Investigating the Gene Expression Strategy of Grapevine Red Blotch Virus

Honors Thesis

Biology & Society Honors Program

Cornell University

by

Harris Liou

May 2018

Mentor: Jeremy R. Thompson

Abstract

Grapevine red blotch virus (GRBV), a geminivirus, is the recently identified causative agent of the emerging grapevine disease known as red blotch. The monopartite circular ssDNA GRBV genome consists of genome-sense open reading frames (ORFs) (V1, V2 and V3), and complementary sense ORFs (C1, C2, and C3). Based on sequence comparisons to related viruses, C1 is hypothesized to be spliced with C2 to give rise to a C1-2 mRNA transcript. Recent unpublished data provides mRNA evidence of a possible seventh ORF referred to as V0 as well as evidence of V0 splicing with V2 to form a V0-2 mRNA transcript. So far, none of the predicted protein products have been detected, and gene expression mechanisms are primarily deduced from comparative sequence analyses. Thus, one aim of this study was to investigate the protein expression profile of GRBV in part by analyzing infected grapevine protein extracts with tandem mass spectrometry. Current GRBV diagnostic resources all depend on enzymatic amplification of viral DNA sequences, though false-negative results have been problematic. Another aim of the study was thus to determine the relative translation levels of the ORFs by expressing them in *Nicotiana benthamiana* and analyzing protein products by western blot in order to identify potential diagnostic protein biomarkers. Additionally, expressing the ORFs in *N. benthamiana* generated mRNA transcripts that were analyzed to test the hypothesized C1-2 and V0-2 splicing mechanisms. C3 and V2 proteins were detected by tandem mass-spectrometry in infected grapevine. Furthermore, C3 was found to have the highest translation level, followed by V2, and then by C1-2; other proteins were not translated at levels high enough for detection. Considering several factors, V2 appears to be the best candidate for a diagnostic biomarker against which antibodies can be developed. Lastly, PCR analysis of mRNA transcripts provides evidence for the hypothesized C1-2 and V0-2 splicing mechanisms.

Introduction

Plant diseases are estimated to cost $33 billion in damage to the United States economy annually, causing reduction of crop yield and necessitating costly management efforts (Madden & Wheelis, 2003). Modern-day agricultural practices and climate change have driven the emergence of new plant viruses that threaten plants important to human welfare. Viruses are estimated to be responsible for 47% of emerging infectious diseases affecting plants, more than any other type of infectious agent (Anderson et al., 2004). Particularly, around 80 different virus species belonging to nearly 30 different genera are documented to infect grapevine (*Vitis* spp.), and the number is rising (Martelli, 2017).

Grapevine, a fruit crop widely grown around the world, is a deciduous woody perennial vine. Currently, the most cultivated grapevine species are the Eurasian *V. vinifera*, the Northeastern American *V. labrusca*, the Southeastern United States *V. rotundifolia*, the Asian *V. amurensis*, as well as several hybrids. Grapes are consumed as fresh fruits, wines, juices, raisins, jellies, vinegars, and seed oils. The grape industries of the United States, one of the top grape producing nations under only China and Italy, are fast-growing areas of production with a commercial presence in most states. In 2010, the total value of grape crop in the United states was $3.47 billion (Naidu et al., 2014).

Viruses with circular DNA genomes interact with all domains of life, infecting humans, other animals, bacteria, archaea, and plants. Recent advances in rolling-circle amplification techniques have allowed for the rapid discovery of many novel circular DNA viruses without *a priori* sequence information (Rector et al., 2004). One such family of viruses that infects grapevine and other crops is *Geminiviridae*, named for their twinned quasi-icosahedral shape (Fig.1) Geminiviruses have small, circular, single-stranded deoxyribonucleic acid (ssDNA), ambisense

genomes and constitute one of the largest and most economically damaging plant virus families (Fondong, 2013). Nine genera have thus far been identified within the family, including the newly established *Grablovirus* originally conceived to categorize the species *Grapevine red blotch virus* (GRBV) (Varsani et al., 2017).

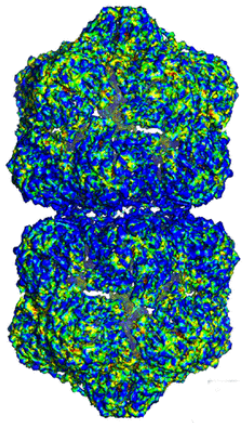


Figure 1. Geminivirus structure. This electron cryomicroscopy image of African cassava mosaic virus exhibits the twinned quasi-icosahedral shape characteristic of geminiviruses. Size - 22nm width 38nm length. Taken from (Hipp et al., 2017).

GRBV is the causative agent of the relatively recently recognized grapevine disease, red blotch (Yepes et al., 2018). Exhibiting symptoms similar to that of leafroll disease but testing negative by leafroll diagnostic tools, red blotch was first described in 2008 as it affected a vineyard of *V. vinifera* 'Cabernet Sauvignon' in Oakville, California (Calvi, 2011). Further investigation led to the characterization and identification of the virus now known as GRBV (Krenz et al., 2012; Rwahnih et al., 2013).

Red blotch disease has foliar and fruit symptoms (Fig. 2) in which the onset varies with cultivar, location, and growing season (Cieniewicz et al., 2017). In red-berried cultivars, GRBV-infected grapevine leaves exhibit blotches of shades varying from crimson to purple early in the growing season. Later in the season, most of the leaf blade becomes red and heavily symptomatic

leaves drop off prematurely. In white-berried cultivars, infected grapevine leaves exhibit chlorotic patches that can become necrotic later in the season. Fruit symptoms include delays in ripening, altered fruit juice chemistry indices (including the decrease of total soluble solids from −1° to −4° Brix), and lower anthocyanin contents in berry skin (Calvi, 2011). A study estimated the financial damage of red blotch to range from $2,213 to $68,548 per hectare over a 25-year productive life span of Cabernet Sauvignon and Merlot vineyards, depending on initial infection levels and the price penalty for suboptimal quality fruit (Ricketts et al., 2017). Currently, there is no cure for red blotch disease. A study using economic analyses recommends replacing symptomatic vines with clean vines derived from virus-tested stocks if red blotch incidence is below 30%, and a full vineyard replacement if disease incidence is above 30% (Ricketts et al., 2017).

GRBV is widespread throughout most of the major grape growing regions of the United States as well as in the Canadian provinces of British Columbia and Ontario (Krenz et al., 2014; Poojari et al., 2017). Furthermore, GRBV was reported in a Swiss experimental vineyard in grapevine imported from California (Reynard et al., 2018). The virus was also present in a Korean vineyard with grapevines of unknown origins (Lim et al., 2016). The three-cornered alfalfa treehopper, *Spissistilus festinus* (Fig. 3), has been shown to transmit GRBV from grapevine to grapevine under greenhouse conditions, though its potential as a vineyard GRBV vector has yet to be determined (Bahder et al., 2016). The virus is graft-transmissible, likely also having spread through the dissemination of infected propagation material (Poojari et al., 2013; Rwahnih et al., 2013).

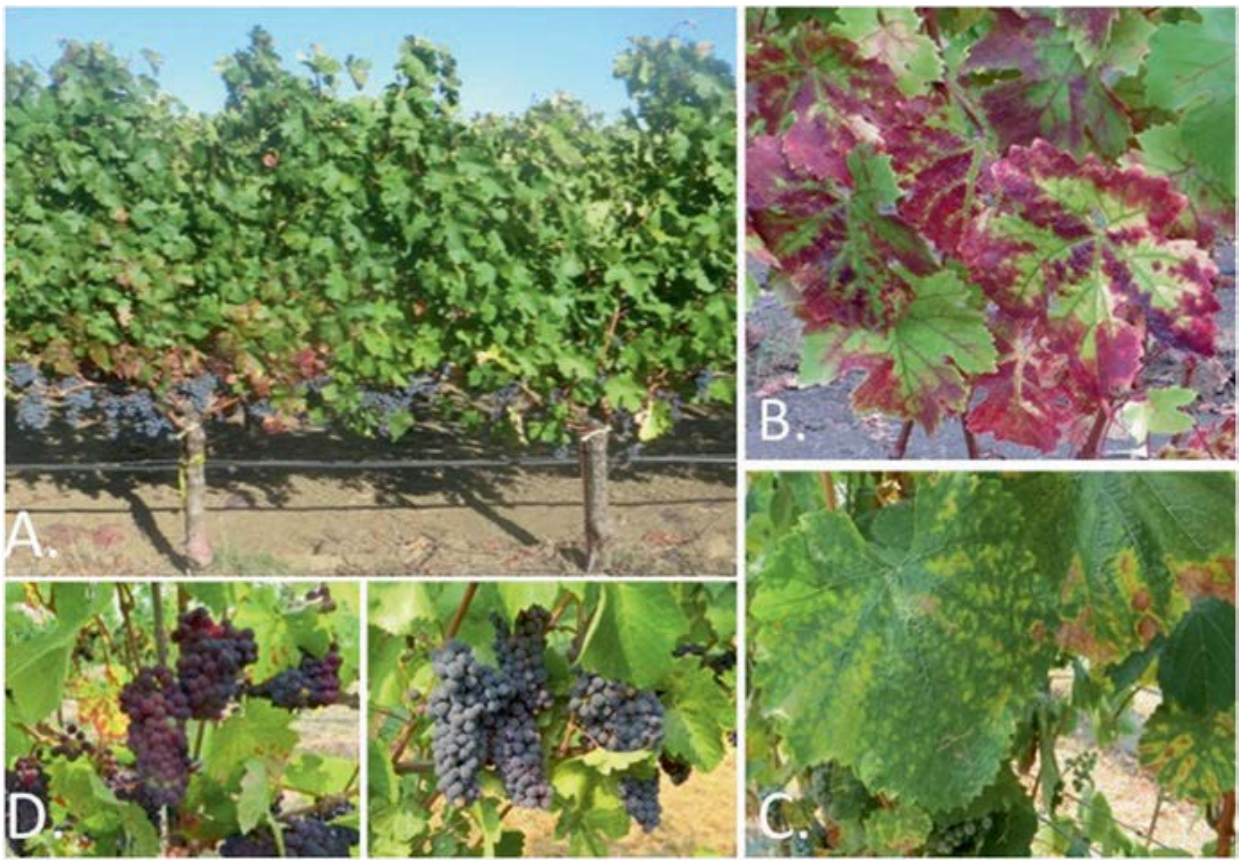


Figure 2. Foliar symptoms of red blotch. (a) at the bottom of the canopy of a diseased (left) compared to a healthy (right) Cabernet franc, close-up of foliar symptoms on (b) Cabernet franc and (c) Chardonnay, and d) fruit symptoms on a diseased (left) compared to a healthy (right) Pinot noir. Taken from (Cieniewicz et al., 2017)

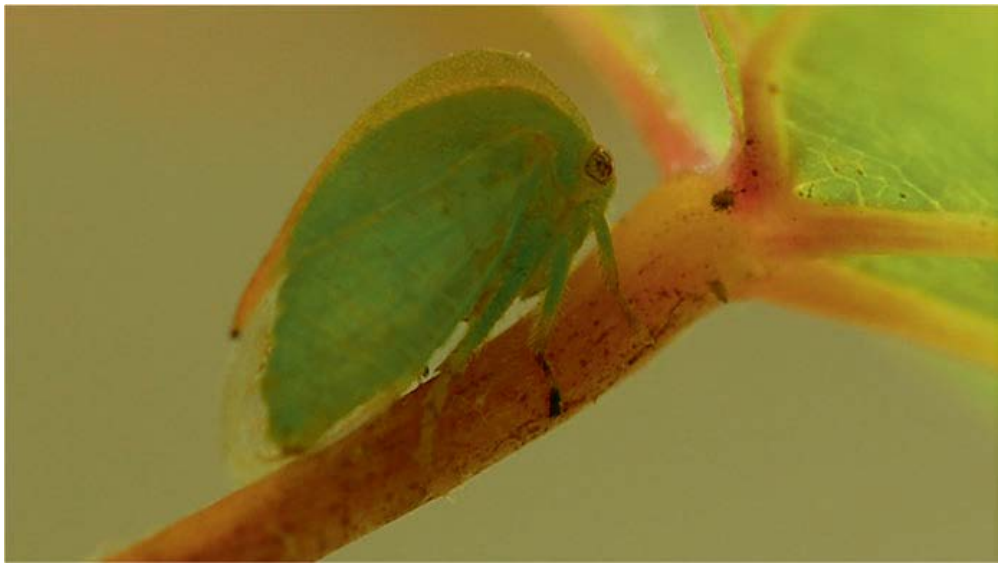


Figure 3. Adult *Spissistilus festinus.* This treehopper rests on a petiole of a *V. vinifera* 'Cabernet Sauvignon' leaf. Taken from (Cieniewicz et al., 2017)

Analysis of the GRBV monopartite circular ssDNA genome consisting of 3206 nucleotides predicts six open reading frames (ORFs) (Rwahnih et al., 2013). The genome consists of a conserved origin of replication, the genome-sense ORFs (V2, V1, and V3), and the complementary sense ORFs (C1, C2, and C3) (Fig. 4). Recent unpublished data provides mRNA evidence of a possible seventh ORF referred to as V0. Based on comparisons of the GRBV genome to those in the *Mastrevirus* genus and other geminiviruses, the complementary sense ORFs are predicted to encode replication-associated proteins while V0, V2 and V3 are hypothesized to encode movement-related proteins (Cieniewicz et al., 2017; Krenz et al., 2014). Specifically, C1 likely encodes the replication protein, RepA, and C1 is likely spliced with C2 to give rise to a C1-2 mRNA transcript encoding another replication protein, Rep. V1 is predicted to code for the coat protein. Recent unpublished work has shown some evidence of V0 splicing with V2 to form a V0-2 mRNA transcript.

