CHARACTERIZATION OF A SYMPTOM DETERMINANT OF GRAPEVINE FANLEAF VIRUS BY REVERSE GENETICS AND PROTEOMICS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Larissa Osterbaan

August 2019

CHARACTERIZATION OF A SYMPTOM DETERMINANT OF GRAPEVINE FANLEAF VIRUS BY REVERSE GENETICS AND PROTEOMICS

Larissa Osterbaan, Ph.D.

Cornell University 2019

Grapevine fanleaf virus (GFLV) is one of the most adverse viral pathogens of grapevine worldwide. GFLV is a species of the genus *Nepovirus* in the family *Secoviridae* in the order *Picornavirales* and is the causative agent of fanleaf degeneration of grapevines. This disease is characterized by symptoms including developmental abnormalities such as double nodes, shortened internodes, fasciations, zigzag growth; foliar symptoms include cupping, asymmetrical leaf shape, and the namesake "fanleaf" consisting of abnormally open petiolar sinuses and aberrant vein spacing in a radial pattern; and also poor and uneven berry set, leading to yield losses. This degeneration may lead to premature vine death. The molecular mechanisms underlying the development of GFLV symptoms are not well studied. In the herbaceous systemic host *Nicotiana benthamiana*, most GFLV isolates produce an asymptomatic infection, while a recently described isolate, GFLV strain GHu, induces distinct vein clearing symptoms on the apical leaves of *N. benthamiana.* Using a reverse genetics approach, I determined that symptom production in *N. benthamiana* by GFLV-GHu is modulated by a single residue of protein $1E^{Pol}$, the RNA-dependent RNA polymerase. Mutation of GFLV-GHu $1E^{Pol}$ residue 802 (lysine) modulates symptom development. While some substitutions (e.g. glycine, alanine) abolish symptoms in GFLV-GHu, other substitutions (arginine) have no effect on symptoms. Some mutations (e.g. proline, serine, threonine) produce intermediate, faint vein

clearing that is distinct from symptoms produced by wild type GFLV-GHu. Lysine 802 is not sufficient for symptom production in *N. benthamiana* and this residue appears to be flanked by motifs or structures that are highly conserved among GFLV isolates. I also probed the protein interaction network of $1E^{Pol}$ through affinity purification of epitope-tagged $1E^{Pol}$ from systemically infected *N. benthamiana* tissue, followed by protein identification by high-resolution tandem mass spectrometry. I found that $1E^{Pol}$, regardless of strain, appears to be in complex with GFLV proteins $1D^{Pro}$ and $1B^{Hel}$. GFLV-GHu $1E^{Pol}$ affinity purification products were specifically enriched in proteins with various cellular functions, most notably a number of proteins associated with chloroplasts. These data lay the groundwork for mapping the $1E^{Pol}$ protein interaction network and for identifying the mechanism by which GFLV-GHu induces vein clearing in *N. benthamiana*. These insights into the molecular biology of GFLV symptom development will aid efforts to develop GFLV management strategies in grapevine by providing information on virus-host interactions that play a central role in GFLV biology.

BIOGRAPHICAL SKETCH

Larissa Joy Osterbaan was raised in Hudsonville, Michigan. She attended Calvin College in Grand Rapids, Michigan as an undergraduate where she earned a Bachelor of Science with Honors degree in biotechnology in 2013. During the pursuit of her doctoral dissertation work, she held a number of leadership positions within the Cornell graduate student community, including secretary of the Plant Pathology Graduate Student Association; secretary, research symposium coordinator, treasurer, and volunteer coordinator for the Student Association of the Geneva Experiment Station; and coordinator for the Plant Pathology Graduate Student Seminar Exchange Program within the Geneva Plant Pathology and Plant-Microbe Biology seminar series. She also mentored six undergraduate research interns through the Summer Research Scholars program at Cornell AgriTech, all of whom pursued graduate studies. Kaitley Wozer and Jessica Carpenter assisted in the development and optimization of the GFLV agroinoculation system. Maddie Flasco, Jess Choi, and Jaimie Flasco created and characterized many of the GFLV mutants used to identify the GFLV-GHu symptom determinant. And Tori Hoyle played a major role in the development and optimization of the GFLV $1E^{Pol}$ affinity purification protocols, especially the creation of the V5-tagged GFLV-F13 construct.

ACKNOWLEDGMENTS

To my doctoral thesis advisor, Dr. Marc Fuchs. Thank you for your unfailing encouragement and patience over the years. I could not have asked for a more dedicated and attentive mentor to guide me in becoming an independent scientist, and one whom I hope you can be proud of.

To my thesis committee members, Drs. Michelle Heck and Jian Hua. Thank you for your illuminating and thoughtful advice and encouragement. Your insights and fresh perspectives have helped me past roadblocks that at the time felt insurmountable.

To my officemates, past and present: Ali Cala, Maryn Carlson, Libby Cieniewicz, Adrienne Gorny, Lori Koenick, Rachel Kreis, Martha Sudermann, and Anna Wallis. Thank you for being an invaluable source of help, guidance, encouragement, reassurance, and often-chocolatey snacks throughout my time at Cornell. Our discussions, both scientific and not, have been a great source of learning, entertainment, and personal growth. And thank you for reminding me that my self-worth is not defined by my research data. Particular thanks are owed to my labmate, Libby Cieniewicz, for being a sounding board for manuscripts and project ideas, a listener and co-commiserater when things were not going as planned, and just an overall most excellent friend.

To the Cornell AgriTech community. Thank you for making the New York State Agricultural Experiment Station an enjoyable and fulfilling place to work. The culture of collaboration and comradery which permeates this campus has influenced who I am as a scientist and as a person. I hope to find, or build, a similar sense of work community wherever life may take me.

To my family and friends, particularly my mom and dad, Beth and Larry Osterbaan, and my best friend, Lauren Huisman. Thank you for helping me become the person I am today and for your endless fount of emotional support. Thank you for being a solid

rock on which to ground my reality, no matter how my research was progressing.
To my husband, David Strickland. I could write a whole thesis enumerating the ways in which you have encouraged, supported, and cared for me and still far fall short of the gratitude and love I feel towards you. “Do you feel appreciated?” YES. “Do you feel supported?” YES.
There are many more people without whom this work would not have been possible and to them I owe many thanks. To the members of the Fuchs lab, past and present, including my mentees from the Summer Scholars program, for their support, fellowship, and encouragement: Keiran Cantilina, Jessica Carpenter, Mei Cheung, Fu-Wah Choi, Jess Choi, Alex Clarke, Rosemary Cox, Maddie Flasco, John Gottula, Kyle Hegel, Tori Hoyle, Jaimie Kenney, Dave MacUmber, Pat Marsella-Herrick, and Kaitley Wozer. To my collaborators and colleagues at INRA (France) for their critical feedback and assistance: Jean-Michel Hily, Corinne Schmitt-Keichinger, and Emmanuelle Vigne. To Stewart Gray and members of his lab, particularly Jason Ingram, for providing me with a lab-home on the Ithaca campus. To the Heck lab, particularly Stacy DeBlasio, Annie Kruse, and Jenny Wilson, for their help and seemingly limitless patience. To Kerik Cox, Noel Knight, Chris Smart, and Tom Burr, for their mentorship and guidance. To Kate Keagle, Holly King, and Gemma Osbourne for making all the non-science stuff (and some of the science stuff) easier. To Keith Rivera of Cold Spring Harbor Laboratory for conducting mass spectrometry. To all the PPPMB, SIPS, and AgriTech graduate students with whom I shared time at Cornell, who are too numerous to list here, for their comradery and friendship.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

DYNAMIC INTERACTIONS BETWEEN PLANT VIRUSES AND THEIR HOSTS FOR SYMPTOM DEVELOPMENT[1]

ABSTRACT

As obligate intracellular parasites, plant viruses catalyze drastic alterations in the cellular physiology of host cells in order to support their own replication. This disruption often, but not always, manifests macroscopically as disease symptoms. The search for what distinguishes symptom-inducing virus strains from their asymptomatic counterparts has long been a central component of plant virology research. A consistent through line has been the conclusion that symptoms arise from specific interactions between viral and host components. The identification of viral components responsible for symptom development (i.e. viral symptom determinants) followed by the identification and characterization of interactions with host components has led to concrete mechanistic linkages between the viral and host interactants for some symptoms. The rise of systems biology approaches (e.g. transcriptomics and proteomics) has allowed host responses to be described in greater detail, providing a broad view of the molecular events of plant virus infections. Here, we review the most recent literature describing plant virus symptom determinants. This includes studies detailing specific virus-host interactions which lead to symptom development, as well as those which utilize systems biology approaches such as transcriptomics to probe the molecular changes underlying the development of virus

[1]This chapter was published as: Osterbaan, L.J., and Fuchs, M. 2019. Dynamic interactions between plant viruses and their hosts for symptom development. J. Plant Pathol. https://doi.org/10.1007/s42161-019-00323-5

symptoms. Emerging trends, and how they might inform the future of plant virus symptomatology research, are discussed.

INTRODUCTION

Viruses are obligate parasites that hijack host metabolic functions and cellular machinery to enable their multiplication. This reorganization of cellular resources often triggers wide-ranging physiological disruptions in the host, leading to the expression of disease symptoms. In plants, symptoms of virus infection may include developmental defects such as stunting, malformed leaves, shoots, or roots, or a range of foliar symptoms such as chlorosis, mosaics, vein clearing or banding and even localized or systemic necrosis. Some infections may even lead to premature death of the plant. An early theory on how plant viruses induce symptoms in their hosts called the competitive disease model proposed that symptoms arise from these intracellular parasites merely overwhelming their hosts, with the host diverting too many resources to viral replication to maintain its normal cellular function. Nonetheless, this theory failed to explain the accumulating examples showing a lack of correlation between symptom severity and virus titer. It has become increasingly accepted that symptoms arise from specific interactions between virus and host components (Culver and Padmanabhan 2007; Pallas and García 2011). This interaction disease model has been bolstered by the results of the simple but powerful "classical" approach to plant virus symptom determinant research: the use of chimeras and reverse genetics to identify specific virus components responsible for symptom development, along with some characterization of host interactions with the viral component.

