 μM reverse primer</td>
<td>0.5 μl</td>
<td>0.2 μM</td>
</tr>
<tr>
<td>10 ng/μl Template DNA</td>
<td>1 μl</td>
<td>0.4 ng/μl</td>
</tr>
<tr>
<td>Platinum™ Taq DNA Polymerase High Fidelity</td>
<td>0.1 μl</td>
<td>0.5 U/rxn</td>
</tr>
</tbody>
</table>
Table 3. Thermocycler protocol for amplifying GVCV fragments with Plantinum *Taq* DNA Polymerase.

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
<th>Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>94 °C</td>
<td>2 minutes</td>
<td>1</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94 °C</td>
<td>30 seconds</td>
<td></td>
</tr>
<tr>
<td>Annealing</td>
<td>55 °C</td>
<td>40 seconds</td>
<td>35</td>
</tr>
<tr>
<td>Extension</td>
<td>68 °C</td>
<td>4 minutes</td>
<td></td>
</tr>
<tr>
<td>Final extension</td>
<td>68 °C</td>
<td>10 minutes</td>
<td>1</td>
</tr>
<tr>
<td>Hold</td>
<td>4 °C</td>
<td>Indefinitely</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Comparative analysis of whole genome, intergenic region (IGR) and major open reading frames (ORFs) in seven GVCV isolates.

<table>
<thead>
<tr>
<th>GVCV isolate</th>
<th>Genome Length</th>
<th>IGR Start-Stop (Length)</th>
<th>ORF1 Start-Stop (Length)</th>
<th>ORFII Start-Stop (Length)</th>
<th>ORFIII Start-Stop (Length)</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHA</td>
<td>7,753 nt</td>
<td>7,321-7,753; 1-484 (917 nt)</td>
<td>485-1,111 (627 nt)</td>
<td>1,112-1,495 (384 nt)</td>
<td>1,495-7,320 (5,826 nt)</td>
<td>JF301669.2</td>
</tr>
<tr>
<td>VRU1</td>
<td>7,755 nt</td>
<td>7,332-7,755; 1-483 (907 nt)</td>
<td>484-1,110 (627 nt)</td>
<td>1,111-1,503 (393 nt)</td>
<td>1,503-7,331 (5,829 nt)</td>
<td>KJ725346.1</td>
</tr>
<tr>
<td>VRU2</td>
<td>7,726 nt</td>
<td>7,317-7,726; 1-474 (884 nt)</td>
<td>475-1,104 (630 nt)</td>
<td>1,105-1,488 (384 nt)</td>
<td>1,488-7,316 (5,829 nt)</td>
<td>KT907478.1</td>
</tr>
<tr>
<td>AMP1</td>
<td>7,749 nt</td>
<td>7,336-7,749; 1-481 (895 nt)</td>
<td>482-1,108 (627 nt)</td>
<td>1,109-1,501 (393 nt)</td>
<td>1,501-7,335 (5,835 nt)</td>
<td>KX610316.1</td>
</tr>
<tr>
<td>AMP2</td>
<td>7,765 nt</td>
<td>7,341-7,765; 1-486 (911 nt)</td>
<td>487-1,116 (630 nt)</td>
<td>1,117-1,509 (393 nt)</td>
<td>1,509-7,340 (5,832 nt)</td>
<td>KX610317.1</td>
</tr>
<tr>
<td>AMP3</td>
<td>7,742 nt</td>
<td>7,326-7,742; 1-483 (899 nt)</td>
<td>484-1,110 (627 nt)</td>
<td>1,111-1,494 (384 nt)</td>
<td>1,494-7,325 (5,832 nt)</td>
<td></td>
</tr>
<tr>
<td>CHA2</td>
<td>7,742 nt</td>
<td>7,326-7,742; 1-483 (899 nt)</td>
<td>484-1,110 (627 nt)</td>
<td>1,111-1,494 (384 nt)</td>
<td>1,494-7,325 (5,832 nt)</td>
<td></td>
</tr>
</tbody>
</table>
Table 5. Nucleotide identity of intergenic region (IGR) among seven *Grapevine vein clearing virus* (GVCV) isolates.