Current GRBV detection and diagnostic resources all depend on nucleic acid based amplification of viral DNA sequences, though false-negative results have been problematic (Cieniewicz et al., 2017). In unpublished studies, antibodies have been produced against a number of versions of the most theoretically abundant and conserved viral protein: the GRBV V1 coat protein. However, these antibodies did not exhibit differential reactions when tested against infected and uninfected plant extracts in both western blot assays and an enzyme-linked immunosorbent assay (ELISA). Thus, the main applied aim of this project is to identify GRBV proteins expressed at levels high enough to serve as potential diagnostic biomarkers that can be detected by antibodies.

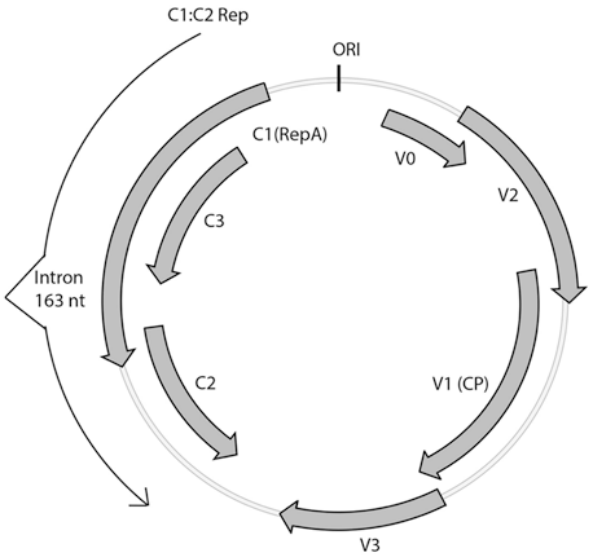


Figure 4. Predicted GRBV genome organization. Thick arrows represent the open reading frames, while the thin arrow represents the C1-2 sequence predicted to be spliced. The conserved origin of replication (ORI) at the top is surrounded by genome-sense ORFs clockwise to the right (V2, V1, and V3) and complementary sense ORFs counterclockwise to the left (C1, C2, and C3). Taken from Cieniewicz et al., 2017.

So far, none of the predicted protein products have been detected in plants, and gene expression mechanisms are primarily deduced from comparative sequence analyses. Thus, another aim of this study was to investigate the protein expression profile of GRBV in part by analyzing infected grapevine protein extracts with tandem mass spectrometry. Additionally, translation levels the predicted ORFs in agroinfiltration assays were to be assessed in order to develop a method of protein expression for downstream antibody production. Furthermore, this study seeks to test the hypothesized splicing of C1 to C2 as well as V0 to V2 to determine if indeed the ORFs give rise to spliced transcripts and their derived protein products.

The results of this work demonstrate that C3 and V2 proteins are present at levels high enough for detection by tandem mass-spectrometry. Furthermore, C3 was found to have the highest translation level, followed by V2, and then by C1-2; other proteins were not translated at

levels high enough for detection. Lastly, PCR analysis of mRNA transcripts provides evidence for the hypothesized C1-2 and V0-2 splicing mechanisms.

Materials and Methods

Grapevine Tissue Protein Extraction

Protein from leaf and petiole of GRBV-infected grapevine (clone GV32 from original plant isolate NY358) was extracted via a trichloroacetic acid (TCA) precipitation protocol adapted from Wang et al. 2006. Approximately 0.25 g of tissue was ground with liquid nitrogen. Approximately 0.1 g of ground sample was then transferred into each of two 5 ml centrifuge tubes along with 5 ml of chilled 1:9 TCA:Acetone +20 mM dithiothreitol (DTT). The tubes were vortexed and incubated with shaking at 4°C for 15 minutes, then centrifuged at 16,000 xG for 15 minutes at 4°C and the supernatant discarded. Pellets were resuspended in 1 ml chilled 80% methanol containing 0.1 M ammonium acetate, centrifuged at 16,000 xG for 15 min at 4°C, and the supernatant discarded. The wash was then repeated. Pellets were resuspended in 1 ml of cold acetone, centrifuged at 16,000 xG for 15 min at 4°C, and the supernatant was discarded. This wash step was then repeated once. Pellets were air dried at room temperature for approximately 5 minutes, resuspended in 2 ml SDS buffer pH 8.0 (0.1 M Tris-HCl, 30% sucrose, 1% SDS, 1 mM EDTA, 1% PVPP, 5% 2-mercaptoethanol) and incubated at room temperature for 1 hour. The tubes were centrifuged at 15,000 xG for 10 min at room temperature, and the supernatant transferred to new 5 ml centrifuge tubes. An equal volume of tris-buffered phenol (pH 8) was added to the tubes, which were then vortexed for 4 minutes. The tubes were centrifuged at 15,000 xG for 5 minutes at room temperature. For each tube, the lower yellow phenol phase was collected and distributed in 500 µl portions into 5 ml centrifuge tubes. Chilled methanol containing 0.1 M ammonium acetate was added to each tube to a final volume of 5 ml. The tubes were vortexed for 30 seconds and incubated at –20°C overnight, after which they were centrifuged at 16,000 xG for 15 minutes at 4°C, and the supernatant discarded. Pellets were resuspended in 1 ml methanol, centrifuged at

16,000 xG for 15 min at 4°C, and the supernatant discarded. This wash step was then repeated. Pellets were resuspended in 1 ml 80% acetone, centrifuged at 16,000 xG for 15 min at 4°C, and the supernatant discarded. This wash step was then repeated. Pellets were stored in 80% acetone at −80°C. This was done for 3 petiole and base leaf samples, which were pooled and designated as "LP protein" (leaf and petiole).

Protein from 0.25 g of GRBV infected grapevine leaf and petiole was prepared using the TCA precipitation method but was resuspended in Laemmli buffer containing 5M urea to address possible issues of protein aggregation during the subsequent gel electrophoresis step. This was done for 50 petiole and base leaf samples, which were pooled together and designated as "LPU protein" (leaf and petiole with urea). Additionally, protein from 0.25 g GRBV infected grapevine cane was extracted with the objective of enriching for phloem tissue where GRBV is hypothesized to predominantly reside. This was done for 5 cane samples, pooled, and designated as "C protein" (cane).

Polyacrylamide gel electrophoresis (PAGE)

Taking the TCA protein samples from storage at -80°C the 80% acetone was removed by centrifugation from the protein pellets, which were resuspended in 200µl of 2x Laemmli sample buffer with 710 mM 2-mercaptoethanol. For only the LPU protein, urea was added to the sample to be solubilized to a concentration of 5M. The resuspended samples were then boiled for 10 minutes after which they were run on Mini-PROTEAN® TGX Any kD™ precast gels, in which 20 µl of sample were loaded. Ten µl of Precision Plus Protein™ Dual Color Standards (BioRad, Hercules, CA) was added in one lane of each gel. Gels were run at 150 V and 250 mA for 45 minutes. Gels were then stained for total protein using Coomassie blue.

Tandem Mass-Spectrometry

For the LP protein, 3 horizontal strips were cut from the stained gels to separate proteins of three discrete protein size ranges based on the size standards' migration: 15-20, 20-25, and 25-37 kDa. Each strip was separately submitted to the tandem mass-spectrometry (MS-MS) facility at the Cornell Proteomics Core Facility (http://www.biotech.cornell.edu/brc/proteomics-mass-spectrometry-facility) to detect predicted GRBV proteins.

For LPU protein, a single strip containing proteins of 15-27 kDa was excised and submitted to the MS-MS facility. For the C protein, 4 separate strips containing proteins of 15-17, 17-19, 19-22, and 22-25 kDa were excised and submitted.

GRBV Open Reading Frame Cloning

DNA cloning of the predicted GRBV open reading frames (ORFs) of V1, V2, V3, C1-2, and C3 was carried out as follows. Individual ORFs were amplified by the polymerase chain reaction (PCR) using as a template a full-length GRBV genome cloned into the plasmid pUC19 (Invitrogen, Carlsbad, CA) with primers (Appendix A) complementary to regions flanking targeted ORFs. The primers were selected so that the amplified ORFs would include three nucleotides upstream of the translation start site. DNA coding for V0-2 was generated in a similar way, but with GRBV virus DNA extracted from the GRBV-infected grapevine line of GV32 as the template due to problems cloning from the pUC19-GRBV clone. In each PCR tube the contents were as follows: 4.0 µl 10X AccuPrime™ PCR Buffer I (Invitrogen, Carlsbad, CA), 1 µl 10 µM forward primer, 1 µl 10 µM reverse primer, 1 µl template DNA, 1 µl AccuPrime™ Taq DNA Polymerase (Thermo Fisher, Waltham, MA). A control was also included in which the template DNA was substituted by water. The PCR program was as follows: 95°C for 3 minutes, followed by 38 cycles

of 94°C for 1 minute, 55°C for 1 minute, 72°C for 1 minute, followed by 72°C for 1 minute, and stored at 4 °C.

Each tube of PCR products was mixed with 5 µl loading dye, loaded onto 1% agarose gels along with BenchTop 1 kb DNA Ladder (Promega, Madison, WI), run at 120 V for 20-30 minutes, and stained with ethidium bromide. The gels were visualized under UV light and the bands containing the expected PCR product (based on size) were cut out. The DNA was then purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Irvine, CA) and cloned into pGEM®-T Easy (Promega, Madison, WI) cloning vector using the Blunt/TA Ligase Master Mix (New England Biolabs, Ipswich, MA) according to manufacturer's directions.

Each plasmid (pGEM®-T Easy vector with GRBV ORF inserts) was transformed into competent *Escherichia coli* (strain DH5α) using the following method. A tube of competent *E. coli* was thawed on ice and mixed gently. 50 µl of *E. coli* was pipetted into a separate microcentrifuge tube, along with 2 µl (~100 ng) of the plasmid. The tube was flicked 5 times to mix then placed on ice for 5 minutes. The tube was then heat shocked at 42°C for 30 seconds and placed on ice for 5 minutes. 950 µl of Lysogeny Broth (LB) medium broth was pipetted into the tube, which was then shaken at 250 rpm at 30°C for 60 minutes. The tube was centrifuged at 1180 xG for 2 minutes. Nine hundred µl of the supernatant was removed, and the pellet resuspended in the remaining supernatant and plated on LB medium agar plates containing ampicillin (100 µl/ml) and coated with 1mg of X-gal (IBI Scientific, Dubuque, IA) dissolved in dimethylformamide. The plate was incubated overnight at 37°C.

For each plate, 4 white colonies were picked and grown in 2 ml of LB medium containing ampicillin (100 µl/ml). Plasmid clones were purified according to the protocol described by Ahn et al. 2000. To verify the presence of the insert in the plasmid, a sample of each plasmid preparation

was digested with *Eco*RI by mixing in a microcentrifuge tube 1 µl *EcoR*I (New England Biolabs, Ipswich, MA), 1 µl CutSmart® buffer, 5 µl midi-prep plasmid product, and 1 µl water. Each tube was incubated for 2 hours at 37°C. Five µl of each digested plasmid mixed with 2 µl of loading dye was loaded onto a 1% agarose gel along with BenchTop 1 kb DNA Ladder, run at 120 V for 20 minutes, and stained with ethidium bromide. The gels were then visualized under UV to verify the presence of bands representing the cloning vector and insert. Samples of minipreps that exhibited the expected *Eco*RI digestion products were sent off for sequencing to verify the sequence of the insert in the vector.

The cloned ORFs were then subcloned from the pGEM®-T Easy vectors into pEAQ-HT expression vector (Sainsbury et al., 2009) using a ligation strategy that allows for the tagging of cloned proteins with polyhistidine-tags at the C terminus. To accomplish this, each insert as well as the pEAQ-HT vector was digested by *Sma*I and *Age*I-HF® (New England Biolabs, Ipswich, MA). For each digest, 4 µl plasmid, 2 µl CutSmart® buffer, 2 µl *Sma*I (NEB) and 12 µl water were mixed into a microcentrifuge tube and incubated for an hour at 25°C. Then, 1 µl of *Sma*I was added to the mixture, which was incubated for another hour at 25°C. Two µl of *Age*I-HF® were then added to the mixture, which was incubated for an hour at 37 °C. One µl of *Age*I-HF® was then added to the tube, which was incubated for another hour at 37 °C. Each digested plasmid was mixed with 5 µl loading dye, loaded onto 1% agarose gels along with BenchTop 1 kb DNA Ladder, run at 120 V for 20-30 minutes, and stained with ethidium bromide. The gels were visualized under UV light and the bands containing the desired DNA (based on size) were cut out. The DNA was then purified using the Zymoclean™ Gel DNA Recovery Kit. The purified inserts were cloned into the purified pEAQ-HT expression vector using the Blunt/TA Ligase Master Mix according to its protocol.