Recently García and Pallás (2015) reviewed the literature on plant virus pathogenicity factors. This review was largely comprised of case studies affirming the notion that the molecular mechanisms leading to plant virus disease symptoms are

highly specific to given host-virus-environment contexts (Culver and Padmanabhan 2007; Pallas and García 2011). These authors concluded this seminal article by predicting that *"[d]esigning of novel experimental approaches, mainly based on systems biology, and development of more powerful technological tools will be necessary to try to understand the molecular basis of these more complex plant virus pathologies in the coming years*." In light of the recent explosion of high-throughput profiling techniques brought to bear on plant virus symptomology research, the vision of García and Pallás (2015) has become reality. Indeed, the application of "omics" techniques to plant virus symptomatology research has made it increasingly obvious that symptom development is a dynamic process, composed of complex interplays between multiple viral and host interactants. In this review, we provide an update on the state of plant virus symptomatology research provided by García and Pallás (2015). We critically analyze the recent literature on the identification of plant virus symptom determinants and characterization of host responses to virus infection for symptom development through the use of advanced profiling techniques (transcriptomics, proteomics, metabolomics, etc.). From these analyses, we note some emerging trends in plant-virus interaction research that may advance this field in the future.

VIRUS COMPONENTS INVOLVED IN SYMPTOM DEVELOPMENT

The interaction disease model posits that symptoms arise from perturbations in host physiology arising from specific interactions between virus and host components (Culver and Padmanabhan 2007). These virus components are often called symptom

determinants or pathogenicity determinants. Changes to the interactions between these host and virus components often lead to alterations in symptom development.

Non-protein virus-host interactions for symptom development

Plant virus symptom determinants may map to either of the main two components of the virus: the nucleic acid genome itself or the proteins which it encodes. There have been cases of symptom determinants being mapped to non-coding regions of the genome (Rodríguez-Cerezo et al. 1991), to a non-protein-coding viral satellite (sat) RNA (Shimura et al. 2011; Smith et al. 2011), or to silent mutations (Krause-Sakate et al. 2005; Hirata et al. 2003). In particular, the work of Shimura et al. (2011) and Smith et al. (2011) generated significant interest by delineating a mechanism for symptom development in which sequence complementarity between a satRNA of cucumber mosaic virus (CMV, genus *Cucumovirus,* family *Bromoviridae)* and a host mRNA coding for a chlorophyll synthesis gene leads to silencing of the host mRNA and production of chlorosis symptoms. More recent examples of this mechanism of symptom production include a 25 nt segment of the tomato yellow leaf curl virus (TYLCV, genus *Begomovirus,* family *Geminiviridae*) intergenic region that has near perfect complementarity to a long noncoding RNA (lncRNA) of tomato. This complementarity between the TYLCV genome and host lncRNA leads to RNA silencing of the lncRNA (Yang et al. 2019). The lncRNA was found to be involved in tomato development and is upregulated during TYLCV infection of a susceptible tomato cultivar and downregulated in a resistant cultivar. Interestingly, the lncRNA sequence in a TYLCV-resistant tomato cultivar contains a 14 nt deletion that abolishes

the complementarity between the lncRNA sequence and viral small interfering (vsi) RNAs derived from the TYLCV intergenic region (Yang et al. 2019).

Similarly, twisting and stunting in *N. benthamiana* caused by rice stripe virus (RSV, genus *Tenuivirus,* family *Phenuviridae*) has been linked to silencing of *Nbe1F4A* via targeting by vsiRNAs derived from a segment of the RSV RNA4 which shares homology with the *Nbe1F4A* sequence (Shi et al. 2016).

Direct virus-host protein interactions for symptom development

In contrast to the handful of examples of nucleic acid-based symptom determinants, the vast majority of plant virus symptom determinants are proteins. Viral protein symptom determinants include the coat protein (CP) and associated products such as the read through protein (RTP, a minor capsid protein); movement protein (MP) and associated triple gene block (TGB) protein; viral suppressor of RNA silencing (VSR); genome-linked viral protein (VPg); and replicases, such as RNA-dependent RNA polymerase (RdRp) and associated domains such as the methytransferase (MT) or helicase (HEL). The disruption of the interaction between virus and host protein components may be direct or indirect.

Direct disruption of virus-host protein interactions requires the alteration of molecular contact points between the separate components. This could be through the targeted mutation of individual residues required for the interaction to occur or through mutations which disrupt protein structural domains, preventing proper intermolecular contacts. As an example of the former, a serendipitous mutation of Arg284 to cysteine in the MT domain of the replicase protein of cucumber green mottle mosaic virus (CGMMV, genus *Tobamovirus*, family *Virgaviridae*) produced

attenuated symptoms and delayed virus accumulation in *Nicotiana benthamiana* (Liu et al. 2017). Modeling of mutant replicases revealed that substitution of the native arginine did not alter the domain structural organization, and thus the functionality, of the replicase. The surface-exposed nature of the residue led these authors to speculate that mutations producing attenuated symptoms were likely altering interactions between the replicase and its viral and/or host partners (Liu et al. 2017). Follow up work to identify these partners and delineate how interruption of these interactions leads to altered symptoms would be of interest.

Similarly, mutation K67E (lysine to glutamic acid) in the triple gene block 3 protein (TGB3) of pepino mosaic virus (PepMV, genus *Potexvirus*, family *Alphaflexiviridae)* converts a mild strain to aggressive in both tomato and *Datura inoxia* (Hasiów-Jaroszewska et al. 2011). Furthermore, long-term passaging of PepMV revealed that this mutation is under positive selective pressure and predicted, along with other mutations found to be associated with increase symptom severity during long-term passage, to lie on the surface of the TGB3 protein (Minicka et al. 2015). As with most other symptom determinants, this suggests that the alteration in symptoms may arise from changes to protein-protein interactions induced by the mutation.

Another example of a single amino acid acting as a symptom determinant is the MP of broad bean wilt virus 2 (BBWV-2, genus *Fabavirus,* family *Secoviridae).* Symptom determination of BBWV-2 in both *N. benthamiana* and *Capsicum annuum* was mapped to a single amino acid, Arg424, of this protein and this residue was both necessary and sufficient to induce symptoms (severe mosaic, vein clearing, and leaf

malformation). However, the mechanistic underpinnings of symptom development remain unknown and it was not established if there is a difference in titer between the parent isolates used for this study (Seo et al. 2017).

An intriguing example of how protein-protein interactions can alter symptomatology is that of CMV infection in *N. clevelandii*, where it was found that symptoms were modulated not based on the amount of a viral component but on the intensity of the binding interaction between host and viral proteins. While CMV strain R induces green mosaic and malformation symptoms on *N. clevelandii*, a mutant bearing a single amino acid change (glutamic acid to arginine) to a surface-exposed β loop of the CP induces chlorotic spots on inoculated leaves and induces severe necrosis leading to plant death (Salánki et al. 2011). While both the wild type and mutant CP were able to interact with the F1 complex of ATP synthase, it was found that the mutated CP formed much stronger molecular contacts with the F1 complex, lethally interfering with its function in cellular respiration, leading to severe necrosis and plant death (Gellért et al. 2018).

The fidelity of the overall structure of the viral component, rather than the presence of specific residues, may be crucial for virus-host interactions for symptom development. Deletion of five amino acids near the C-terminus of the RTP of potato leafroll virus (PLRV, genus *Polerovirus*, family *Luteoviridae*) was sufficient to significantly decrease virus titer and reduce the appearance of interveinal chlorosis in hairy nightshade (*Solanum sarrachoides*) (Xu et al. 2018). This decrease in virus accumulation held true for other hosts, including *N. benthamiana* and *Physalis floridana* (ground cherry), along with a delay in symptom expression. Protein

modeling revealed that the deleted residues formed a distinctive ordered domain in this otherwise characteristically disordered portion of the RTP. Revertant isolates (which regained high viral accumulation and symptom induction) contained duplicated sequences which restored a crucial α helix structure. Interestingly, these duplicated sequences were not identical to those deleted in the original mutant. Further study demonstrated that the α-helix itself, not the particular residues it contains, was necessary for normal virus accumulation and symptom development, with the only specific residue requirement being an aromatic residue (Y, H, or F) at the terminus of the helix (Xu et al. 2018). This suggests that in some systems, considerations of alterations to the overall protein structure may be more informative than identification of single amino acid mutations which affect symptom development.

The determinant of vein clearing symptoms by grapevine fanleaf virus (GFLV, genus *Nepovirus*, family *Secoviridae)* strain GHu in *N. benthamiana* had previously been mapped to the ultimate 408 nt of the $1E^{Pol}$ (RdRp) coding region, though it could not be resolved whether symptoms were determined by the nucleotide or amino acid sequence of this region (Vigne et al. 2013). Recently, it has been shown that vein clearing symptoms are modulated by a single amino acid, Lys802, of the $1E^{Pol}$ protein (Osterbaan et al. 2019a). This residue lies between a highly structured region and a predicted alpha-helix motif of unknown function, both of which are highly conserved among GFLV isolates, suggesting that the identity of this residue modulates the overall structure of the $1E^{Pol}$ C-terminus and thus its ability to interact with other components to induce symptom development (Osterbaan et al. 2019a).

Alanine-scanning mutagenesis of the ourmia melon virus (OuMV, genus *Ourmiavirus*) MP identified two mutants showing necrosis (Pro202 to alanine) and pronounced mosaic symptoms (Gly137 to alanine) in *N. benthamiana*, in contrast to the mild chlorosis symptoms induced by the wild type virus. Interestingly, neither of the mutated amino acids were predicted to be solvent exposed; thus it is unlikely that these mutations directly altered or abolished an intermolecular contact point. However, both of the mutated residues (proline and glycine) are structurally unique among amino acids and substitution of these residues often disrupts protein structure (Betts and Russell 2007). Such a change in protein confirmation could disrupt intermolecular contacts. One of the residues (Gly137) was strictly conserved among OuMV isolates and closely related viruses, suggesting that pressure exists to self-attenuate the virus' potential to induce symptoms (Margaria et al. 2016).

In another case of self-attenuation of symptom induction, deletion of amino acids 58 to 84 of wheat streak mosaic virus (WSMV, genus *Tritimovirus*, family *Potyviridae*) increased the severity of chlorosis symptoms in a range of systemic hosts, including wheat, barley, maize, and rye. This increase in severity was accompanied by an increased accumulation of both viral RNA and CP (Tatineni et al. 2017). Since the mite vector of WSMV prefers to feed on green tissue, these authors suggest that the self-attenuation of chlorosis symptoms may be an adaptive trait to promote acquisition and spread of the virus by its vector, a conclusion supported by a recent report that these residues are necessary for efficient transmission by the wheat curl mite (Tatineni et al. 2018).