<table>
<thead>
<tr>
<th>IGR</th>
<th>Isolate</th>
<th>CHA</th>
<th>VRU1</th>
<th>VRU2</th>
<th>AMP1</th>
<th>AMP2</th>
<th>AMP3</th>
<th>CHA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHA</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VRU1</td>
<td>89.8</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VRU2</td>
<td>89.4</td>
<td>90.8</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMP1</td>
<td>93.1</td>
<td>92.5</td>
<td>91.1</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMP2</td>
<td>93.5</td>
<td>91.1</td>
<td>90.3</td>
<td>93.6</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMP3</td>
<td>93.0</td>
<td>93.0</td>
<td>91.2</td>
<td>96.4</td>
<td>93.7</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHA2</td>
<td>93.2</td>
<td>93.1</td>
<td>91.4</td>
<td>96.6</td>
<td>93.8</td>
<td>99.7</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
Table 6. Nucleotide identity and amino acid identity (in bold) of ORF I among seven *Grapevine vein clearing virus* (GVCV) isolates.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>CHA</th>
<th>VRU1</th>
<th>VRU2</th>
<th>AMP1</th>
<th>AMP2</th>
<th>AMP3</th>
<th>CHA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHA</td>
<td>-</td>
<td>97.6</td>
<td>95.2</td>
<td>97.1</td>
<td>95.2</td>
<td>97.1</td>
<td>97.1</td>
</tr>
<tr>
<td>VRU1</td>
<td>94.3</td>
<td>-</td>
<td>95.7</td>
<td>96.2</td>
<td>95.7</td>
<td>96.6</td>
<td>96.6</td>
</tr>
<tr>
<td>VRU2</td>
<td>92.4</td>
<td>94.3</td>
<td>-</td>
<td>93.8</td>
<td>95.2</td>
<td>95.2</td>
<td>95.2</td>
</tr>
<tr>
<td>AMP1</td>
<td>94.6</td>
<td>92.3</td>
<td>92.4</td>
<td>-</td>
<td>95.7</td>
<td>95.7</td>
<td>95.7</td>
</tr>
<tr>
<td>AMP2</td>
<td>93.2</td>
<td>93.3</td>
<td>93.3</td>
<td>92.2</td>
<td>-</td>
<td>94.3</td>
<td>94.3</td>
</tr>
<tr>
<td>AMP3</td>
<td>95.4</td>
<td>94.6</td>
<td>93.5</td>
<td>93.5</td>
<td>93.3</td>
<td>-</td>
<td>100</td>
</tr>
<tr>
<td>CHA2</td>
<td>95.4</td>
<td>94.6</td>
<td>93.5</td>
<td>93.5</td>
<td>93.3</td>
<td>99.7</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 7. Nucleotide identity and amino acid identity (in bold) of ORF II among seven *Grapevine vein clearing virus* (GVCV) isolates.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>CHA</th>
<th>VRU1</th>
<th>VRU2</th>
<th>AMP1</th>
<th>AMP2</th>
<th>AMP3</th>
<th>CHA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHA</td>
<td>-</td>
<td>86.2</td>
<td>93.7</td>
<td>88.5</td>
<td>90.8</td>
<td>94.5</td>
<td>94.5</td>
</tr>
<tr>
<td>VRU1</td>
<td>83.9</td>
<td>-</td>
<td>90.8</td>
<td>93.1</td>
<td>94.6</td>
<td>90.8</td>
<td>90.8</td>
</tr>
<tr>
<td>VRU2</td>
<td>88.9</td>
<td>88.6</td>
<td>-</td>
<td>92.3</td>
<td>91.5</td>
<td>99.2</td>
<td>99.2</td>
</tr>
<tr>
<td>AMP1</td>
<td>83.4</td>
<td>88.5</td>
<td>89.3</td>
<td>-</td>
<td>93.8</td>
<td>92.3</td>
<td>92.3</td>
</tr>
<tr>
<td>AMP2</td>
<td>88.4</td>
<td>93.2</td>
<td>88.7</td>
<td>89.4</td>
<td>-</td>
<td>91.5</td>
<td>91.5</td>
</tr>
<tr>
<td>AMP3</td>
<td>90.4</td>