Each plasmid (pEAQ-HT with inserts) was transformed into competent *E. coli* (DH5α) cells using the same method described earlier, but plated on 50 μM kanamycin LB medium agar plates. For each plate, 4 colonies were picked and grown in 2 ml of Luria broth with 50 μM kanamycin. Small scale plasmid isolations mini-preps were performed to isolate the cloned plasmids, with approximately 500 μl of culture left behind and stored at 4°C for the subsequent midi-prep. PCRs were performed using the mini-prep products as templates and the p1321 and p1322 primers to verify subcloning success. The reactant volumes and PCR program were the same as previously described. The reserved cultures for which their mini-preps exhibited successful subcloning were mixed with 150 ml LB medium with 50 μM kanamycin and grown overnight at 37°C. Midi-preps were performed on the resulting cultures using the E.Z.N.A.® Plasmid Midi Kit (Omega Bio-tek, Norcross, GA). A midi-prep was not done for the V0-2 plasmid, as its mini-prep DNA concentration was sufficient for subsequent agrobacterium transformation. The final midi-prep products (and the final mini-prep product for V0-2) were prepared with p1321 and p1322 primers and sent off for sequencing to verify the orientation and successful in-frame cloning of the inserts. Their concentrations were measured by a NanoVue™ Spectrophotometer.

Transformation of competent cells of *Agrobacterium tumefaciens*

A colony of *Agrobacterium tumefaciens* (strain GV2260) was picked and inoculated in 3 ml of LB medium with 50 μM rifampicin. The culture was grown at 30°C on a roller drum overnight. The culture was added to 50 ml of LB medium in a 250 ml flask and grown at 30°C until at mid-log growth phase. The resulting culture was chilled for 5 minutes on ice, transferred to a 30 ml centrifuge tube and centrifuged at 664 xG for 5 min at 4°C. After the supernatant had been discarded, the tube was drained by inversion for 60 seconds, and the pellet resuspended in 1

ml of ice cold 20 mM $CaCl_2$. Bacterial suspension (0.1 ml) was placed into a microcentrifuge tube on ice for each plasmid as well as an extra one as a control. One µg of each of the plasmids (pEAQ-HT with inserts) was pipetted into their respective microcentrifuge tubes, which was then mixed by tapping. This was also done with 1 µg of pEAQ-HT with no insert to generate an empty vector control. A negative control tube containing only *E. coli* without plasmid was also used. All tubes were frozen in liquid nitrogen and thawed for 5 minutes at 37°C. One ml of LB medium was added to each tube, which were then incubated for 2 hours on a roller drum at 30°C. The tubes were centrifuged for 5 minutes at 4000 rpm. After 900µl of the supernatant from each tube had been removed, the pellets were resuspended and plated on 50 µM kanamycin 50 µM rifampicin LB medium agar plates. The plates were incubated for 48 hours at 30°C.

Agroinfiltration in plants of *Nicotiana benthamiana*

A. tumefaciens cells transformed with each of the expression vectors (V0-2, V1, V2, V3, C1-2, C3, GFP, and empty vector control) were infiltrated into separate *Nicotiana benthamiana* plants using the protocol described by Velásquez et al. 2009. The *N. benthamiana* plants were then grown in a controlled growth room set to a day length of 18 hours at 22°C for 4-6 days.

Polyacrylamide Gel Electrophoresis and Coomassie Staining

Sixteen mm diameter discs of infiltrated leaf sections were punched out using a sterile metal borer. The discs were placed in separate microcentrifuge tubes along with 200 µL 2x Laemmli sample buffer with 710 mM 2-mercaptoethanol. They were ground with a micropestle and boiled for 10 minutes. They were then run on Mini-PROTEAN® TGX Any kD™ precast gels (Bio-Rad, Hercules, CA), in which 20 µl of sample were loaded per lane along with 10 µl of

Precision Plus Protein™ Dual Color Standards ladder (Bio-Rad, Hercules, CA) in 1 lane, at 150 V for 45 minutes. The gels were then Coomassie stained and scanned.

Western Blot

Each transformed culture of *A. tumefaciens* was infiltrated into separate *Nicotiana benthamiana* plants in the same way described earlier but each plant was grown for 4 days between infiltration and leaf collection. For additional splicing experiments, *A. tumefaciens* cultures transformed with V0-2 were each infiltrated into 2 plants, one grown for 4 days and the other grown for 6 days. Five 16 mm infiltrated leaf discs from each plant were harvested into separate microcentrifuge tubes. A tube was frozen for subsequent TNA extraction and the remaining 4 were suspended in 2x Laemmli sample buffer with 710 mM 2-mercaptoethanol in the same manner described earlier. A single tube of leaf sample infiltrated with GFP-transformed *A. tumefaciens* culture was designated to be run with each leaf sample for subsequent normalization. For each infiltrated plant, 20 µl of the 4 suspended leaf discs, 20 µl of an empty vector control, 20 µl of the GFP control, and 10 µl of Precision Plus Protein™ Dual Color Standards ladder were loaded into Mini-PROTEAN® TGX Stain-Free Any kD™ gels and run at 150 V for 45 minutes. The gels were removed from their cases and imaged under UV light as total protein scans for subsequent normalization purposes.

The upper well portion of the gels was cut and removed. Tris-glycine buffer with 20% methanol was prepared as western blot transfer buffer. Each gel was placed on a piece of wet Whatman paper that had been dipped in transfer buffer. A nitrocellulose membrane dipped in transfer buffer was placed on the gel, and bubbles were removed using a pipette tip. Another wet

piece of Whatman was placed on top of the membrane, and everything was placed in the western transfer apparatus. The transfer was run at 80 V for 1 hour at 4°C with stirring.

The membranes were removed and incubated in the primary antibody, (anti-Tetra-His. Qiagen, Hilden, Germany. Cat No.: 34670) diluted 1:2000 in 3% bovine serum albumin (BSA) with shaking at 4°C overnight. The primary antibody was removed, and the membranes were washed with Tris-buffered saline plus 0.05% Tween-20 (TBST) for 10 minutes 3 times. The membranes were then incubated in the secondary antibody (goat anti-mouse alkaline phosphatase conjugate. Sigma-Aldrich, St. Louis, MO. Cat No.: A3562) diluted 1:2000 in 3% BSA with shaking at room temperature for 2 hours. The membranes were washed 3 times with TBST and then incubated with BCIP®/NBT-Blue Liquid Substrate (Sigma-Aldrich, St. Louis, MO) until adequately developed. The reaction was stopped by soaking in water after which the membranes were dried and scanned.

Quantification

The C1-2, C3, V0-2, and V2 protein samples were analyzed by western blot in the same manner, but they were transferred to polyvinylidene difluoride (PVDF) membranes (several attempts) rather than nitrocellulose membranes for clearer imaging. After the membranes were developed and washed, they were scanned and used for quantification.

ImageJ software (https://imagej.nih.gov/ij/) was used to roughly quantify the bands on the western blots to calculate relative translation levels of each cloned GRBV ORF (Heidebrecht et al., 2009). The bands were quantified based on their relative intensity and size. Each western blot band representing the GRBV His-tagged protein product was normalized to the GFP band on the blot. On the total protein scans, the corresponding rubisco band for each product lane was

normalized to a designated ladder band. The resulting GRBV protein product quantities (four biological replicates per ORF) were then normalized to their corresponding rubisco quantities. The calculated relative quantities were then averaged between replicates.

Spliced mRNA PCR

In order to test for the presence of spliced C1-2 and V0-2 mRNA in *N. benthamiana* plants infiltrated with the corresponding pEAQ-HT transformed *A. tumefaciens* culture, a reverse transcription polymerase chain reaction (RT-PCR) were performed according to the following process. Extraction buffer consisting of 2% CTAB, 2.5% PVP-40, 2 M NaCl,100 mM Tris-HCl pH 8.0, and 25 mM EDTA pH 8.0 was prepared for a total nucleic acid (TNA) extraction. For each sample of leaf infiltrated with *A. tumefaciens* cells transformed with C1-2, V0-2, and empty vector, the disc was ground in liquid nitrogen with a mortar and pestle. For V0-2, both the sample collected 4 days after infiltration and the sample collected 6 days after placement were processed. One thousand µl of extraction buffer warmed to 65°C was added to the mortar, and the sample was ground further. The resulting homogenate was quickly added to a microfuge tube along with 10 µl 2-mercaptoethanol and inverted to mix. The tube was incubated at 60°C with continual mixing for 1 hour. An equal volume of chloroform: isoamyl alcohol (24:1 v/v) was added to the tube, which was then vortexed vigorously and centrifuged at 11,000 xG for 10 min at 4°C. The supernatant was transferred to a fresh tube with an equivalent volume of ice-cold isopropanol. The tube was centrifuged at full speed for 15 minutes at 4°C. The liquid was removed, and the pellet resuspended in 500 µl 70% ethanol. The tube was centrifuged for 2 minutes at full speed at 4°C, drained, and air-dried for 5 minutes. The pellet was resuspended in 50 µl of sterile distilled water. The resulting TNA was measured by a NanoVue™ Spectrophotometer to determine concentration. One µg TNA

was treated with DNase according to the protocol provided by the RQ1 RNase-Free DNase enzyme (Promega, Madison, WI). The resulting RNA was run on a 1% agarose gel as previously described to visually verify elimination of DNA from the sample. The resulting RNA was then used in RT-PCR as follows. For every sample, a PCR tube was filled with 400 ng RNA, 2 µl of 10 µM random primers, 2 µl of 10 mM dNTPs, and sterile distilled water to a total of 13 µl. The tube was then incubated at 70°C for 5 minutes in a thermocycler. A mixture of 4 µl 5x First Strand Buffer (Invitrogen, Carlsbad, CA), 2 µl 0.1M DTT, and 1 µl MMLV-RT (Invitrogen, Carlsbad, CA) was then added to the tube. The tube was incubated in the thermocycler at 37°C for 60 minutes, followed by 70°C for 10 minutes, and held at 4°C.

PCRs were then performed using the resulting cDNA using the same PCR method described earlier. The primers for the C1-2 PCR were p1533 and p1376, and the primers for the V0-2 PCR were p1680 and p1402. The PCR products were mixed with 5 µl loading dye, loaded onto 1.5% agarose gels along with BenchTop 1 kb DNA Ladder, run at 120V for approximately 40 minutes, stained with ethidium bromide, and imaged under UV. The expression vectors (pEAQ-HT with inserts) were also included on the gel as controls.

Due to the appearance of two bands in the V2 western blot, the mRNA from leaf infiltrated by *A. tumefaciens* transformed with V2 was analyzed in the same way. The primers used for the PCR were p1684 and p1689.

Results

Tandem Mass-Spectrometry identifies two GRBV predicted proteins: C3 and V2

In order to develop a serological method for the detection of GRBV, knowledge of which virus-associated protein markers might be amenable to antibody development was first sought. The mass-spectrometry results were produced in a joint project with Jeremy Thompson.

Petiole+Base Leaf Samples: From LP protein, only the predicted C3 protein was detected. It was found in the gel slice containing the protein size range of 15-20 kDa (Fig. 5), consistent with its predicted size of 17 kDa. A single peptide resembling the 11 screened predicted peptides of C3 following trypsin digestion was detected.

Cane Samples: From the C protein, only peptides from the predicted V2 were detected in gel slices containing the protein size ranges of 18-23 kDa and 23-27 kDa. V2 was thus detected at sizes at and slightly above its predicted size of 19.2 kDa. In the 18-23 kDa gel section, 6 peptides resembling the 4 of the 16 screened predicted peptides of V2 following trypsin digestion were detected. In the 23-27 kDa gel section, only 1 peptide resembling the 16 screened predicted peptides was detected, and a peptide of the same sequence was also found in the 18-23 kDa section.

Petiole+Base Leaf Samples (Urea): No peptides resembling any predicted protein digestion peptides were detected in the gel section from the LPU sample.