Systemic necrosis, or lethal systemic necrosis, is widely accepted to be the result of a system-wide runaway induction of the hypersensitive response (HR) in certain plant-virus infections. Since it shares molecular event similarities with HR, it is assumed that systemic necrosis is the result of programmed cell death triggered by R-gene recognition of a pathogen effector gene. The 11K domain of tomato torrado virus (ToTV, genus *Torradovirus*, family *Secoviridae*) lies at the N-terminus of its RNA1-encoded polyprotein, just upstream of the protease cofactor domain. The 11K domain induces an HR-like response during transient expression in *N. benthamiana* and enhances the pathogenicity of a heterologous virus such as potato virus X (PVX, genus *Potexvirus*, family *Alphaflexiviridae*). PVX typically causes mild mosaic on systemically infected leaves, while PVX expressing the 11K domain induced severe necrosis both on inoculated and systemic leaves. This domain did not exhibit any silencing suppressor activity (Wieczorek and Obrępalska-Stęplowska 2016). It will be interesting to see if these results in a heterologous system translate to a distinct function for the 11K domain in tomato.

The C-terminus of the $2A^{HP}$ putative homing protein of GFLV strain F13 functions as an avirulence (avr) factor in *N. occidentalis*, eliciting a genuine HR on inoculated leaves and via transient expression. However, this HR is not sufficient to perfectly restrict the virus as the authors noted that some plants did become systemically infected with variants harboring the avr portion of the $2A^{HP}$ protein (Martin et al. 2018). It would be interesting to investigate the dynamics of HR establishment and maintenance to determine which environmental factors play a role in the inconsistent ability of strain F13 to escape HR-mediated restriction.

The C4 MP of beet severe curly top virus (BSCTV, genus *Curtovirus,* family *Geminiviridae*) interacts with the catalytic domains of CLAVATA1, a receptor kinase with a central role in shoot meristem regulation (Li et al. 2018). This interaction is dependent on the S-acylation (addition of a fatty acid chain) of the C4 protein. In *Arabidopsis thaliana*, the S-acylation site of C4 serves as a determinant for leaf-curling and vein-swelling symptoms. BSCTV mutants producing a version of C4 in which the S-acylation site had been abolished were systemically infectious but did not produce symptoms (Li et al. 2018). Thus, mutation of the symptom determinant site does not directly abolish an interaction with a host component, but instead prevents the post-translational modification needed for the interaction. Similarly, the C4 protein of apple geminivirus (AGV, unassigned genus, family *Geminiviridae*) induced severe upward curling in *N. benthamiana* when expressed from a PVX vector. As with BSCTV (Li et al. 2018), mutation of a putative S-palmitoylation site altered C4 localization and reduced its ability to induce symptoms (Zhan et al. 2018).

Similarly, the βC1 protein of the betasatellite of tomato yellow leaf curl China virus (TYLCCNB, genus *Begomovirus,* family *Geminiviridae*), which is a pathogenicity factor which functions as a VSR, can be phosphorylated by SnRK1, a protein kinase of *N. benthamiana* involved in host defense against geminiviruses. When co-inoculated with its helper virus, wild type TYLCCNB produces severe leaf curling and shoot twisting in *N. benthamiana*. Mutations which mimic phosphorylation of βC1 leads to attenuation of disease symptoms. This attenuation was also seen with heterologous expression in PVX vector. The mutations had no effect on βC1 localization, though they did impair the ability of βC1 to interact with

NbAS1, a transcription factor which controls leaf development (Zhong et al. 2017). This is another example of virus-host interactions being modulated through post-translational modification of viral components, leading to alterations in symptom development.

Interaction between the viral protein genome-linked (VPg) of potyviruses and proteins in the elongation initiation factor 4E (eIF4E) family is required for successful potyvirus infection and thus eIF4E alleles are a common focus of recessive resistant breeding against potyviruses (Robaglia and Caranta 2006). Recently Gomez et al. (2019) found that the VPg of cassava brown streak virus (CBSV, genus *Ipomovirus,* family *Potyviridae*) can interact with a distinct clade of eIF4E proteins called novel cap-bind protein (nCBPs). Simultaneous introduction of indels to two nCBP genes via CRISPR/Cas9 in cassava caused CBSV to produce delayed and attenuated symptoms, though the endpoint titer in wild type and edited plants was not significantly different.

Indirect virus-host protein interactions for symptom development

Symptom development may be affected if the localization of a symptom-associated component (host or virus) is altered even if its functionality remains intact. For example, the difference in symptom expression in *N. benthamiana* of two isolates of pepper mild mottle virus (PMMoV, genus *Tobamovirus,* family *Virgaviridae*) was mapped to the 126 kDa replicase protein, which is a VSR. No qualitative difference in the silencing suppressor activity of the 126 kDa protein of these two isolates could be found but their cellular localization differed. The 126 kDa protein of a mild isolate, which produced no symptoms on inoculated leaves and minimal leaf distortions and chlorotic mottle on systemically infected leaves, was localized to the nucleus and the

cell periphery. In the severe isolate, which induced necrosis on inoculated leaves and chlorotic mottle on upper leaves, the 126 kDa protein formed small cytoplasmic punctae and was strongly aggregated along the cell periphery. Therefore, the authors hypothesized that aberrant localization of the 126 kDa protein from the mild isolate explains the attenuated host response (Han et al. 2017).

Similarly, the cysteine-rich proteins (CRPs) of six different carlaviruses enhanced the accumulation of a heterologous PVX vector. Along with other lines of evidence, this suggests that CRPs act as weak VSRs (Fujita et al. 2018). The determinant for symptoms induced by PVX harboring carlavirus CRPs, including leaf malformations and necrosis in *Nicotiana occidentalis*, was mapped to the highly-variable N-terminus region of these proteins, though both the highly conserved nuclear localization signal (NLS) and zinc finger (ZF) domains of the CRP were required for symptom induction (Fujita et al. 2018). How precisely this N-terminus region functions in symptom induction remains to be determined, though the authors hint that this region may be involved in cellular localization of the CRP, particularly in the proportioning of the protein to different cellular compartments.

In another example, both transient and heterologous expression of the radish mosaic virus (RaMV, genus *Comovirus*, family *Comoviridae*) HEL domain induced cell death in *N. benthamiana* (Hashimoto et al. 2015). These authors linked the membrane-modification functions of this domain to the programmed cell death response and found that this function of the HEL protein was conserved in closely related viruses, despite little sequence consensus of this region (Hashimoto et al. 2015). Interesting, the HEL domain only causes cell death when expressed transiently.

The viruses encoding the HEL domains do not induce cell death when infecting *N. benthamiana.* This confirms that plant viruses often self-attenuate their potential to drastically, or lethally, disrupt host physiology (Paudel and Sanfaçon, 2018).

Plant shoot apical meristems (SAMs) are typically resistant to most viral infections. While some viruses can transiently infect the SAM during early infection via deployment of VSRs, this invasion is cleared later in the infection cycle by strong RNA silencing activity in the SAM. In the case of CMV, it appears that this clearing of viral infection from the SAM may be aided by the virus itself through self-attenuation. Mutation of a conserved N-terminal arginine-rich region with RNA-binding capacity in the CMV CP enabled the mutant to persistently infect the SAM of *A. thaliana*, resulting in severe stunting and loss of apical dominance (Zhang et al. 2017). While wild type CMV CP was able to suppress the silencing suppressor activity of the 2b protein in green fluorescence protein (GFP) patch-infiltration assays, the arginine-mutated form had loss this function. Furthermore, the wild type CP had higher RNA binding activity than the mutated form. Both forms of the CP co-localized with host RNA silencing components RDR6 and SGS3 during infection. These authors proposed that wild type CP binds CMV RNAs, blocking their translation and increasing their use as templates for secondary siRNA production, thereby increasing host RNA silencing activity against CMV and allowing the host to clear CMV from the SAM (Zhang et al. 2017).

It is widely assumed and accepted that chlorosis symptoms arise from alterations to chloroplast structure and/or function that occur during viral infection. This assumption is now being validated by the multitude of transcriptomics and

proteomics papers demonstrating that chloroplast-related genes (and proteins) are widely represented among differentially-expressed genes identified during viral infection (see below). We are also starting to see these broad data confirmed by specific functional characterization of the identified genes. For example, the CP of a chlorosis-inducing strain of CMV (CMV-M) interacts with chloroplast ferredoxin I (Fd I) protein, while the CP of a green mosaic-inducing strain (CMV-Q) does not (Qiu et al. 2018). Additionally, these authors were able to demonstrate that CMV-M CP interacts with a precursor form of Fd I, preventing its proper localization to chloroplasts and thus inhibiting chloroplast function (Qiu et al. 2018).

Symptom remission, or "recovery", is a common phenomenon among geminiviruses. In a functional comparison of two geminiviruses which differ in their symptom remission phenotypes in *N. tabacum,* it was found that the CP precursor (AV2 protein) of the remission-type geminivirus tomato leaf curl Gujarat virus (ToLCGV, genus *Begomovirus,* family *Geminiviridae*) lacked the ability to block RDR1-mediated antiviral silencing. While tobacco RDR1 is able to direct the hypermethylation of the ToLCGV genome, leading to reduced virus titer and concomitant symptom remission in newly emerged leaves, this activity is blocked by the AV2 protein of a non-remission type virus, tomato leaf curl New Delhi virus (ToCNDV, genus *Begomovirus*, family *Geminiviridae),* thus allowing ToLCNDV to maintain high virus titer and induce symptoms throughout the course of infection (Basu et al. 2018).

HOST COMPONENTS INVOLVED IN SYMPTOM DEVELOPMENT

Advanced profiling techniques have enabled great strides in advancing our understanding of virus-host interactions for symptom development. Nonetheless, it is still common for transcriptomic studies of plant virus infection to be largely descriptive rather than hypothesis-driven. Comparison of the transcriptomes of healthy tissue to those from virus-infected tissue provide a broad overview of the impacts of viral infection on host gene transcription. However, this one-host-one-virus experimental design leaves us with few helpful lines of enquiry with which to probe the data. Fortunately, as the price of high-throughput transcriptomic services has fallen, it has become more feasible to integrate more variables for comparison in transcriptomic studies, such as infections with different virus variants, comparisons of several hosts, or time-course monitoring of gene expression changes. This increased granularity allows us to ask more precise questions of the data such as "what is the order in which gene expression changes during viral infection?" or "which genes are specifically regulated by virus X but not by virus Y" or "which pathways are commonly altered during most virus infections, regardless of host?"