<td>88.3</td>
<td>93.8</td>
<td>89.1</td>
<td>88.6</td>
<td>-</td>
<td>100</td>
</tr>
<tr>
<td>CHA2</td>
<td>90.4</td>
<td>88.3</td>
<td>93.8</td>
<td>89.1</td>
<td>88.6</td>
<td>100.0</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 8. Nucleotide identity and amino acid identity (in bold) of ORF III among seven Grapevine vein clearing virus (GVCV) isolates.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>CHA</th>
<th>VRU1</th>
<th>VRU2</th>
<th>AMP1</th>
<th>AMP2</th>
<th>AMP3</th>
<th>CHA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHA</td>
<td>-</td>
<td>96.3</td>
<td>96.7</td>
<td>97.2</td>
<td>97.2</td>
<td>97.5</td>
<td>97.5</td>
</tr>
<tr>
<td>VRU1</td>
<td>91.8</td>
<td>-</td>
<td>96.6</td>
<td>96.1</td>
<td>96.1</td>
<td>96.1</td>
<td>96.0</td>
</tr>
<tr>
<td>VRU2</td>
<td>92.0</td>
<td>93.8</td>
<td>-</td>
<td>96.6</td>
<td>96.0</td>
<td>96.6</td>
<td>96.6</td>
</tr>
<tr>
<td>AMP1</td>
<td>92.4</td>
<td>91.8</td>
<td>92.0</td>
<td>-</td>
<td>96.8</td>
<td>98.3</td>
<td>98.3</td>
</tr>
<tr>
<td>AMP2</td>
<td>92.7</td>
<td>92.1</td>
<td>91.8</td>
<td>92.5</td>
<td>-</td>
<td>96.7</td>
<td>96.7</td>
</tr>
<tr>
<td>AMP3</td>
<td>93.0</td>
<td>92.1</td>
<td>92.0</td>
<td>94.1</td>
<td>92.9</td>
<td>-</td>
<td>99.9</td>
</tr>
<tr>
<td>CHA2</td>
<td>92.9</td>
<td>92.1</td>
<td>92.1</td>
<td>94.1</td>
<td>92.9</td>
<td>99.8</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 5. Amplification and assembling strategy of GVCV genome sequences. The genome is divided into three overlapping fragments, with start to end site shown in the figure.
Figure 6. GVCV associated symptoms on the two infected vines. (A) AMP3 vine exhibiting mild vein clearing. (B) CHA2 vine exhibiting severe vein clearing and leaf deformation and with curling of leaf margins.
Figure 7. Pairwise nucleotide identity analysis of IGR and three ORF regions among seven GVCV isolates.
Figure 8. Pairwise amino acid identity analysis of three ORFs among seven GVCV isolates.
Figure 9. GVCV genome structure. The promoter region and three ORFs were labeled.
Figure 10. Graphical representation of the conserved domains in two GVCV isolates.
Figure 11. Phylogenetic relationship among seven GVCV isolates. Validation of branches was performed by 1000 bootstrap replicates. The tree is built with branch lengths (beside the branches) predicting evolutionary distance calculated using p-distance method. The *Gooseberry vein banding virus* (GVBaV, GenBank accession number NC_018105.1) was set as outgroup for constructing the rooted tree.
Figure 12. Predicted structure of leader RNA using the program mFold. The blue arrows indicate the position of transcription initiation site. The red arrows indicate the position of the fifth codon in the ORF I. The direction of the arrows shows the direction of the translation.
References


Elad Y, 1997. Responses of plants to infection by Botrytis cinerea and novel means involved in reducing their susceptibility to infection. *Biological Reviews* 72, 381-422.