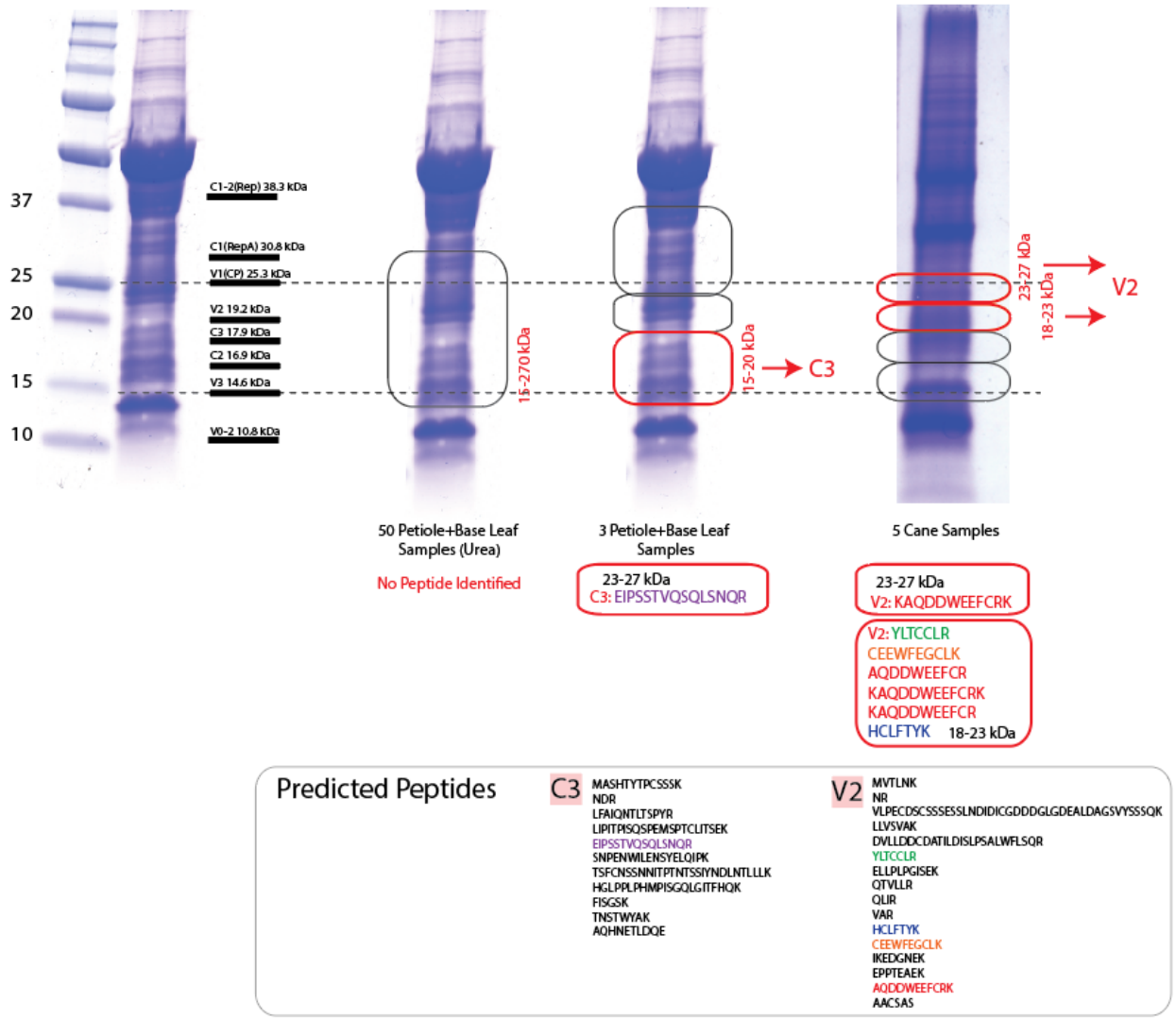


Figure 5. Tandem mass spectrometry analysis of GRBV-infected grape tissue. On the left, the sizes of the 8 predicted GRBV proteins are displayed in relation to the protein ladder used to separate gel slices. The submitted gel slices are enclosed, and red enclosures signify that peptides consistent with predicted GRBV peptides were detected. On the bottom, the predicted peptides following trypsin digestion are tabulated for C3 and V2. All predicted peptides for each predicted protein were screened for, but there were only matches for C3 and V2. Sequences that exhibit resemblance between predicted and detected peptides are depicted in matching text colors. Original figure courtesy of Jeremy Thompson.

Total protein profiles reveal a specific fragment for the C3 agroinfiltrated plants

A distinct band was visible in the lanes in which *N. benthamiana* leaf samples infiltrated by pEAQ-HT:C3His-transformed agrobacterium were run (Fig. 6). This band was not present in the GFP or empty vector controls. Given its position relative to the protein ladder, the band was

consistent with the predicted C3 size of 17 kDa. For none of the other gels in which samples of leaf infiltrated by other GRBV protein-transformed agrobacterium were run were distinct bands distinguishable (Appendix B).

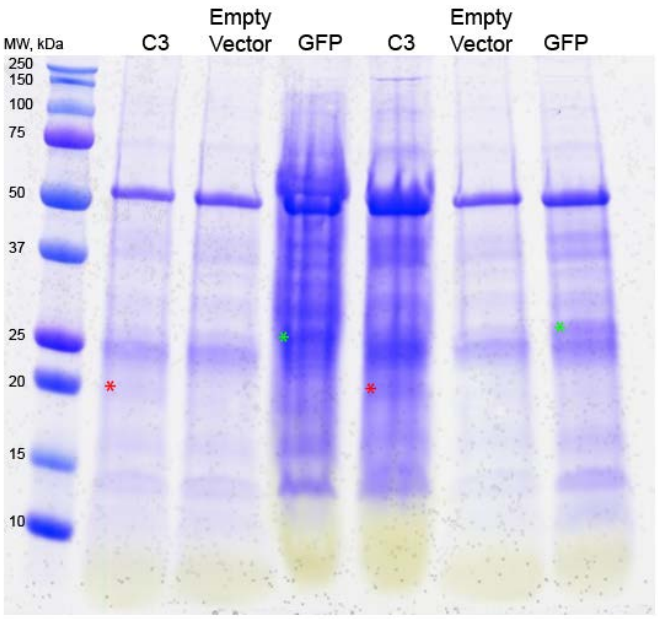


Figure 6. Coomassie stained polyacrylamide gel of leaf samples infiltrated with pEAQ-HT:C3His-transformed agrobacterium cells. The bands consistent with the predicted size of the C3 protein (17 kDa) are marked by red asterisks. Bands of that size are not visible in the control lanes loaded with leaf samples infiltrated by GFP and empty vector transformed leaf. The GFP (27 kDa) bands are marked by green asterisks.

Western Blot analyses of the protein expression in *N.benthamiana* plants showed detectable specific expression of proteins for C3, C1-2 and V2 constructs

Western blot analysis of protein from *N. benthamiana* leaf samples infiltrated by pEAQ-HT:C3His-transformed, pEAQ-HT:V2His-transformed, and pEAQ-HT:C1-2His-transformed agrobacterium on nitrocellulose membranes detected the presence of the expected GRBV proteins. Western blots of the C3, V2, and C1-2 samples probed on PVDF membranes are displayed in Figures 7-9. Western blot analysis of the V1, V3, and V0-2 samples did not detect the expected proteins (Appendix C). However, the blot of the V0-2 samples produced similar results as the blot

of the V2 samples (Fig. 10).

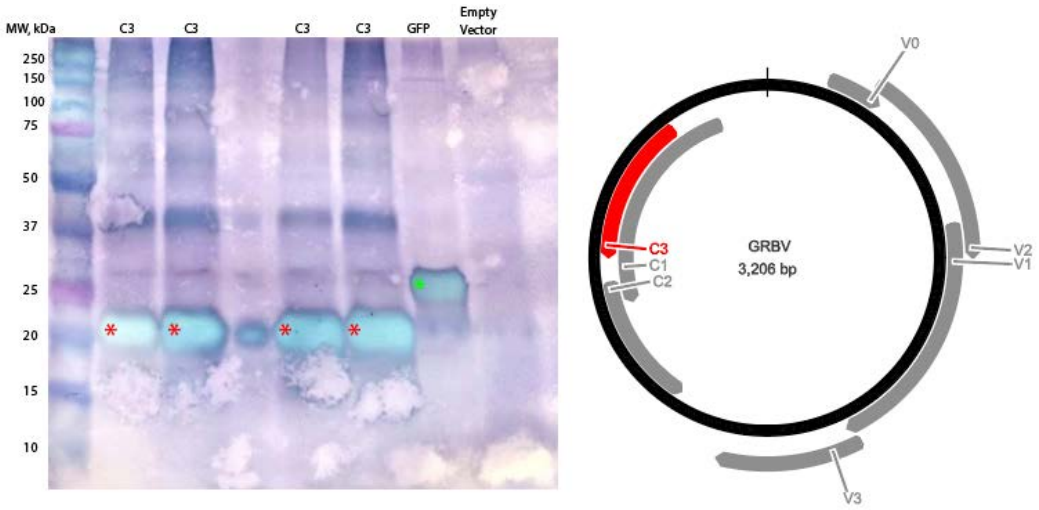


Figure 7. Western blot analysis of C3. The protein from leaf infiltrated with pEAQ-HT:C3His-transformed agrobacterium was transferred to PVDF membrane and probed for His-tagged protein (left). The bands marked by red asterisks represent C3 protein of predicted size of 17.9 kDa. Protein from leaf infiltrated with His-tagged GFP (predicted size 27 kDa, marked by a green asterisk) and empty vector serve as positive and negative controls, respectively. The predicted ORF of the GRBV genome that was expressed to synthesize the protein is in red in the image on the right.

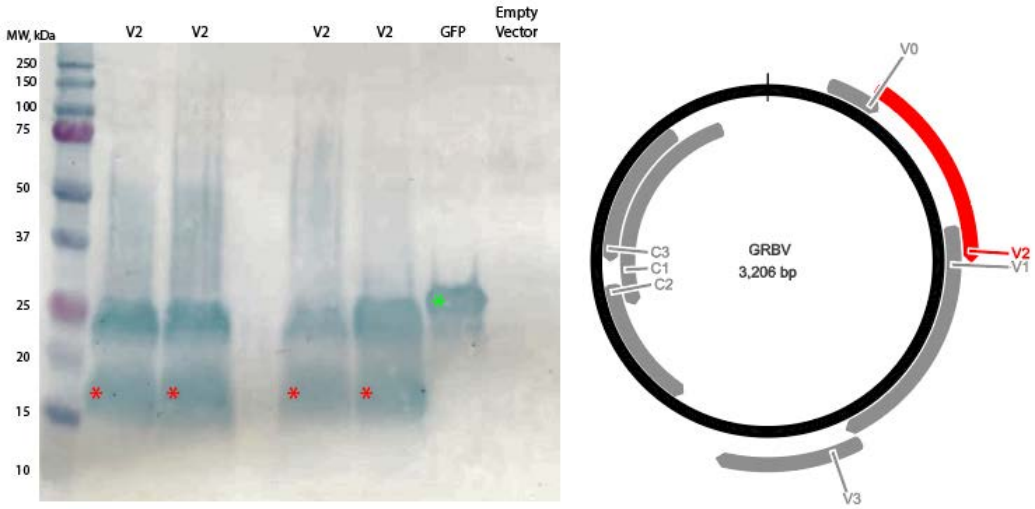


Figure 8. Western blot analysis of V2. The protein from leaf infiltrated with pEAQ-HT:V2His-transformed agrobacterium was transferred to PVDF and his-tagged V2 protein was probed for (left). The bands marked by the red asterisks represent V2 protein of predicted size of 19.2 kDa. The additional bands of greater molecular weight may be due to modifications (see discussion). Protein from leaf infiltrated with his-tagged GFP (predicted size 27 kDa, marked by a green asterisk) and empty vector serve as positive and negative controls, respectively. The section of the GRBV genome that was expressed to synthesize the protein is in red in the image on the right.

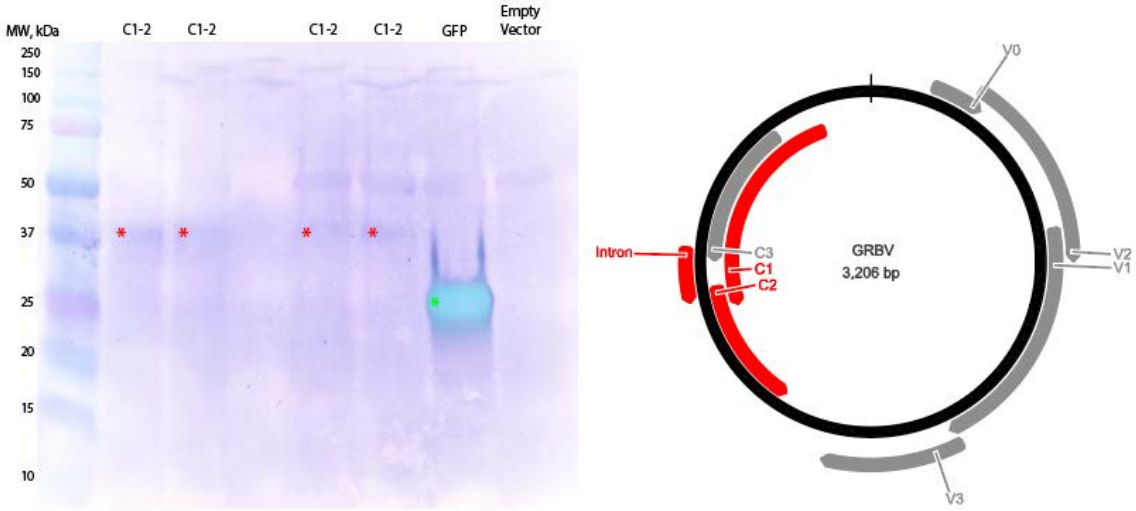


Figure 9. Western blot analysis of C1-2. The protein from leaf infiltrated with pEAQ-HT:C1-2His-transformed agrobacterium was transferred to PVDF and his-tagged C1-2 protein was probed for (left). The bands marked by the red asterisks represent the Rep protein (a fusion of the C1 (RepA) and C2 protein) of predicted size 38.3 kDa. Protein from leaf infiltrated with his-tagged GFP (predicted size 27 kDa, marked by a green asterisk) and empty vector serve as positive and negative controls, respectively. The section of the GRBV genome that was expressed to synthesize the protein is in red in the image on the right. The intron marks the section of the mRNA that was predicted to be spliced before translation.