In some cases, it is easy to generate hypotheses linking visual virus symptoms to those host genes which must be perturbed to produce them. For example, chlorosis is assumed to arise from alteration of genes controlling chlorophyll content. Reddening or other aberrant red-purple pigmentations is assumed to be the result of alterations to anthocyanin or flavonoid production. To confirm the former hypothesis, Chen et al. (2018) produced transgenic *A. thaliana* expressing the full length genome of brassica yellows virus (BrYV, genus *Polerovirus*, family *Luteorividae*) which

caused purple coloration of old leaves, among other symptoms. These authors analyzed two such transgenic lines, comparing the global transcriptome of these lines against wild type *A. thaliana*. By using two independent transgenic lines for comparative transcriptomics, they were able to identify differentially expressed genes (DEGs) which were shared in the two transgenic lines, allowing them to confidently identify DEGs that were the result of BrYV expression, rather than incidental DEGs due to the genomic context of the BrYV transgene. As expected, many host genes involved in flavonoid metabolism were disturbed in both transgenic lines, concomitant with increased anthocyanin content in these plants (Chen et al. 2018). This comparative transcriptomics approach to identify host genes perturbed by virus infection is particularly elegant to better distinguish background noise from true changes to host physiology.

The link between chlorosis and chloroplast related genes was confirmed in *N. benthamiana* plants infected with RSV (Shi et al. 2016). Importantly, these authors followed up their microarray analyses with virus-induced gene silencing (VIGS) experiments showing that silencing of differentially expressed genes identified by the microarrays lead to chlorosis. Interestingly, one of the silencing constructs (targeting eIF4A) produced leaf twisting and overall stunting, characteristic symptoms of RSV infection of *N. benthamiana.* It was later confirmed that leaf twisting and stunting caused by RSV is the result of silencing of *Nbe1F4A* targeted by vsiRNAs derived from a segment of the RSV RNA4 which shares homology with the *Nbe1F4A* sequence. A similar mechanism for symptom etiology was previously demonstrated

for chlorosis induced by the CMV Y satellite RNA (Shimura et al. 2011; Smith et al. 2011).

While transcriptomic studies can easily generate large datasets indicating genes affected by virus infection, it remains necessary to follow-up these surveys with functional characterization of the identified genes. The miRNA nbe-miR166h-p5 was found to be upregulated during PVX infection of *N. benthamiana* (Wang et al., 2018). Suppression of nb-miR166h-p5 was found to attenuate leaf yellowing symptoms of both PVX and turnip mosaic virus (TuMV, genus *Potyvirus,* family *Potyviridae).* Some, but not all, of the targets of nbe-miR166h-p5, including a receptor-like kinase and an auxin-responsive protein, were downregulated in PVX-infected plants, consistent with the upregulation of their controlling miRNA, nbe-miR166h-p5. However, individually silencing of these targets was not sufficient to reproduce PVX-like chlorosis, suggesting that additional layers of regulation exist between PVX-induced upregulation of nbe-miR166h-p5 and the induction of PVX symptoms (Wang et al. 2018).

Infection of wheat (*Tritium aestivum)* by barley yellow dwarf virus (BYDV, genus *Luteovirus,* family *Luteoviridae*) isolate GAV is marked by leaf yellowing and plant dwarfing. Transcriptomic analysis by microarrays revealed that genes controlling chlorophyll biosynthesis and chloroplast functioning, along with hormone metabolism, were well represented among the DEGs identified (Rong et al. 2018). Ultrastructural examination showed that leaf yellowing was associated with abnormal chloroplast development. It is hoped that data such as these will be followed up with functional characterization of the identified differentially expressed genes.

While comparison of transcriptomes from healthy tissue to those from virus-infected tissue provide an overview of the general impacts of viral infection on host gene transcription, a more informative experimental design is to compare two different states of virus infection to each other and to a healthy control. This allows for the possible identification of specific genes involved in symptom development by uncovering those genes impacted more strongly by the more severe virus state. For example, infection of *N. benthamiana* by peanut stunt virus (PSV, genus *Cucumovirus,* family *Bromoviridae)* induces mosaics, leaf malformations and plant stunting. These symptoms are more severe when PSV is co-inoculated with a satRNA. In comparing the transcriptomic profiles of *N. benthamiana* infected with PSV alone or PSV with satRNA, Obrepalska-Steplowska et al. (2018) found that while genes differentially expressed in response to either infection largely overlapped, the main difference was the intensity with which expression was altered, with the more severe infection (PSV+satRNA) inducing larger fold-changes in expression of many gene families. However, it should be noted that DEGs unique to each virus state were detected, with PSV+satRNA having more uniquely regulated genes than PSV alone. One of the most abundant classes of genes uniquely regulated by PSV+satRNA were those associated with protein modification processes, such as phosphorylation. The authors also noted changes in many of the pathways found to be altered in other plant virus systems, including hormone signaling, chloroplast functioning, cellular metabolism, protein translation, etc.

In another example, the transcriptomic changes induced by viruses causing distinct suites of symptoms were investigated. In tomato (*Solanum lycopersicum)*,

tomato chlorosis virus (ToCV, genus *Crinivirus*, family *Closteroviridae)* induces chlorosis and anthocyanin accumulation (purple pigmentation of infected leaves), while TYLCV causes stunting, reduced leaf size, and yellowing and curling of young leaves. Co-infection with these two viruses leads to simultaneous expression of both symptom suites: stunted plants with smaller leaves sporting purple pigmentation. As seen with PSV+satRNA (Obrepalska-Steplowska et al. 2018), a co-infection which induced more severe symptoms was marked by a larger number of DEGs than detected in either single infection. Interestingly, over 500 genes were specifically regulated by co-infection, suggesting that the marked increase in DEGs during co-infection is synergistic rather than merely additive. Genes involved in pigment metabolism were enriched among DEGs in ToCV infection, while those enriched in TYLCV viruses included pathways for plant hormone signal transduction and development. This observation was followed up by showing that brassinosteroid levels were greatly reduced in TYLCV-infected plants and that TYLCV stunting symptoms could be alleviated by foliar brassinosteroid applications. Genes involved in photosynthesis were downregulated in all three infection conditions, consistent with findings from a wide variety of plant virus systems.

Overall expression patterns for wild-type and recessive resistance-breaking isolates of melon necrotic spot virus (MNSV, genus *Carmovirus*, family *Tombusviridae)* were similar for both isolates in a susceptible melon cultivar lacking the recessive resistance allele. However, some specific genes from a wide range of functional classes were found to be specifically regulated by the resistance-breaking isolate. This suggests that on a global level, both isolates interact with the host in a

similar manner, with only a surprisingly small, though potentially significant, proportion of genes responding in an isolate-specific manner. The authors noted that they could not find a common regulatory or functional feature among the genes identified as being uniquely regulated by the resistance-breaking isolate (Gómez-Aix et al. 2016). This is important because it highlights the fact that while transcriptomic studies can identify host responses which are specific to a given virus isolate, they cannot yet identify the functional or mechanistic significance of the specific responses. Follow-up functional characterization of such host components are still needed to understand their role in virus infection.

Papaya sticky disease (PSD) is marked by spontaneous exudation of plant latex leading to unmarketable fruit. PSD develops post-flowering in *Carica papaya* plants co-infected with papaya meleira virus (PMeV, dsRNA, unassigned) and papaya meleira virus 2 (PMeV2, ssRNA, unassigned); infected plants remain asymptomatic prior to flowering. RNA sequencing of pre- and post-flowering plants infected with the PMeV complex revealed that more genes are differentially expressed during the asymptomatic (pre-flowering) stage of the disease than during the symptomatic (post-flowering) stage, a marked contrast to other studies that typically identified a correlation between the breadth of transcriptomic changes and the severity of symptom expression. It is interesting to note that while genes classified into the functional categories typically seen in plant-virus transcriptomics studies (defense responses, hormone signaling, chloroplast functioning, protein translation and cellular transport) were differentially expressed in both pre- and post-flowering PSD plants, some of the most highly accumulated genes in both datasets were genes with unknown

function (Madroñero et al. 2018). This highlights the reality that the conclusions we can draw from transcriptomics is only as good as the available annotations.

Ever since the discovery that chlorosis symptoms in *N. tabacum* induced by infection with a satRNA of CMV were due to sequence homology between the satRNA and a host chlorophyll synthesis gene leading to targeting the host gene mRNA by RNA silencing (Shimura et al. 2011; Smith et al. 2011), there have been searches for further examples of vsiRNA-targeting of host transcripts as the underlying mechanism of virus-induced symptoms. Additional concrete examples of this mechanism have been reported recently (Yang et al. 2019; Shi et al. 2016). Thanks to reduced costs of small RNA (sRNA) sequencing (also called deep sequencing) and the proliferation of sequenced plant genomes, it is fairly easy to find putative host gene targets among the vsiRNAs identified during virus infection. However, without experimental validation of decreased expression of the putative targets, we are left with little insight into what real impacts these vsiRNAs have on host physiology (Qiu et al. 2017). When follow up studies are conducted, the results may be inconclusive. For example, Ramesh et al. (2017) found little correlation between putative targeting of tomato transcripts by TSWV vsiRNAs and down-regulation. In fact, some of the putatively targeted transcripts were in fact upregulated during TSWV infection. These authors astutely point out that the study of virus-host interactions at the RNA silencing interface is bound to be complicated, given that both parties encode a gauntlet of proteins specifically intended to manipulate RNA silencing machinery, likely precluding the possibility of a tidy one-to-one relationship between vsiRNA targeting and host transcriptomic responses. Likewise, Lan et al.

(2018) followed up their sRNA deep-sequencing of rice infected with rice black-streaked dwarf virus (RBSDV, genus *Fijivirus,* family *Reoviridae)* with degradome sequencing to detect host targets actually degraded following targeting by vsiRNAs. A total of 25 target genes was found within the degradome which mapped to confirmed RBSDV vsiRNAs. This is a far cry from the nearly 200 putative targets of CMV vsiRNAs identified within the tobacco genome (Qiu et al. 2017). This suggests that the joint interpretation of deep sequencing and transcriptomic data will require nuance and extensive experimental validation.