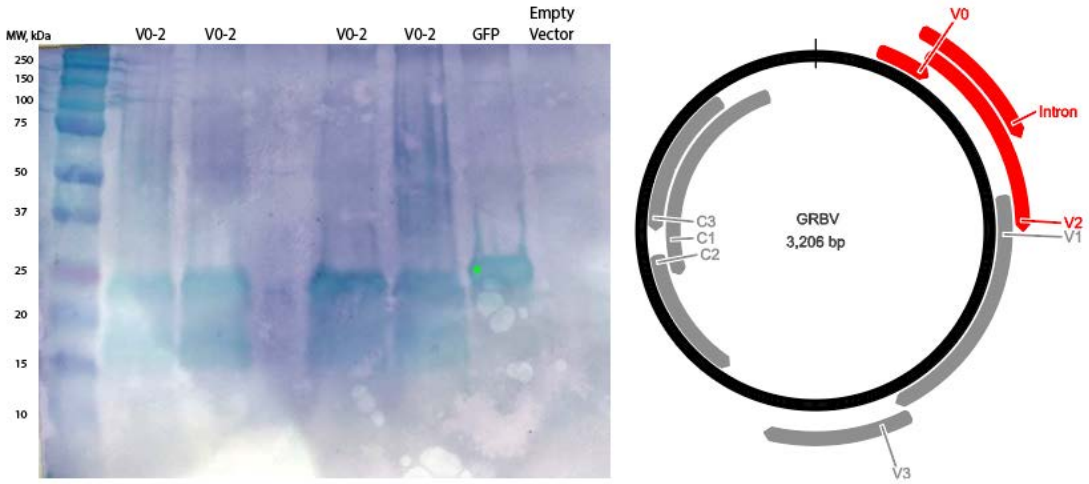


Figure 10. Western blot analysis leaf infiltrated with pEAQ-HT:V0-2His-transformed agrobacterium. The protein was transferred to PVDF and his-tagged proteins were probed for (left). No bands were detected at the predicted spliced V0-2 mRNA protein product of 10.8 kDa. However, bands resembling those detected in the V2 (predicted size 19.2 kDa) western blot (Fig. 8) were present. Protein from leaf infiltrated with his-tagged GFP (predicted size 27 kDa, marked by a green asterisk) and empty vector serve as positive and negative controls, respectively. The section of the GRBV genome that is predicted to be expressed to synthesize the protein is in red in the image on the right. The intron marks the section of the mRNA that is predicted to be spliced out before translation.

Quantification

Based on normalized western blot band intensity quantifications (Appendix D) of leaf samples infiltrated by pEAQ-HT:C3His-transformed, pEAQ-HT:V2His-transformed, and pEAQ-HT:C1-2His-transformed agrobacterium, the relative translation levels of the proteins were calculated (Fig. 11). Calculated values represent the translation levels of the ORFs relative to measured rubisco levels within the same sample.

Protein	*Average Relative Translation Level*	*Standard Deviation*
C3	73.4	40.1
V2	5.63	2.42
C1-2	0.415	0.117
V1	n.d.	-
V3	n.d.	-
V0-2	n.d.	-

Figure 11. Relative translation levels of predicted GRBV proteins. The band intensities from western blots were used to quantify relative translation levels. There were 4 biological replicates for each construct. Since no bands were detected from western blots of V1, V3, and V0-2, their relative translation levels were designated as not detected (n.d.).

An ANOVA test comparing the average relative translation levels between C3, V2, and C1-2 generated a p-value of 0.0066, indicating that the difference between the means is statistically significant. T-tests were also performed to compare the means of the relative translation levels of C3, V2, and C1-2 in pairs (Fig. 12). All differences between pairs of means were statistically significant.

Means Compared	*P-Value*
V2 and C1-2	0.0335
V2 and C3	0.00996
C1-2 and C3	0.00749

Figure 12. *T*-tests between pairs of means. All *p*-values were derived from two-tailed tests.

RT-PCR confirms C1-2 and V0-2 splicing

The PCR results indicated the presence of a spliced C1-2 mRNA transcript in leaf infiltrated by pEAQ-HT:C1-2His-transformed agrobacterium (Fig. 13). Based on the primers selected (Appendix A), the expected size of the PCR product of the transcript with the intron spliced out at the predicted site would be 203 base pairs (bp). A band of the appropriate size was detected. Sequencing data of the mRNA confirmed that the transcript was spliced in that the intron was missing in the expected location (Fig. 14) and that an in-frame fusion of the C1 and C2 ORFs had occurred to produce the predicted Rep protein as observed in the protein assay above (Fig. 9)

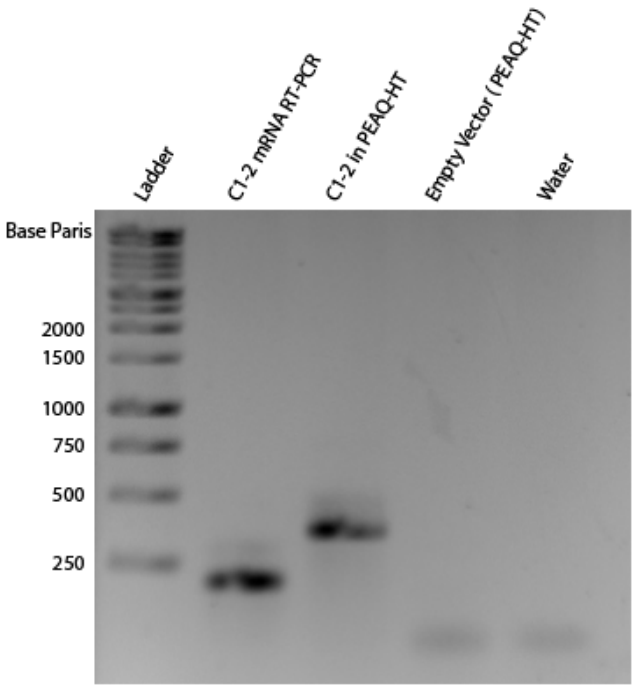


Figure 13. PCR analysis of mRNA-derived cDNA from leaf infiltrated by pEAQ-HT:C1-2His-transformed agrobacterium. Primers (p1533 and p1376) were selected so that the PCR product would be 203 bp if the predicted spliced C1-2 mRNA transcript was present and 366 bp if the C1-2 mRNA was unspliced. The band in the C1-2 mRNA lane is consistent with the expected spliced transcript size (203 bp), and the band in the C1-2 in pEAQ-HT negative control lane was consistent with the expected unspliced C1-2 size (366 bp). Lanes containing mRNA-derived cDNA from leaf infiltrated by empty pEAQ-HT-transformed agrobacterium as well as water serve as additional negative controls.

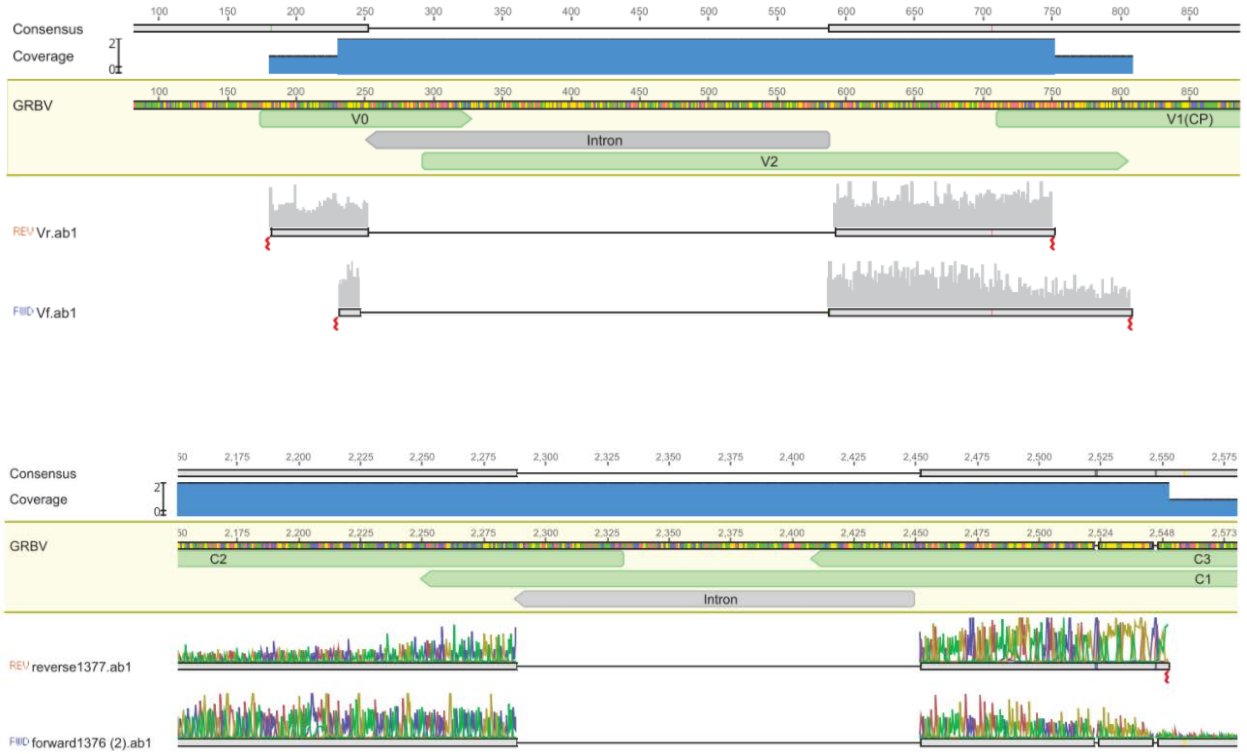


Figure 14. Sequencing data of the V0-2 (top) and C1-2 (bottom) mRNA PCR product. The sequences, when compared to the GRBV genome, indicate that the mRNA was spliced in the hypothesized positions since the introns were missing. Courtesy of Jeremy Thompson.

Similarly, PCR results revealed spliced V0-2 mRNA transcript in leaves infiltrated by pEAQ-HT:V0-2His-transformed agrobacterium (Fig. 15). Based on the primers selected (Appendix A), the expected size of the PCR product of the transcript with the intron spliced out at the predicted site would be 280 bp. Bands of that size were detected. However, a band that corresponded with the expected size of unspliced V0-2 mRNA (620 bp) was also detected in the leaf sample harvested 6 days after infiltration. Sequencing data of the 280 bp band confirmed that the transcript was spliced, as the predicted intron was missing (Fig. 14).

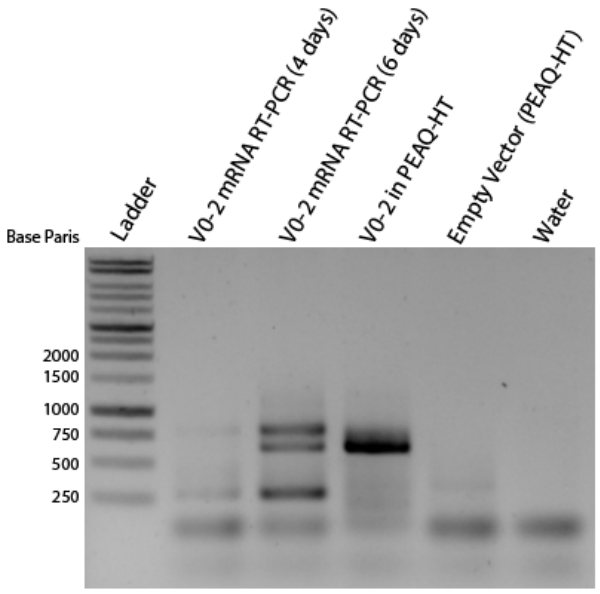


Figure 15. PCR analysis of mRNA-derived cDNA from leaves infiltrated by pEAQ-HT:V0-2His-transformed agrobacterium. Leaf harvested 4 days after infiltration as well as 6 days after infiltration were used. Primers (p1680 and p1402) were selected so that the PCR product would be 280 bp if the predicted spliced V0-2 mRNA transcript was present and 620 bp if the V0-2 mRNA was unspliced. The lower band in the V0-2 mRNA lane is consistent with the expected spliced transcript size (280 bp), and the band in the V0-2 in PEAQ-HT negative control lane was consistent with the expected unspliced V0-2 size (620 bp). In the mRNA from the sample harvested 6 days after infiltration, there was also a band that was consistent with the expected unspliced V0-2 size (620 bp). Lanes containing mRNA-derived cDNA from leaf infiltrated by empty pEAQ-HT-transformed agrobacterium as well as water serve as additional negative controls.

No evidence of spliced mRNA was found for V2 when the corresponding mRNA samples were screened in the same manner. A PCR product of the transcript matched that of V2 in pEAQ-HT in terms of size (Appendix E).

Discussion

The study aimed to investigate the protein expression profile of GRBV by analyzing infected grapevine protein extracts with tandem mass spectrometry. Another goal was to determine the relative translation levels of the ORFs. This data would identify potential diagnostic protein biomarkers. Furthermore, the study sought to test the hypothesized C1-2 and V0-2 splicing mechanisms.

From the proteins extracted from GRBV extracted grapevine tissue, only C3 and V2 were detected using TCA precipitation and tandem mass spectrometry. Prior to this study, there have been no successful attempts at directly detecting any GRBV proteins in infected plants (Vargas-Asencio et al., 2017). Thus, the detection of C3 and V2 represent the first empirical evidence of the predicted GRBV proteins.

Possible reasons for the limited number of proteins detected may be that the other predicted proteins aggregated in the wells of the gel and did not enter it. An attempt was made to address this by increasing the potential solubility of the proteins by the addition of urea. However, for the urea-treated sample, one large gel section with more protein was sent. This approach had the aim of increasing the amount of total protein analyzed but in the end appeared to "drown out" viral proteins altogether. This reflects the limited resolution of the method, in which only small sections of protein suspended in PAGE gel can be effectively analyzed at once. No gel sections were sent in which a putative V0 or V0-2 derived protein at its predicted size could be detected. Another likely reason for the detectable absence of other viral proteins is simply that the virus is at low titers in the plant and as such its proteins are below the limit of detection.

Of the 4 V2 predicted peptides that matched peptides detected in the leaf sample, one, YLTCCLR, is encoded in the predicted V0-2 intron. Thus, the detected protein was most likely

not the putative V0-2 but in fact V2. The gel sections (18-23 kDa and 23-27 kDa) from which the V2 peptides were detected were also consistent with its predicted size of 19.2 kDa, while the predicted size of the putative V0 and V0-2 proteins are 6.2 and 10.8 kDa, respectively. The V2 peptide detected in the 23-27 kDa gel slice may be attributable to poor protein migration in the gel, post-translational modifications, or re-dimerization of the protein following denaturing. There is evidence of other geminiviruses encoding movement-associated proteins that undergo post-translational modifications in their hosts (Pascal et al., 1994) and dimerization (Boulton, 2002).

V2 was only detected in cane tissue, while C3 was only detected in petiole and base leaf tissue (near the petiole). While this may reflect the limitations of tandem mass spectrometry, the results may also indicate that different types of grapevine tissues give rise to different TCA precipitation protein profiles. Previous unpublished data based on nucleic acid analysis demonstrates that petioles consistently contain more GRBV compared to their corresponding leaves, though cane was not analyzed (Setiono, 2016). Thus, a protein biomarker diagnostic test for GRBV may be effective if it samples proteins from petiole, cane, and perhaps leaf near the petiole (base leaf).

Upon Coomassie staining the gels on which proteins from the infiltrated *N. benthamiana* leaves were run, only C3 exhibited a distinct band of its predicted size that was not present in the controls. This may be a consequence of the high quantity of C3 produced by the infiltrated plants. Furthermore, the presence of host-derived protein bands may have obscured the lighter bands of the other GRBV protein constructs.

Western blot analysis using antibodies targeting His-tagged proteins detected bands representing V2, C3, and C1-2 in their respectively infiltrated *N. benthamiana* leaf samples. Separate infiltrations of agrobacterium transformed with C1 and C2 were not attempted due to the

assumption that complementary-sense encoded (replication-associated) proteins are less expressed and could not serve as potential diagnostic biomarkers. Infiltration with V0-transformed agrobacterium was not attempted. The small size of the predicted protein would make it difficult to express in *N. benthamiana* and also difficult to detect by western blot. V1, V3, V0-2, were not detected, which may indicate low levels of translation or may be the consequence of the detection limitations of western blot.

The western blot of the C3 samples detected two bands distinct from the controls. One is of the predicted C3 size of 17.9 kDa, and the other is around 40 kDa. Since it is about double the size of the lower band, the band of greater molecular weight may be the consequence of the reformation of C3 homodimers following denaturing. Alternatively, it may represent C3 proteins with a great deal of post-translational modifications. The western blot of the V2 proteins also exhibits two distinct bands, one of the predicted size of 19.2 kDa and the other about 24 kDa. Consistent with how V2 from GRBV infected grapevine was detected by mass spectrometry in a gel section containing proteins of 23-27 kDa, the upper band is likely the consequence of post-translational modifications to the V2 protein. PCR analysis of the mRNA extracted from the leaf infiltrated by V2-transformed agrobacterium did not detect any spliced product. Thus, it is unlikely that the upper band was the translated product of alternative splicing of the V2 transcript.

While *N. benthamiana* leaf infiltrated by agrobacterium transformed with V0-2 did not produce putative V0-2 at its predicted size of 10.8 kDa in quantities detectable by western blot, the samples produced bands nearly identical in size and appearance to those produced by leaf infiltrated by V2-transformed agrobacterium. This may be the result of the host ribosome initiating translation at the first methionine encoded by the V2 and bypassing the V0.

GRBV coat protein, encoded by V1, was not detected by western blot or Coomassie staining when transiently expressed in *N. benthamiana*, or by mass spectrometry in GRBV-infected grapevine. There have been no successful attempts at visualizing GRBV virions within or purified from infected plants, and there are no electron microscopic images of GRBV (Vargas-Asencio et al., 2017). Previous attempts to develop an antibody-based diagnostic method targeted the V1 coat protein as the biomarker, but this study indicates that V1 may be expressed at levels too low to serve as a biomarker. This observation is contrary to what one would expect of a viral coat protein. An alternative function for the V1 protein other than encapsidation is plausible but the virus' transmission by *S. festinus* implies that encapsidation of the viral DNA must occur.

Quantification and normalization of the western blot bands of the C3, V2, and C1-2 samples from infiltrated *N. benthamiana* allowed for relative translation levels to be calculated for the three proteins. As each construct was cloned into identical pEAQ-HT vectors which include identical transcription start sites, transcription levels of the GRBV sequences in *N. benthamiana* can be assumed to be equal. Thus, expression levels in the gene expression system used in this study are dependent on translational factors. Translation efficiency is determined in part by codon bias (different translation levels of the same peptides encoded by different codons) and mRNA folding patterns (Tuller et al., 2010). Furthermore, the nucleotide context surrounding individual codons has also been shown to affect translation efficiency (Shpaer, 1986). Also important to translation is the optimality of the starting codon context, known as the Kozak sequence (Kozak, 1986). The constructs of this study included the three nucleotides upstream (-3, -2, -1) of the ORF starting site at position +1, therefore allowing for sufficiently "native" codon context. Sequence data of the GRBV starting contexts and the optimal starting codon contexts are tabulated in

Appendix F. The stability of the synthesized proteins and their half-lives may have affected the calculated translation levels, representing a limitation to the experimental method.

Unpublished data on GRBV mRNA abundance in infected grape tissues indicates that virion-sense ORFs (V1, V2, V3) are more highly transcribed than complementary-sense ORFs (C1, C2, C3), as is the case for other geminiviruses (Shivaprasad et al., 2005). Specifically, the accumulated RNA-seq data suggest there may be about 15 times as many V-sense transcripts than C-sense transcripts (unpublished RNA-seq data). Furthermore, 7 V2 peptides were detected while only 1 C3 peptide was detected in the tandem mass spectrometry results, indicating that V2 may be at higher quantities in infected grapevine than C3. Thus, V2 may serve as a good protein biomarker for a diagnostic test even though the transient assays indicate that C3 has a translation level about 13 times higher that of V2. Indeed, the V2 protein was at high enough levels in GRBV-infected grapevine for detection by mass spectrometry and the V2 ORF is highly transcribed as well as highly translated.

The western blot detection of GRBV C1-2 protein expressed in *N. benthamiana* provides additional evidence that GRBV employs the C1-2 splicing mechanism common to geminiviruses (Rwahnih et al., 2013). PCR analysis of the mRNA extracted from *N. benthamiana* infiltrated with agrobacterium transformed with C1-2 and sequencing of the PCR product confirms that the transcripts were spliced by the host. Thus, mRNA transcript containing GRBV C1 adjacent to C2 was spliced into C1-2, and the spliced transcript was translated at levels high enough for western blot detection.

Although a putative V0-2 protein was not detected by western blot, PCR analysis of the mRNA in *N. benthamiana* leaf infiltrated with agrobacterium transformed by V0-2 revealed that splicing of the transcript occurs. Sequencing of the PCR product and comparing it to the unspliced

V0 and V2 ORFs identified the location of the intron and the splicing site, which were consistent with GRBV-infected grapevine mRNA sequencing (Sanger and Illumina) data (unpublished). Thus, mRNA transcript including the V0 ORF adjacent to the V2 ORF was spliced into V0-2 by a plant host. Unlike the analysis on C1-2 mRNA, PCR analysis of V0-2 mRNA showed that unspliced V0-2 transcript was present at detectable quantities in the infiltrated *N. benthamiana* leaf. Thus, it was likely that the unspliced transcript was translated to give rise to the V2-like proteins as detected by western blot.

Further work is necessary to attempt to detect V0 and V0-2 proteins by mass spectrometry, since gel slices that potentially contain them were not analyzed in this study. Considering the discovery of the relatively high translation levels of C3 and C1-2, the complementary-sense proteins of C1 and C2 should also be expressed separately to investigate their translation levels. V2 protein expressed by infiltrated *N. benthamiana* should be analyzed by mass spectrometry to determine what, if any, post-translational modifications distinguish the proteins forming the two bands detected by western blot. Using the same gene expression system employed by this study, large quantities of V2 protein could be synthesized and purified to develop antibodies for diagnostic purposes. Compared to bacterial expression methods (as was attempted previously for antibody production, unpublished), expression of V2 protein in *N. benthamiana* is likely to produce proteins more similar to the V2 protein found in infected grapevine. This is because *N. benthamiana*, a plant, has the machinery to make appropriate post-translational modifications to GRBV proteins.

Except for C1-2, C1, and V1, the predicted proteins encoded by GRBV do not have clearly deduced functions. Additionally, efficient viral proteins often exhibit multiple functions. As this study developed a method to express GRBV ORFs in *N. benthamiana*, there is the opportunity to

express the ORFs in the transgenic green fluorescent protein (GFP) expressing line of *N. benthamiana* designated as 16c. Given that 16c *N. benthamiana* can exhibit transgene silencing of GFP, the ability of GRBV proteins to enable the detection of GFP would provide preliminary evidence for their transgene silencing suppression activity (Stephan et al., 2011). Thus, another future direction is to screen GRBV proteins for their potential to suppress transgene silencing, a defense mechanism used by plants to protect themselves from viral infection. Other proteins of geminiviruses have been demonstrated to carry out this activity (Bisaro, 2006).

In conclusion, this study found that in infected grapevine, the GRBV C3 and V2 proteins are present in quantities high enough for detection by tandem mass-spectrometry. Furthermore, C3 was found to have the highest translation level, followed by V2, and then by C1-2; other proteins were not translated at levels high enough for detection. Considering this and previous data on transcription levels, V2 appears to be the best candidate for a diagnostic biomarker against which antibodies can be developed. Lastly, the study provides clear evidence for the splicing of C1-2 and V0-2 transcripts.

Acknowledgement

I would like to express my gratitude to Dr. Jeremy Thompson for his dedicated mentorship, stimulating guidance, and unwavering patience throughout the course of the project. I would also like to thank Dr. Keith Perry for his instruction and for giving me the opportunity to conduct research in his lab. Furthermore, I am grateful for the excellent company of the rest of the lab members including Heather McLane and Dr. Jose Vargas-Asencio for their generous and insightful advice.

Works Cited

Ahn, S. C., Baek, B. S., Oh, T., Song, C. S., & Chatterjee, B. (2000). Rapid mini-scale plasmid isolation for DNA sequencing and restriction mapping. *BioTechniques*, *29*(3), 466–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10997259

Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., & Daszak, P. (2004). Emerging Infectious Diseases of Plants: Pathogen Pollution, Climate Change and Agrotechnology Drivers. *Trends in Ecology & Evolution*, *19*(10), 535–544. http://doi.org/10.1016/J.TREE.2004.07.021

Bahder, B. W., Zalom, F. G., Jayanth, M., & Sudarshana, M. R. (2016). Phylogeny of Geminivirus Coat Protein Sequences and Digital PCR Aid in Identifying *Spissistilus festinus* as a Vector of Grapevine red blotch-associated virus. *Phytopathology*, *106*(10), 1223–1230. http://doi.org/10.1094/PHYTO-03-16-0125-FI

Bisaro, D. M. (2006). Silencing suppression by geminivirus proteins. *Virology*, *344*(1), 158–168. http://doi.org/10.1016/J.VIROL.2005.09.041

Boulton, M. I. (2002). Functions and Interactions of Mastrevirus Gene Products. *Physiological and Molecular Plant Pathology*, *60*(5), 243–255. http://doi.org/10.1006/PMPP.2002.0403

Calvi, B. L. (2011). *Effects of Red-leaf Disease on Cabernet Sauvignon at the Oakville Experimental Vineyard and Mitigation by Harvest Delay and Crop Adjustment*. University of California, Davis. Retrieved from https://search-proquest-com.proxy.library.cornell.edu/docview/908636906/fulltextPDF/7157B20FE787486APQ/1?accountid=10267

Cieniewicz, E., Perry, K., & Fuchs, M. (2017). Grapevine Red Blotch: Molecular Biology of the Virus and Management of the Disease. In *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 303–314). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-57706-7_14

Fondong, V. N. (2013). Geminivirus protein structure and function. *Molecular Plant Pathology*, *14*(6), 635–649. http://doi.org/10.1111/mpp.12032

Heidebrecht, F., Heidebrecht, A., Schulz, I., Behrens, S.-E., & Bader, A. (2009). Improved Semiquantitative Western Blot Technique with Increased Quantification Range. *Journal of Immunological Methods*, *345*(1–2), 40–48. http://doi.org/10.1016/J.JIM.2009.03.018

Hipp, K., Grimm, C., Jeske, H., & Böttcher, B. (2017). Near-Atomic Resolution Structure of a Plant Geminivirus Determined by Electron Cryomicroscopy. *Structure*, *25*(8), 1303–1309.e3. http://doi.org/10.1016/j.str.2017.06.013

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, *44*(2), 283–292. http://doi.org/10.1016/0092-8674(86)90762-2

Krenz, B., Thompson, J. R., Fuchs, M., & Perry, K. L. (2012). Complete Genome Sequence of a New Circular DNA Virus from Grapevine. *Journal of Virology*, *86*(14), 7715. http://doi.org/10.1128/JVI.00943-12