ENVIRONMENTAL FACTORS INVOLVED IN SYMPTOM DEVELOPMENT

It has long been observed that temperature can have a significant impact on the outcome of plant virus infections. In some virus-host combinations, increased temperatures can exacerbate virus symptoms (Selman and Grant 1957). In other cases, elevated temperatures trigger a phenomenon called "heat masking" in which symptoms are reduced or eliminated at higher temperatures despite continued infection of the host (Kassanis 1952). While modern studies continue to describe both positive (Wosula et al. 2016; Chung et al. 2018) and negative (Chung et al. 2016; del Toro et al. 2017; Godfrey et al. 2017; Masinde et al. 2018) correlations between temperature and the severity of plant viral diseases, there have been fewer studies which probe the underlying molecular mechanisms of these phenomena.

In CMV-infected *N. tabacum,* incubation at higher temperatures (28°C versus 18°C) led to higher viral accumulation and more severe vein clearing symptoms. Molecular characterization of CMV-infected plants revealed that several salicylic acid (SA) associated genes were upregulated at lower temperatures, while jasmonic acid

(JA) associated genes were upregulated at higher temperatures (Zhao et al. 2016). SA-dependent responses are typically associated with plant defense against biotrophic pathogens, while JA pathways are often triggered in response to necrotrophic pathogens. These pathways are antagonistic to one another such that plants in which JA pathways are induced are more susceptible to biotrophic pathogens and vice versa.

Similarly, tracking of heat shock protein (HSP) and pathogenesis-related (PR) genes in both resistant and susceptible *S. tuberosum* plants during potato virus Y (PVY, genus *Potyvirus*, family *Potyviridae*) infection at different temperatures revealed that both cultivars upregulated HSP genes in response to heat stress, but that the resistant cultivar had higher HSP levels overall. PR genes were upregulated in the resistant cultivar during heat stress, but downregulated in the susceptible cultivar. This correlated with an increase in symptom severity in the susceptible cultivar.

A severe strain of alternanthera mosaic virus (AltMV, genus *Potexvirus*, family *Alphaflexiviridae)* induces mosaic and localized necrosis in *N. benthamiana* at 25°C. At 15°C, these symptoms escalate to systemic necrosis and plant death by 30 dpi. A chimera containing the CP of a mild AltMV strain reduced severe strain symptoms to systemic mosaic at 15°C. The severe strain CP was competent to induce systemic necrosis via heterologous expression within a PVX vector, suggesting that the CP is specifically a symptom elicitor. While there was no significant difference in viral titer of the severe and mild strain at a given temperature, the titer of both strains was significantly higher at 15°C compared to 25°C. This fits the trend of symptom severity correlating with the absolute titer of a symptom elicitor. It was found that the

SN-inducing form of the CP interacts with a boron transporter, likely perturbing boron metabolism and triggering systemic necrosis (Lim et al. 2014).

It is surprising that given the notable increase in applying transcriptomic techniques to the study of plant virus symptoms, there appears to be a shortage of studies which include temperature as a variable. Indeed, in their review of plant transcriptomic responses to biotic and abiotic stresses, Zhang and Sonnewald (2017) noted that the number of studies "applying combined stress scenarios…is still very low." As mentioned earlier, the falling prices of transcriptomic and other systems biology approaches should allow the field of plant virology to expand into this largely unexplored area of research.

THE OTHER DIMENSIONS: SPACE AND TIME

In the resurgence of the "grind-and-find" era of plant virus research, it can be easy to forget that plant tissues do not represent a homogenous swath of cells. Most advanced profiling techniques focus on leaf tissue, yet leaves are complicated plant organs comprised of a varied population of cells which may respond to virus infection differently. Yang et al. (2007) spatially dissected the transcriptomic responses of *A. thaliana* to TuMV infection. They found that the magnitude of mRNA changes correlated with the local accumulation in TuMV as it spread from initial infection foci. This framework could be used to explain the fact that comparative transcriptomics of *A. thaliana* plants infected with TuMV revealed no significant DEGs at a very early timepoint (4 dpi) (Sánchez et al. 2015). In fact, these authors posited that "there was probably a dilution effect with uninfected tissue." This serves as a forewarning in the design of time-course studies of plant virus systems which utilize systems biology

techniques: namely, that time-course studies should account for the uneven spatial distribution of viruses within plant tissue.

Profiling of the sRNA populations of *A. thaliana* infected with oilseed rape mosaic virus (ORMV) or crucifer-infecting tobacco mosaic virus (TMV-Cg, both of genus *Tobamovirus*, family *Virgaviridae*) revealed that both infections led to increased miRNA accumulations by 4 dpi, though the mild variant (TMV-Cg) actually caused a decrease in micro RNA (miRNA) accumulation at 2 dpi before catching up with the more severe variant (ORMV) by 4 dpi. Notably, this increase in miRNAs did not translate to a decrease in the accumulation of the mRNAs targeted by these miRNAs (Zavallo et al. 2015), leaving us with no mechanistic insight into how differences in the sRNA populations of plants infected with these two different viruses explains their different symptoms.

CONCLUSIONS AND PERSPECTIVES

Recent advances in our understanding of the development of plant virus symptoms have revealed some emerging trends. First, we continue to see evidence that symptoms arise from specific interactions between plant viruses and their hosts. In accordance with this model, it is possible for virus-host interactions to be changed directly (i.e. alteration of molecular contact points between components) or indirectly (e.g. aberrant localization of components), resulting in a change in symptom development. Furthermore, evidence is mounting that plant viruses commonly self-attenuate symptom development, possibly for evolutionary advantage. This trend is thoroughly reviewed in Paudel and Sanfaçon (2018). The rise of advanced profiling techniques (particularly transcriptomics) applied to plant virus symptomatology has

revealed some common host responses to virus infection, particularly disruption of photosynthesis-related genes. However, such studies often reveal more differences than similarities, highlighting the need for hypothesis driven follow-up studies for validation of systems biology data. Thus, a chasm still looms between the conclusions reached through “classical” plant virus symptom determinant research and those reached via “omics” studies. Concrete, mechanistic linkages between identified viral symptom determinants and global host physiological responses remain rare. Together, now more than ever, it is apparent that plant virus disease symptoms arise from complex and dynamic mechanisms of interplays between virus and host interactants that are influenced by environmental factors, as well as the temporal and spatial contexts of virus infections. It is expected that new discoveries will shed light on the mechanistic underpinnings of virus-host interactions for symptom development.

REFERENCES

Basu, S., Kumar Kushwaha, N., Kumar Singh, A., Pankaj Sahu, P., Vinoth Kumar, R., and Chakraborty, S. 2018. Dynamics of a geminivirus-encoded pre-coat protein and host RNA-dependent RNA polymerase 1 in regulating symptom recovery in tobacco. J. Exp. Bot. 69:2085–2102

Betts, M. J., and Russell, R. B. 2007. Amino-acid properties and consequences of substitutions. Pages 311–342 in: Bioinformatics for Geneticists, M.R. Barnes, ed. John Wiley & Sons, Ltd, Chichester, UK.

Chen, X.-R., Wang, Y., Zhao, H.-H., Zhang, X.-Y., Wang, X.-B., Li, D.-W., Yu, J.-L., and Han, C.-G. 2018. Brassica yellows virus' movement protein upregulates anthocyanin accumulation, leading to the development of purple leaf symptoms on *Arabidopsis thaliana*. Sci. Rep. 8:16273

Chung, B. N., Canto, T., Tenllado, F., Choi, K. S., Joa, J. H., Ahn, J. J., Kim, C. H., and Do, K. S. 2016. The effects of high temperature on infection by *Potato virus Y*, *Potato virus A*, and *Potato leafroll virus*. Plant Pathol. J. 32:321–328

Chung, B. N., Lee, J.-H., Kang, B.-C., Koh, S. W., Joa, J. H., Choi, K. S., and Ahn, J. J. 2018. HR-mediated defense response is overcome at high temperatures in *Capsicum* species. Plant Pathol. J. 34:71–77

Culver, J. N., and Padmanabhan, M. S. 2007. Virus-induced disease: altering host physiology one interaction at a time. Annu. Rev. Phytopathol. 45:221–243

Fujita, N., Komatsu, K., Ayukawa, Y., Matsuo, Y., Hashimoto, M., Netsu, O., Teraoka, T., Yamaji, Y., Namba, S., and Arie, T. 2018. N-terminal region of cysteine-rich protein (CRP) in carlaviruses is involved in the determination of symptom types. Mol. Plant Pathol. 19:180–190

García, J. A., and Pallás, V. 2015. Viral factors involved in plant pathogenesis. Curr. Opin. Virol. 11:21–30

Gellért, Á., Pósa, T., Fábián, A., Szabó, L., Bóka, K., Forró, B., Salánki, K., Drahos, L., Tóth, E., Juhász, A., and Balázs, E. 2018. A single point mutation on the cucumber mosaic virus surface induces an unexpected and strong interaction with the F1 complex of the ATP synthase in *Nicotiana clevelandii* plants. Virus Res. 251:47–55

Godfrey, S., Settumba, M., and Samuel, K. 2017. Effect of temperature on sweet potato virus disease symptom expression. African J. Agric. Res. 12:2295–2309

Gómez-Aix, C., Pascual, L., Cañizares, J., Sánchez-Pina, M. A., and Aranda, M. A. 2016. Transcriptomic profiling of *Melon necrotic spot virus*-infected melon plants revealed virus strain and plant cultivar-specific alterations. BMC Genomics 17:429

Gomez, M. A., Lin, Z. D., Moll, T., Chauhan, R. D., Hayden, L., Renninger, K., Beyene, G., Taylor, N. J., Carrington, J. C., Staskawicz, B. J., and Bart, R. S. 2019. Simultaneous CRISPR/Cas9-mediated editing of cassava eIF4E isoforms nCBP-1 and nCBP-2 reduces cassava brown streak disease symptom severity and incidence. Plant Biotechnol. J. 17:421–434

Han, S.-H., Park, J.-S., Han, J.-Y., Gong, J.-S., Park, C.-H., Kim, J.-K., Seo, E.-Y., Domier, L. L., Hammond, J., and Lim, H.-S. 2017. New Korean isolates of *Pepper mild mottle virus* (PMMoV) differ in symptom severity and subcellular localization of the 126 kDa protein. Virus Genes 53:434–445

Hashimoto, M., Komatsu, K., Iwai, R., Keima, T., Maejima, K., Shiraishi, T., Ishikawa, K., Yoshida, T., Kitazawa, Y., Okano, Y., Yamaji, Y., and Namba, S. 2015. Cell death triggered by a putative amphipathic helix of *Radish mosaic virus* helicase protein is tightly correlated with host membrane modification. Mol. Plant-Microbe Interact. 28:675–688