Krenz, B., Thompson, J. R., McLane, H. L., Fuchs, M., & Perry, K. L. (2014). Grapevine red blotch-associated virus is Widespread in the United States. *Phytopathology*, *104*(11), 1232–1240. http://doi.org/10.1094/PHYTO-02-14-0053-R

Lim, S., Igori, D., Zhao, F., Moon, J. S., Cho, I.-S., & Choi, G.-S. (2016). First Report of *Grapevine red blotch-associated virus* on Grapevine in Korea. *Plant Disease*, *100*(9), 1957–1957. http://doi.org/10.1094/PDIS-03-16-0283-PDN

Lütcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F., & Scheele, G. A. (1987). Selection of AUG initiation codons differs in plants and animals. *The EMBO Journal*, *6*(1), 43–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3556162

Madden, L. V, & Wheelis, M. (2003). THE THREAT OF PLANT PATHOGENS AS WEAPONS AGAINST U.S. CROPS. *Annu. Rev. Phytopathol*, *41*, 155–76. http://doi.org/10.1146/annurev.phyto.41.121902.102839

Martelli, G. P. (2017). An Overview on Grapevine Viruses, Viroids, and the Diseases They Cause. In *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 31–46). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-57706-7_2

Naidu, R., Rowhani, A., Fuchs, M., Golino, D., & Martelli, G. P. (2014). Grapevine Leafroll: A Complex Viral Disease Affecting a High-Value Fruit Crop. *Plant Disease*, *98*(9), 1172–1185. http://doi.org/10.1094/PDIS-08-13-0880-FE

Pascal, E., Sanderfoot, A. A., Ward, B. M., Medville, R., Turgeon, R., & Lazarowitz, S. G. (1994). The Geminivirus BRI Movement Protein Binds Single-Stranded DNA and Localizes to the Cell Nucleus. *The Plant Cell*, *6*, 995–1006. Retrieved from http://www.plantcell.org.proxy.library.cornell.edu/content/plantcell/6/7/995.full.pdf

Poojari, S., Alabi, O. J., Fofanov, V. Y., & Naidu, R. A. (2013). A Leafhopper-Transmissible DNA Virus with Novel Evolutionary Lineage in the Family Geminiviridae Implicated in Grapevine Redleaf Disease by Next-Generation Sequencing. *PLoS ONE*, *8*(6), e64194. http://doi.org/10.1371/journal.pone.0064194

Poojari, S., Lowery, D. T., Rott, M., Schmidt, A. M., & Úrbez-Torres, J. R. (2017). Incidence, distribution and genetic diversity of *Grapevine red blotch virus* in British Columbia. *Canadian Journal of Plant Pathology*, *39*(2), 201–211. http://doi.org/10.1080/07060661.2017.1312532

Rector, A., Tachezy, R., & Van Ranst, M. (2004). A sequence-independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification. *Journal of Virology*, *78*(10), 4993–8. http://doi.org/10.1128/JVI.78.10.4993-4998.2004

Reynard, J.-S., Brodard, J., Dubuis, N., Zufferey, V., Schumpp, O., Schaerer, S., & Gugerli, P. (2018). *Grapevine red blotch virus* : Absence in Swiss Vineyards and Analysis of Potential Detrimental Effect on Viticultural Performance. *Plant Disease*, *102*(3), 651–655. http://doi.org/10.1094/PDIS-07-17-1069-RE

Ricketts, K. D., Gómez, M. I., Fuchs, M. F., Martinson, T. E., Smith, R. J., Cooper, M. L., … Wise, A. (2017). Mitigating the Economic Impact of Grapevine Red Blotch: Optimizing Disease Management Strategies in U.S. Vineyards. *American Journal of Enology and Viticulture*, *68*(1), 127–135. http://doi.org/10.5344/ajev.2016.16009

Rwahnih, M. Al, Dave, A., Anderson, M. M., Rowhani, A., Uyemoto, J. K., & Sudarshana, M. R. (2013). Association of a DNA Virus with Grapevines Affected by Red Blotch Disease in California. *Phytopathology*, *103*(10), 1069–1076. http://doi.org/10.1094/PHYTO-10-12-0253-R

Sainsbury, F., Thuenemann, E. C., & Lomonossoff, G. P. (2009). pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnology Journal*, *7*(7), 682–693. http://doi.org/10.1111/j.1467-7652.2009.00434.x

Setiono, F. (2016). *The Development of a Quantitative Assay for the Detection of Grapevine Red Blotch-Associated Virus in Vitis Vinifera Identifies Significant Differences in Virus Distribution*. Cornell University. Retrieved from http://doi.org/10.7298/X4N29TWV

Shivaprasad, P. V, Akbergenov, R., Trinks, D., Rajeswaran, R., Veluthambi, K., Hohn, T., & Pooggin, M. M. (2005). Promoters, transcripts, and regulatory proteins of Mungbean yellow mosaic geminivirus. *Journal of Virology*, *79*(13), 8149–63. http://doi.org/10.1128/JVI.79.13.8149-8163.2005

Shpaer, E. G. (1986). Constraints on codon context in Escherichia coli genes their possible role in modulating the efficiency of translation. *Journal of Molecular Biology*, *188*(4), 555–564. http://doi.org/10.1016/S0022-2836(86)80005-5

Stephan, D., Slabber, C., George, G., Ninov, V., Francis, K. P., & Burger, J. T. (2011). Visualization of plant viral suppressor silencing activity in intact leaf lamina by quantitative fluorescent imaging. *Plant Methods*, *7*(1), 25. http://doi.org/10.1186/1746-4811-7-25

Tuller, T., Waldman, Y. Y., Kupiec, M., & Ruppin, E. (2010). Translation efficiency is determined by both codon bias and folding energy. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(8), 3645–50. http://doi.org/10.1073/pnas.0909910107

Vargas-Asencio, J., Wojciechowska, K., Baskerville, M., Gomez, A. L., Perry, K. L., & Thompson, J. R. (2017). The Complete Nucleotide Sequence and Genomic Characterization of Grapevine Asteroid Mosaic Associated Virus. *Virus Research*, *227*, 82–87. http://doi.org/10.1016/j.virusres.2016.10.001

Varsani, A., Roumagnac, P., Fuchs, M., Navas-Castillo, J., Moriones, E., Idris, A., … Martin, D. P. (2017). Capulavirus and Grablovirus: two new genera in the family Geminiviridae. *Archives of Virology*, *162*(6), 1819–1831. http://doi.org/10.1007/s00705-017-3268-6

Velásquez, A. C., Chakravarthy, S., & Martin, G. B. (2009). Virus-induced gene silencing (VIGS) in Nicotiana benthamiana and tomato. *Journal of Visualized Experiments : JoVE*, (28). http://doi.org/10.3791/1292

Wang, W., Vignani, R., Scali, M., & Cresti, M. (2006). A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. *ELECTROPHORESIS*, *27*(13), 2782–2786. http://doi.org/10.1002/elps.200500722

Yepes, L. M., Cieniewicz, E. J., Krenz, B., McLane, H., Thompson, J. R., Perry, K. L., & Fuchs, M. (2018). Causative Role of Grapevine Red Blotch Virus in Red Blotch Disease. *Phytopathology*, PHYTO-12-17-0419-R. http://doi.org/10.1094/PHYTO-12-17-0419-R

Appendices

Appendix A

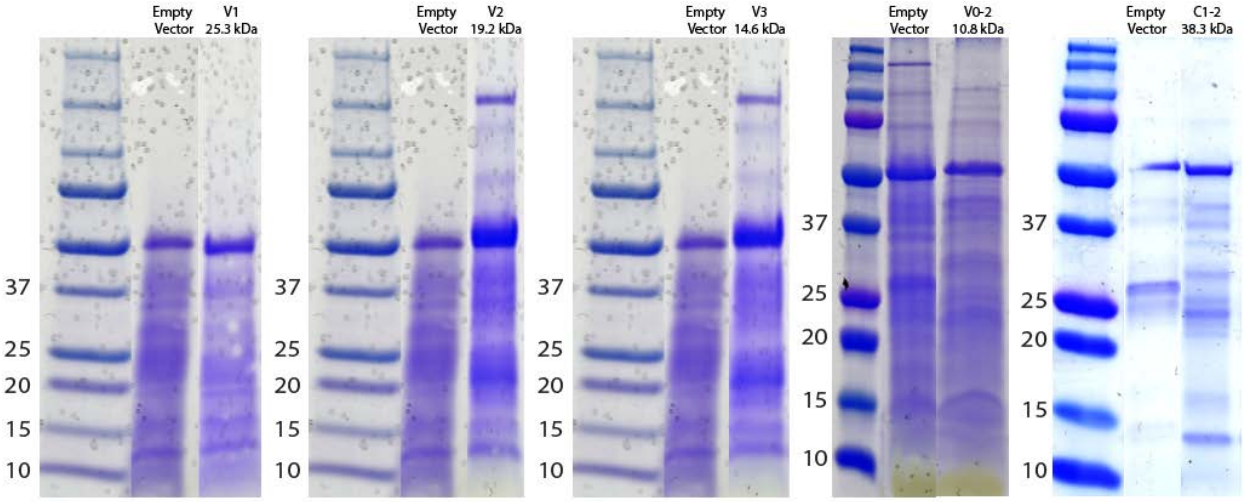
Primer Name	Sequence	Attributes
p1797.pEAQAgeIV3f	CGCGACCGGTATAATGGAGTTATTTTTCAATAGACCACG	*Age*I site, V3 ORF start
p1810.pEAQbluntV3r	GCACCCGGGCATTACTGTACAACGACTGTCTGGACT	*Sma*I site, V3 ORF end
p1799.pEAQAgeIV1f	CGCGACCGGTAGGATGGTAATGAAAAAAAGGAGCCGCCAACGGAA	*Age*I site, V1 ORF start
p1811.pEAQbluntV1r	GCACCCGGGTTGAAAAATAACTCCATTATAAAAGTTT	*Sma*I site, V1 ORF end
p1801.pEAQAgeIV2f	CGCGACCGGUGGAATGGTAACACTGAACAAACGGAATCGC	*Age*I site, V2 ORF start
p1812.pEAQbluntV2r	GCACCCGGGCGAGGCCGAGCACGCCGCCTTACGG	*Sma*I site, V2+V0-2 ORF end
p1803.pEAQAgeIC3f	CGCGACCGGUAGAATGGCGAGCCACACCTACACGCCTTGCTCATCT	*Age*I site, C3 ORF start
p1813.pEAQbluntC3r	GCACCCGGGTTCCTGGTCCAAGGTTTCGTTGTGCT	*Sma*I site, C3 ORF end
p1805.pEAQAgeIREPf	CGCGACCGGUAAGATGGCATCAAGTAGCTCCTTCAATCTCCGCTC	*Age*I site, C1-2 ORF start
p1814.pEAQbluntREPr	GCACCCGGGGTACAACGGTCTATCAATTTTTACAAAGA	*Sma*I site, C1-2 ORF end
p1856.pEAQAgeIV0f	CGCGACCGGTTCGATGCGGGTGGAAACGCACTTTGGCTG	*Age*I site, V3 ORF start

Primers used for ORF amplification. Amplified sequences were cloned into pEAQ-HT

Primer Name	Sequence
p1321	TATTGTTGCCTGTACTTCTTTCTTCTTC
p1322	AAGAAATTAAATTACATAAAATAAACACAC
p1533	CAAAACGAACTCTACGTGGAAG
p1376	TTACAAGGCAAATATTGGAATG
p1680	TCGATGCGGGTGGAAACGCA
p1402	ACTACGAGGCCGAGCACGCCG
p1684	ATGGTAACACTGAACAAACG

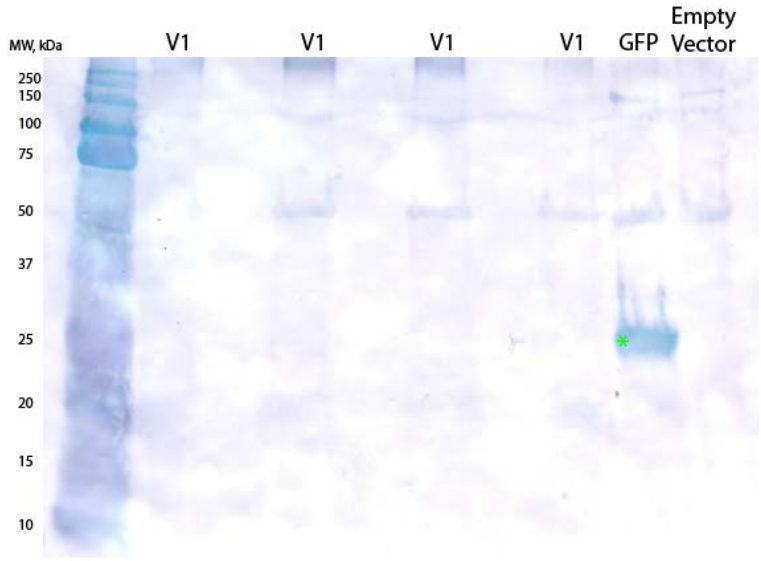
Primers used for PCR analysis and sequencing.

Appendix B

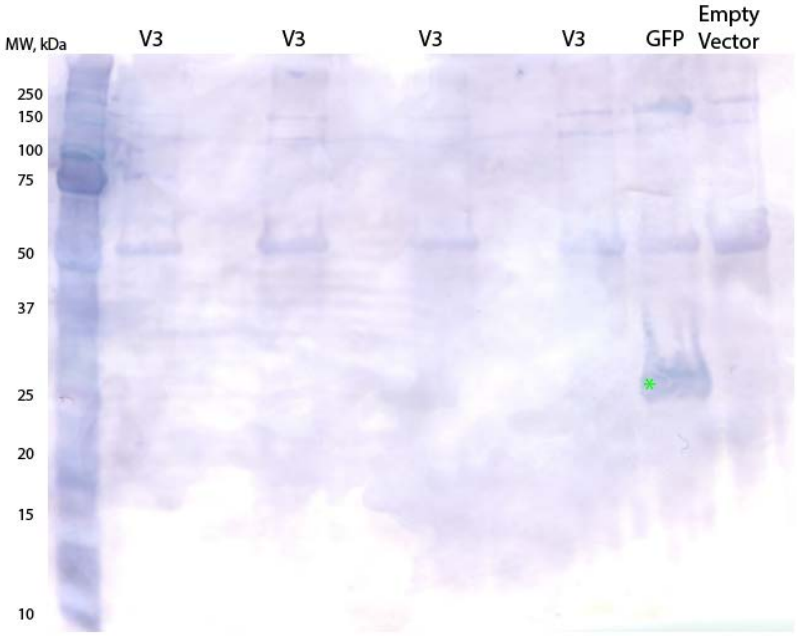


Coomassie stains of samples of leaf infiltrated by agrobacterium transformed with constructs and empty vector controls. Unlike the leaves expressing C3, none of these leaves expressed proteins that produced a distinct band detectable by Coomassie staining.

Appendix C

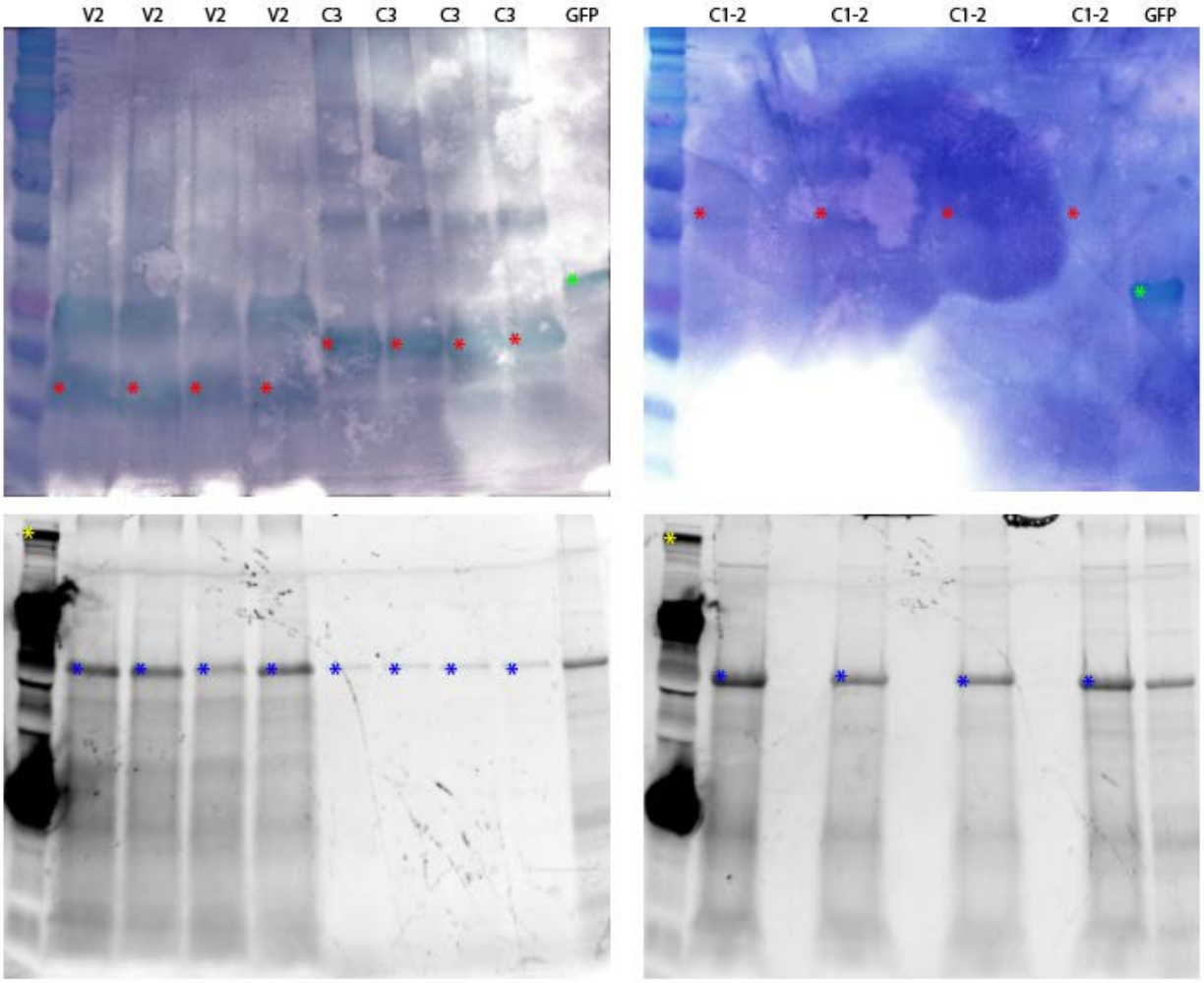


Western blot analysis of V1. The protein from leaf infiltrated with pEAQ-HT:V1His-transformed agrobacterium was transferred to nitrocellulose and his-tagged V1 protein was probed for. V1 protein of predicted size of 25.3 kDa was not detected. Protein from leaf infiltrated with his-tagged GFP (predicted size 27 kDa, marked by a green asterisk) and empty vector serve as positive and negative controls, respectively.



Western blot analysis of V3. The protein from leaf infiltrated with pEAQ-HT:V3His-transformed agrobacterium was transferred to nitrocellulose and his-tagged V3 protein was probed for. V3 protein of predicted size of 14.6 kDa was not detected. Protein from leaf infiltrated with his-tagged GFP (predicted size 27 kDa, marked by a green asterisk) and empty vector serve as positive and negative controls, respectively.

Appendix D

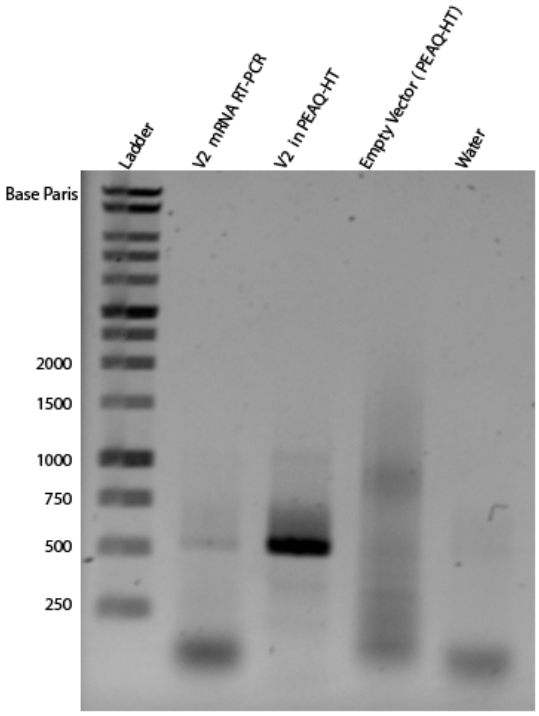


Band size and intensity quantification. Western blots on PVDF (top) were used for quantification of V2, C3, and C1-2 bands. The bands quantified are marked by red asterisks. The GFP bands, marked by green asterisks, were used for normalization. From total protein images (bottom), the rubisco bands (marked by blue asterisks) were quantified, normalized to the selected ladder band (marked by yellow asterisks), and used to normalize western blot band quantifications.

	DNA Ladder Top Band	Rubisco	Rubisco/Ladder	GFP	Construct	Construct/GFP	Relative Intensity
V2							
1	6392.74	3856.175	0.603212	17621.36	23312.588	1.322973	2.193216
2	6392.74	3768.468	0.589492	17621.36	23569.591	1.337558	2.269001
3	6392.74	2251.296	0.352164	17621.36	20382.588	1.156698	3.284538
4	6392.74	3823.225	0.598057	17621.36	39390.633	2.235391	3.737754
C1-2							
1	7487.832	8213.459	1.096907	17956.77	13961.132	0.777486	0.708798
2	7487.832	4478.539	0.598109	17956.77	14104.04	0.785444	1.313213
3	7487.832	3592.569	0.479788	17956.77	6531.12	0.363714	0.758072
4	7487.832	6935.731	0.926267	17956.77	3689.116	0.205444	0.221798
C3							
1	6392.74	376.627	0.058915	17621.36	47082.39	2.671893	45.35181
2	6392.74	452.042	0.070712	17621.36	38955.943	2.210723	31.26385
3	6392.74	622.87	0.097434	17621.36	78529.637	4.456502	45.73869
4	6392.74	644.87	0.100875	17621.36	41070.808	2.33074	23.10514

Band intensity data and normalization calculations.

Appendix E



PCR analysis of mRNA-derived cDNA from leaf infiltrated by pEAQ-HT:V2His-transformed agrobacterium. Primers (p1684 and p1689) were selected so that the PCR product would be 500 bp if the V2 mRNA was unspliced. The band in the V2 mRNA lane is consistent with the expected unspliced transcript size (500 bp), as is the band in the V2 in PEAQ-HT negative control lane. Lanes containing mRNA-derived cDNA from leaf infiltrated by empty PEAQ-HT-transformed agrobacterium as well as water serve as additional negative controls.

Appendix F

ORF	*Virus*	*Position (nt)*[a]	*Start codon context*	*Length (aa)*	*within sp. % aa id*	*between spp. % aa id*[f]
V0	GRBV	174-329	TTA **ATG** GA	46/51[b]	82.6-100	32.0, 29.4
	PrGVA	146-343	TGT **ATG** GA	65	100	32.0, 36.5
	WVV1	123-308	MRM **ATG** GA	61	86.9-96.7	29.4, 36.5
V2	GRBV	295-810	GGR **ATG** GD	171	89.5-100	48.9, 37.8
	PrGVA	258-788	AGG **ATG** CC	176	99.4-100	48.9, 38.6
	WVV1	256-819	AAA **ATG** TC	187	92.5	37.8, 38.6
V1 (CP)	GRBV	710-1384	AGG **ATG** GT	223/224	95.5-100	84.1, 78.7
	PrGVA	691-1365	AAG **ATG** GT	224	99.1-100	84.1, 77.8
	WVV1	716-1393	AAG **ATG** GT	225	98.7-100	78.7, 77.8
V3	GRBV	1365-1736	ATA **ATG** GA	123/124	91.9-100	64.0, 44.1
	PrGVA	1229-1717	AGA **ATG** AT	162[c]	100	64.0, 44.1
	WVV1	1377-1751	TTA **ATG** GA	124	90.3-100	44.1, 44.1
C1(RepA)	GRBV	3052-2250	AAG **ATG** GC	264	89.8-100	62.4, 51.9
	PrGVA	3017-2229	AAG **ATG** GC	262	96.6-100	62.4, 52.3
	WVV1	3038-2286	TAG **ATG** AC	250/261	86.2-100	51.9, 52.3
C2	GRBV	2332-1898	ATT **ATG** TY	144	89.6-100	79.4, 68.6
	PrGVA	2272-1877	CTG **ATG** TC	131	99.2-100	79.4, 71.3
	WVV1	2383-2003	TAG **ATG** WT[e]	126[d],148,149	90.5-100	68.6, 71.3
C3	GRBV	2890-2408	AVR **ATG** GC	160	75.6-100	56.3, 48.4
	PrGVA	2866-2390	AAG **ATG** GA	158	100	56.3, 49.7
	WVV1	2972-2493	TAG **ATG** GC	159[d]	83.6 -100	48.4, 49.7

Distribution and constitution of predicted open reading frames (ORF) of three grabloviruses; grapevine red blotch virus (GRBV), Prunus geminivirus A (PrGVA) and Wild Vitis virus 1 (WVV1). Courtesy of Jeremy Thompson.

[a]PrGVA (strain FT7 (OL3-22)) (Acc. no. MF579394), WVV1 (strain 1358) (Acc. no. MF185005)

[b] Clade 1 46aa, clade 2 51aa.

[c] due to presence of 39 aa at N-terminus

[d]deletion in NY1358

[e] excluding one sequences 1308 which 'starts' on second AUG and Met

[f] % ids in each row refer to a comparison with the other two viruses going from top to bottom.

Table II. AUG initiation codon consensus sequences in animals and plants

Animal

	6	5	4	3	2	−1	+1	2	3	4	5	6
A	22	18	25	80	28	19	100	–	–	29	26	14
G	43	15	11	17	11	17	–	–	100	38	16	37
U	21	28	8	0.5	14	7	–	100	–	17	18	29
C	14	39	56	2	47	56	–	–	–	15	39	20
Consensus:			C	A	C	C	A	U	G			

Plant

A	36	28	51	62	21	56	100	–	–	5	5	13
G	20	21	10	21	5	10	–	–	100	85	15	18
U	33	15	13	10	13	7	–	100	–	2	3	51
C	11	36	26	7	61	28	–	–	–	8	77	18
Consensus:			A	A	C	A	A	U	G	G	C	

Consensus (optimal) translation start context. Taken from Lütcke et al., 1987.