Hasiów-Jaroszewska, B., Borodynko, N., Jackowiak, P., Figlerowicz, M., and Pospieszny, H. 2011. Single mutation converts mild pathotype of the *Pepino mosaic virus* into necrotic one. Virus Res. 159:57–61

Hirata, H., Lu, X., Yamaji, Y., Kagiwada, S., Ugaki, M., and Namba, S. 2003. A single silent substitution in the genome of Apple stem grooving virus causes symptom attenuation. J. Gen. Virol. 84:2579–2583

Kassanis, B. 1952. Some effects of high temperature on the susceptibility of plants to infection with viruses. Ann. Appl. Biol. 39:358–369

Krause-Sakate, R., Redondo, E., Richard-Forget, F., Jadão, A. S., Houvenaghel, M. C., German-Retana, S., Pavan, M. A., Candresse, T., Zerbini, F. M., and Le Gall, O. 2005. Molecular mapping of the viral determinants of systemic wilting induced by a *Lettuce mosaic virus* (LMV) isolate in some lettuce cultivars. Virus Res. 109:175–180

Lan, Y., Li, Y., E, Z., Sun, F., Du, L., Xu, Q., Zhou, T., Zhou, Y., and Fan, Y. 2018. Identification of virus-derived siRNAs and their targets in RBSDV-infected rice by deep sequencing. J. Basic Microbiol. 58:227–237

Li, H., Zeng, R., Chen, Z., Liu, X., Cao, Z., Xie, Q., Yang, C., and Lai, J. 2018. S-acylation of a geminivirus C4 protein is essential for regulating the CLAVATA pathway in symptom determination. J. Exp. Bot. 69:4459–4468

Lim, H.-S., Nam, J., Seo, E.-Y., Nam, M., Vaira, A. M., Bae, H., Jang, C.-Y., Lee, C. H., Kim, H. G., Roh, M., and Hammond, J. 2014. The coat protein of *Alternanthera mosaic virus* is the elicitor of a temperature-sensitive systemic necrosis in *Nicotiana benthamiana*, and interacts with a host boron transporter protein. Virology 452–453:264–278

Liu, L., Peng, B., Zhang, Z., Wu, Y., Miras, M., Aranda, M. A., and Gu, Q. 2017. Exploring different mutations at a single amino acid position of *Cucumber green mottle mosaic virus* replicase to attain stable symptom attenuation. Phytopathology 107:1080–1086

Madroñero, J., Rodrigues, S. P., Antunes, T. F. S., Abreu, P. M. V, Ventura, J. A., Fernandes, A. A. R., and Fernandes, P. M. B. 2018. Transcriptome analysis provides insights into the delayed sticky disease symptoms in *Carica papaya*. Plant Cell Rep. 37:967–980

Margaria, P., Anderson, C. T., Turina, M., and Rosa, C. 2016. Identification of *Ourmiavirus* 30K movement protein amino acid residues involved in symptomatology, viral movement, subcellular localization and tubule formation. Mol. Plant Pathol. 17:1063–1079

Martin, I. R., Vigne, E., Berthold, F., Komar, V., Lemaire, O., Fuchs, M., and Schmitt-Keichinger, C. 2018. The 50 distal amino acids of the $2A^{HP}$ homing protein of *Grapevine fanleaf virus* elicit a hypersensitive reaction on *Nicotiana occidentalis*. Mol. Plant Pathol. 19:731–743

Masinde, E. A., Mkamillo, G., Ogendo, J. O., Hillocks, R., Mulwa, R. M. S., Kimata, B., and Maruthi, M. N. 2018. Genotype by environment interactions in identifying cassava (*Manihot esculenta* Crantz) resistant to cassava brown streak disease. F. Crop. Res. 215:39–48

Minicka, J., Rymelska, N., Elena, S. F., Czerwoniec, A., and Hasiów-Jaroszewska, B. 2015. Molecular evolution of *Pepino mosaic virus* during long-term passaging in different hosts and its impact on virus virulence. Ann. Appl. Biol. 166:389–401

Obrepalska-Steplowska, A., Zmienko, A., Wrzesinska, B., Goralski, M., Figlerowicz, M., Zyprych-Walczak, J., Siatkowski, I., and Pospieszny, H. 2018. The defense response of *Nicotiana benthamiana* to peanut stunt virus infection in the presence of symptom exacerbating satellite RNA. Viruses 10:449

Osterbaan, L. J., Choi, J., Kenney, J., Flasco, M., Vigne, E., Schmitt-Keichinger, C., Rebelo, A. R., Heck, M., and Fuchs, M. 2019. The identity of a single residue of the RNA-dependent RNA polymerase of grapevine fanleaf virus modulates vein clearing in *Nicotiana benthamiana*. Mol. Plant-Microbe Interact. 32:790–801

Pallas, V., and García, J. A. 2011. How do plant viruses induce disease? Interactions and interference with host components. J. Gen. Virol. 92:2691–2705

Paudel, D. B., and Sanfaçon, H. 2018. Exploring the diversity of mechanisms associated with plant tolerance to virus infection. Front. Plant Sci. 9:1575

Qiu, Y., Zhang, Y., Hu, F., and Zhu, S. 2017. Characterization of siRNAs derived from cucumber mosaic virus in infected tobacco plants. Arch. Virol. 162:2077–2082

Qiu, Y., Zhang, Y., Wang, C., Lei, R., Wu, Y., Li, X., and Zhu, S. 2018. *Cucumber mosaic virus* coat protein induces the development of chlorotic symptoms through interacting with the chloroplast ferredoxin I protein. Sci. Rep. 8:1205

Ramesh, S. V, Williams, S., Kappagantu, M., Mitter, N., and Pappu, H. R. 2017. Transcriptome-wide identification of host genes targeted by tomato spotted wilt virus-derived small interfering RNAs. Virus Res. 238:13–23

Robaglia, C., and Caranta, C. 2006. Translation initiation factors: a weak link in plant RNA virus infection. Trends Plant Sci. 11:40–45

Rodríguez-Cerezo, E., Klein, P. G., and Shaw, J. G. 1991. A determinant of disease symptom severity is located in the 3'-terminal noncoding region of the RNA of a plant virus. Proc. Natl. Acad. Sci. 88:9863–9867

Rong, W., Wang, X., Wang, X., Massart, S., and Zhang, Z. 2018. Molecular and ultrastructural mechanisms underlying yellow dwarf symptom formation in wheat after infection of barley yellow dwarf virus. Int. J. Mol. Sci. 19:1187

Salánki, K., Kiss, L., Gellért, Á., and Balázs, E. 2011. Identification a coat protein region of cucumber mosaic virus (CMV) essential for long-distance movement in cucumber. Arch. Virol. 156:2279–2283

Sánchez, F., Manrique, P., Mansilla, C., Lunello, P., Wang, X., Rodrigo, G., López-González, S., Jenner, C., González-Melendi, P., Elena, S. F., Walsh, J., and Ponz, F. 2015. Viral strain-specific differential alterations in *Arabidopsis* developmental patterns. Mol. Plant-Microbe Interact. 28:1304–1315

Selman, I., and Grant, S. A. 1957. The influence of temperature and daylength on spotted wilt virus disease of tomato. Ann. Appl. Biol. 45:312–317

Seo, J.-K., Kwak, H.-R., Choi, B., Han, S.-J., Kim, M.-K., and Choi, H.-S. 2017. Movement protein of broad bean wilt virus 2 serves as a determinant of symptom severity in pepper. Virus Res. 242:141–145

Shi, B., Lin, L., Wang, S., Guo, Q., Zhou, H., Rong, L., Li, J., Peng, J., Lu, Y., Zheng, H., Yang, Y., Chen, Z., Zhao, J., Jiang, T., Song, B., Chen, J., and Yan, F. 2016. Identification and regulation of host genes related to Rice stripe virus symptom production. New Phytol. 209:1106–1119

Shimura, H., Pantaleo, V., Ishihara, T., Myojo, N., Inaba, J., Sueda, K., Burgyán, J., and Masuta, C. 2011. A viral satellite RNA induces yellow symptoms on tobacco by targeting a gene involved in chlorophyll biosynthesis using the RNA silencing machinery. PLoS Pathog. 7:e1002021

Smith, N. A., Eamens, A. L., and Wang, M.-B. 2011. Viral small interfering RNAs target host genes to mediate disease symptoms in plants. PLoS Pathog. 7:e1002022

Tatineni, S., Elowsky, C., and Graybosch, R. A. 2017. *Wheat streak mosaic virus* coat protein deletion mutants elicit more severe symptoms than wild-type virus in multiple cereal hosts. Mol. Plant-Microbe Interact. 30:974–983

Tatineni, S., McMechan, A. J., and Hein, G. L. 2018. *Wheat streak mosaic virus* coat protein is a determinant for vector transmission by the wheat curl mite. Virology 514:42–49

del Toro, F. J., Rakhshandehroo, F., Larruy, B., Aguilar, E., Tenllado, F., and Canto, T. 2017. Effects of simultaneously elevated temperature and CO_2 levels on *Nicotiana benthamiana* and its infection by different positive-sense RNA viruses are cumulative and virus type-specific. Virology 511:184–192

Vigne, E., Gottula, J. W., Schmitt-Keichinger, C., Komar, V., Ackerer, L., Belval, L., Rakotomalala, L., Lemaire, O., Ritzenthaler, C., and Fuchs, M. 2013. A strain-specific segment of the RNA-dependent RNA polymerase of grapevine fanleaf virus determines symptoms in *Nicotiana* species. J. Gen. Virol. 94:2803–2813

Wang, S., Cui, W., Wu, X., Yuan, Q., Zhao, J., Zheng, H., Lu, Y., Peng, J., Lin, L., Chen, J., and Yan, F. 2018. Suppression of nbe-miR166h-p5 attenuates leaf yellowing symptoms of potato virus X on *Nicotiana benthamiana* and reduces virus accumulation. Mol. Plant Pathol. 19:2384–2396

Wieczorek, P., and Obrępalska-Stęplowska, A. 2016. The N-terminal fragment of the tomato torrado virus RNA1-encoded polyprotein induces a hypersensitive response (HR)-like reaction in *Nicotiana benthamiana*. Arch. Virol. 161:1849–1858

Wosula, E. N., Tatineni, S., Wegulo, S. N., and Hein, G. L. 2016. Effect of temperature on wheat streak mosaic disease development in winter wheat. Plant Dis. 101:324–330

Xu, Y., Da Silva, W. L., Qian, Y., and Gray, S. M. 2018. An aromatic amino acid and associated helix in the C-terminus of the potato leafroll virus minor capsid protein regulate systemic infection and symptom expression. PLoS Pathog. 14:1–30

Yang, C., Guo, R., Jie, F., Nettleton, D., Peng, J., Carr, T., Yeakley, J. M., Fan, J.-B., and Whitham, S. A. 2007. Spatial analysis of *Arabidopsis thaliana* gene expression in response to *Turnip mosaic virus* infection. Mol. Plant-Microbe Interact. 20:358–370

Yang, Y., Liu, T., Shen, D., Wang, J., Ling, X., Hu, Z., Chen, T., Hu, J., Huang, J., Yu, W., Dou, D., Wang, M.-B., and Zhang, B. 2019. Tomato yellow leaf curl virus intergenic siRNAs target a host long noncoding RNA to modulate disease symptoms. PLoS One 15:e1007534

Zavallo, D., Debat, H. J., Conti, G., Manacorda, C. A., Rodriguez, M. C., and Asurmendi, S. 2015. Differential mRNA accumulation upon early *Arabidopsis thaliana* infection with ORMV and TMV-Cg is associated with distinct endogenous small RNAs level. PLoS One 10:e0134719

Zhan, B., Zhao, W., Li, S., Yang, X., and Zhou, X. 2018. Functional scanning of apple geminivirus proteins as symptom determinants and suppressors of posttranscriptional gene silencing. Viruses 10:488

Zhang, H., and Sonnewald, U. 2017. Differences and commonalities of plant responses to single and combined stresses. Plant J. 90:839–855

Zhang, X.-P., Liu, D.-S., Yan, T., Fang, X.-D., Dong, K., Xu, J., Wang, Y., Yu, J.-L., and Wang, X.-B. 2017. Cucumber mosaic virus coat protein modulates the accumulation of 2b protein and antiviral silencing that causes symptom recovery *in planta*. PLoS Pathog. 13:e1006522

Zhao, F., Li, Y., Chen, L., Zhu, L., Ren, H., Lin, H., and Xi, D. 2016. Temperature dependent defence of *Nicotiana tabacum* against *Cucumber mosaic virus* and recovery occurs with the formation of dark green islands. J. Plant Biol. 59:293–301

Zhong, X., Wang, Z. Q., Xiao, R., Cao, L., Wang, Y., Xie, Y., and Zhou, X. 2017. Mimic phosphorylation of a βC1 protein encoded by TYLCCNB impairs its functions as a viral suppressor of RNA silencing and a symptom determinant. J. Virol. 91:e00300-17

CHAPTER 2

AN OVERVIEW OF GRAPEVINE FANLEAF VIRUS BIOLOGY

GRAPEVINE FANLEAF VIRUS AND FANLEAF DEGENERATION DISEASE

Grapevine fanleaf virus (GFLV) is one of several viruses in the genus *Nepovirus,* family *Secoviridae,* which causes fanleaf degeneration disease of grapevine (Digiaro et al. 2017). Fanleaf degeneration is traditionally delineated into two distinct syndromes (sets of symptoms): infectious malformations and yellow mosaic (Digiaro et al. 2017). The former is characterized by marked developmental malformations, including double nodes, shortened internodes, fasciations, and zigzag growth. Foliar symptoms include cupping, asymmetrical leaf shape, and the namesake "fanleaf" consisting of abnormally open petiolar sinuses and aberrant vein spacing in a radial pattern, giving leaves the appearance of an open fan. These symptoms are accompanied by poor and uneven berry set, leading to yield losses. In the latter syndrome, malformations are less prominent and the distinguishing symptoms include chrome-yellow foliar discolorations, including vein banding, mottling, and mosaics. Infectious malformation and yellow mosaic are caused by distorting and chromogenic strains of the virus, respectively (Digiaro et al. 2017). Though some work has been done to identify correlations between genetic elements of the virus and the syndromes caused by distinct strains (Elbeaino et al. 2014), the molecular mechanisms underpinning symptom development remain unknown.

The term "nepovirus" derives from nematode-vectored polyhedral virus. GFLV is a nepovirus that is exclusively transmitted by the ectoparasitic dagger nematode *Xiphinema index* in a non-circulative, semi-persistent manner (Hewitt et al.

1958). GFLV virions consist of 30 nm icosahedral particles made up of 60 copies of the coat protein encapsulating separately two genomic RNAs. Vector-virus specificity between *X. index* and GFLV is determined by the coat protein (Andret-Link et al. 2004; Schellenberger et al. 2010). GFLV has been found in most major grape production areas worldwide. However, while GFLV can be mechanically passaged to experimental herbaceous hosts with relative ease, mechanical inoculation of grapevines is all but impossible (Valat et al. 2003) and spread of GFLV by seed in grapevine is economically insignificant since grapevine is almost exclusively vegetatively propagated (Digiaro et al. 2017). Thus, GFLV's impact is greatest in those viticulture regions in which *X. index* is endemic and able to spread the virus to healthy vines.

TAXONOMIC CLASSIFICATION OF GRAPEVINE FANLEAF VIRUS

The genus *Nepovirus* lies within the family *Secoviridae* of the *Picornavirales* order. This order includes the extensively-studied poliovirus and rhinoviruses; in some cases, the functions of some nepovirus proteins, as well as molecular nomenclature, have been imputed or borrowed from these more (in)famous cousins. Nepoviruses are classified into subgroups based on the length of RNA2 and phylogenetic relationships between the coat protein ($2C^{CP}$) sequences and a region spanning the protease and polymerase (Pro-Pol) coding sequences (Digiaro et al. 2017; Fuchs et al. 2017). GFLV is in subgroup A, along with the closely related species *Arabis mosaic virus* (ArMV), *Grapevine deformation virus* (GDeV), *Tobacco ringspot virus* (TRSV), and *Raspberry ringspot virus* (RpRSV). This subgroup is characterized by having the shortest RNA2 sequences of all the nepoviruses. The type species of subgroups B and C are *Tomato*

black ring virus (TBRV) and *Tomato ringspot virus* (ToRSV), respectively, which sport progressively longer RNA2 sequences.

GENOME ORGANIZATION AND EXPRESSION OF GRAPEVINE FANLEAF VIRUS

Like all nepoviruses, GFLV has a bipartite, positive-sense RNA genome (Figure 2-1). The genomic RNAs are termed RNA1 and RNA2. Each RNA contains a single open reading frame encoding a large polyprotein (P1 and P2, respectively), flanked on either end by 5' and 3' untranslated regions (UTRs). The 5' end is covalently linked to the RNA1-encoded VPg (viral protein, genome linked), which, based on analogy to other members of *Picornavirales*, is assumed to function in RNA priming during genome replication (Steil and Barton 2009). The 3'UTR is proceeded by a poly(A) tail. These features enable the genomic RNAs to function as mRNAs for direct translation by host ribosomes for the production of viral proteins. Like other members of *Picornavirales*, GFLV genome expression is by monocistronic translation, followed by proteolytic processing into mature peptides by an internal (virus-encoded) protease (Margis et al. 1991, 1994).

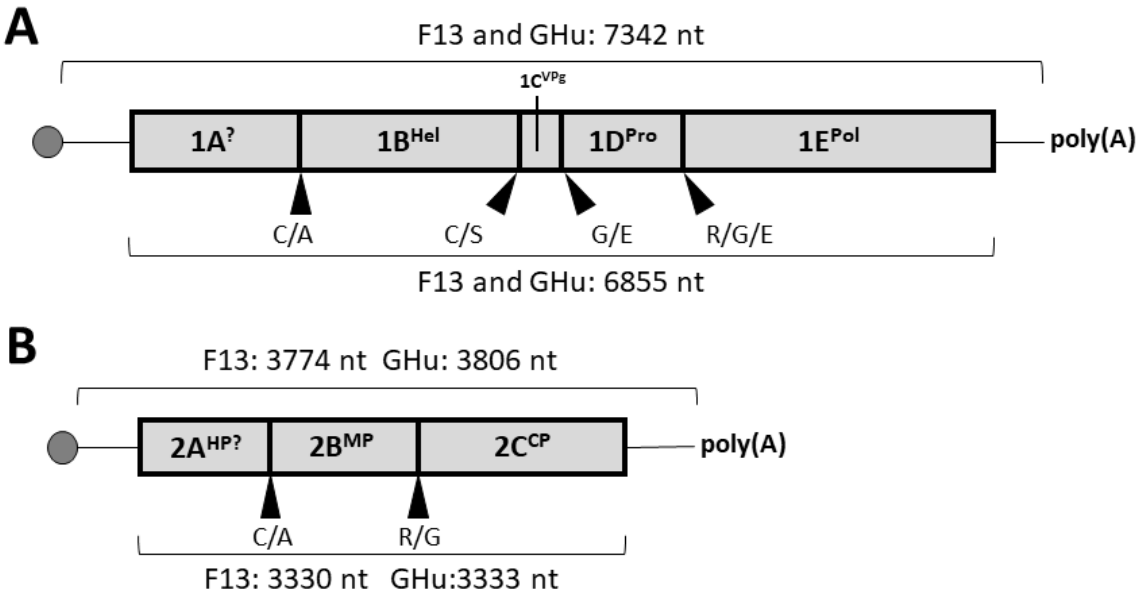


Figure 2-1. Genomic structure of grapevine fanleaf virus (GFLV) genomic RNA1 **(A)** and RNA2 **(B).** Open boxes represent open reading frames (ORFs) with protein maturation products labelled individually: ?, protein of suspected function for which experimental evidence is lacking; Hel, putative helicase; VPg, viral protein, genome-linked; Pro, protease; Pol, RNA-dependent RNA polymerase; HP?, putative homing protein; MP, movement protein; CP, coat protein. Thin lines extending from ORFs designate 5' and 3' untranslated regions. Grey circles represent VPg covalently linked to the genomic RNAs. Black triangles indicate sites of cleavage of the polyprotein by the $1D^{Pro}$ protease with the residues of the cleavage sites listed below. Sequence lengths for full genomic RNAs and ORFs are listed above and below each RNA, respectively, for both GFLV strains F13 and GHu.

RNA1 encodes all viral products necessary for genome replication and thus is capable of independent replication in protoplasts in the absence of RNA2 (Viry et al. 1993). Proteolytic processing of P1 produces five mature peptides (Figure 2-1A). The function of protein 1A has not been conclusively determined. In ToRSV (subgroup C), the 1A analog X1 is further processed into X2 and NTB (nucleotide triphosphate binding) and both play a role in membrane restructuring for formation of the endoplasmic reticulum-derived replication complex (Wang and Sanfaçon 2000; Sanfacon 2013); however, no evidence of a similar mechanism could be found in GFLV (Margis et al. 1994). Protein $1B^{Hel}$ possesses the hallmark motifs of a

Picornavirales helicase and is theorized to function as such (Candresse et al. 1990; Ritzenthaler et al. 1991). Protein $1C^{VPg}$ is the aforementioned genome-linked viral protein which becomes covalently linked to the 5' end of the genomic RNAs (Pinck et al. 1991). Protein $1D^{Pro}$ is a cysteine protease which recognizes and cleaves at cysteine-alanine, cysteine-serine, glycine-glutamic acid, and arginine-glycine sites on P1 and P2 in cis and trans, respectively, to produce mature peptides; it is likely that other residues flanking these dipeptides aid in the specific recognition of these sites by $1D^{Pro}$ (Margis et al. 1991). The $1E^{Pol}$ protein is an 824 residue protein containing the seven canonical motifs of a superfamily I-type RNA-dependent RNA polymerase (RdRp) (Vigne et al. 2013). These seven conserved RdRp motifs lie within the N-terminal 433 residues of GFLV $1E^{Pol}$; this is consistent with the minimal RdRps of picornaviruses, such as the 461 residue poliovirus RdRp (Hansen et al. 1997). The C-terminal extension of the GFLV $1E^{Pol}$ is the site of a symptom determinant for vein clearing in *N. benthamiana* (Vigne et al. 2013; Chapter 4). As with the processing of X1 into X2 and NTB, ToRSV has been shown to produce an additional mature peptide from the C-terminus of $1E^{Pol}$ (Chisholm et al. 2007). However, no evidence of proteolytic cleavage of $1E^{Pol}$ could be found for GFLV (Vigne et al. 2013).

While RNA1 is capable of independent replication, both RNAs must be present for local spread and systemic infection to occur (Viry et al. 1993). This is because RNA2 encodes proteins for cell-to-cell movement, as well as the coat protein. Proteolytic processing of P2 produces three mature peptides (Figure 2-1B). Protein $2A^{HP}$ has been termed a homing protein because it is necessary for RNA2 replication and is believed to function in targeting RNA2 to the site of viral replication (Gaire et

al. 1999). Protein $2B^{MP}$ is a movement protein of the tubule-forming type. Protein $2B^{MP}$ polymerizes to form tubules which extend through plasmodesmata. These tubules enable the shuttling of virions into adjacent cells, thus propagating the infection throughout the plant (Laporte et al. 2003). Protein $2C^{CP}$ is the coat protein and consists of three jelly-roll domains. Sixty copies of $2C^{CP}$ can be assembled into 30 nm isometric virions with pseudo T=3 symmetry (Schellenberger et al. 2011). Genomic RNAs are encapsulated separately (Quacquarelli et al. 1976). Structural predictions of assembled GFLV virions identified a positively charged cavity on the capsid surface (Schellenberger et al. 2011) that corresponds to residues previously identified as determining the specificity between the virus and its nematode vector. It is predicted that this pocket interacts with an as-yet-unidentified nematode component to enable uptake and retention of GFLV by *X. index* for transmission (Schellenberger et al. 2010).

SYMPTOM DETERMINANTS OF GRAPEVINE FANLEAF VIRUS

The natural host range of GFLV is relatively narrow, consisting of grapevine and a few weed species, including *Aristolochia clematitis, Lagenaria siceraria,* Bermuda grass, knotweed, raspberry, Johnson grass, *Melilotus* spp. and *Plantago lanceolate* (Izadpanah et al. 2003b; Horváth et al. 1994; Zaki-Aghl et al. 2015; Izadpanah et al. 2003a). However, these weedy hosts do not appear to play a significant role in GFLV epidemiology (Digiaro et al. 2017). GFLV has been successfully passaged to a range of herbaceous experimental hosts, including members of *Amaranthaceae, Chenodiaecae, Cucurbitaceae, Leguminoseae, Solanaceae,* and *Fabaceae* families (Schmitt-Keichinger et al. 2017).

How plant viruses cause symptoms in their hosts is a long-standing question in plant pathology. GFLV causes a variety of symptoms in grapevine. These symptoms are even delineated into distinct syndromes, infectious malformation and yellow mosaic, which are caused by distorting and yellowing strains of the virus, respectively. While distorting and yellowing strains can be separated by the symptoms they produce on grapevine (Elbeaino et al. 2014), we cannot yet say what sets them apart genetically. Some authors have pointed to evidence of recombination between GFLV and ArMV within $2A^{HP}$ as a frequent hallmark of yellowing strains; in the same vein, phylogenetic analyses of GFLV strains determined using $2A^{HP}$ sequences produced distinct groups for yellowing and distorting strains (Elbeaino et al. 2014). However, this approach has failed to identify specific genetic elements which distinguish yellowing from distorting strains, and mechanistic explanations for how such genetic differences give rise to different syndromes are lacking.

While work on GFLV symptom determinants in grapevine has been sparse, recent advancements have been made in identifying and characterizing symptom determinants in herbaceous hosts. These advancements have been driven, in part, by the fact that GFLV can be mechanically transmitted to herbaceous hosts, such as *Chenopodium* and *Nicotiana* species, with relative ease.

While its precise role in grapevine symptomatology remains inconclusive, the $2A^{HP}$ protein of GFLV strain F13 functions as an avirulence (avr) factor in *Nicotiana occidentalis* (Martin et al. 2018). Inoculation of *N. occidentalis* with GFLV-F13 elicits a genuine hypersensitive response (HR) on inoculated leaves and this response was mapped to the C-terminal 50 amino acids of the protein. HR can also be induced via

transient expression of the $2A^{HP}$ protein. Interestingly, this HR is not sufficient to prevent virus infection as the authors noted that some plants did become systemically infected with variants harboring the avr portion of the $2A^{HP}$ protein (Martin et al. 2018). Further work is needed to investigate the dynamics of HR establishment and determine which environmental factors play a role in the inconsistent ability of GFLV strain F13 to escape HR-mediated restriction in *N. occidentalis*.

While GFLV-F13 produces HR on *N. occidentalis*, this strain produces an asymptomatic infection in the closely-related species *Nicotiana benthamiana* (Figure 2-2B). In contrast, strain GFLV-GHu produces distinct vein clearing on the upper, uninoculated leaves of *N. benthamiana* (Figure 2-2C). The vein clearing determinant of GFLV-GHu had previously been mapped to the 3' end of the $1E^{Pol}$ coding region, though it was not determined if symptoms were determined by the nucleotide or amino acid sequence of this region (Vigne et al. 2013). The workflow for this research required primary infection of *Chenopodium quinoa* via mechanical inoculation with *in vitro* RNA transcripts of mutant GFLV infectious clones, followed by mechanical passage of viral progeny to *N. benthamiana* for symptom phenotyping. Thus the creation and characterization of new GFLV mutants was enormously laborious and time-consuming. Furthermore, the utility of this approach for reverse genetics was limited due to the lack of compatible restriction enzyme sites shared between GFLV-GHu and –F13 in this region of the $1E^{Pol}$ coding sequence.

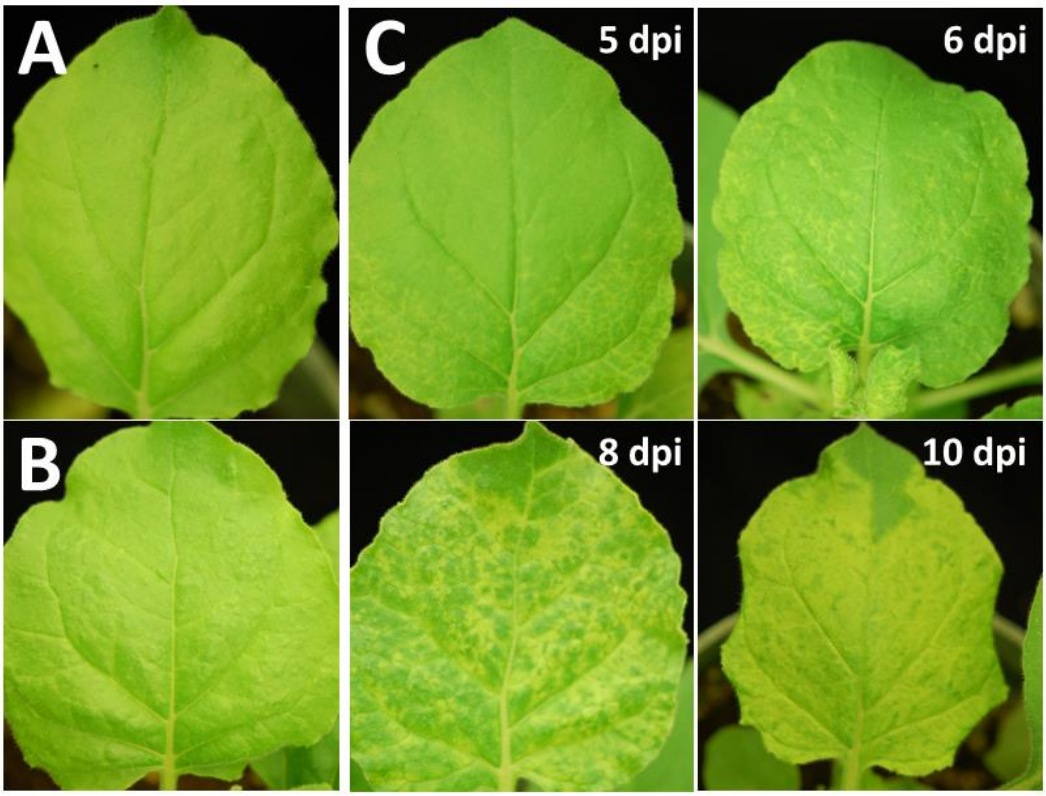


Figure 2-2. Phenotypes of systemic infection of *Nicotiana benthamiana* by strains F13 and GHu of grapevine fanleaf virus (GFLV). **A.** Mock inoculated *N. benthamiana.* **B.** Systemic infection of *N. benthamiana* by GFLV strain F13 is asymptomatic. **C.** Systemic infection of *N. benthamiana* by GFLV strain GHu produces vein clearing on upper, uninoculated leaves. Subpanels show the development and spread of symptoms over time at specific days post inoculation (dpi).

To circumvent these limitations, I have cloned GFLV cDNAs of strains GHu and F13 into plasmids compatible with *Agrobacterium tumefaciens*-mediated inoculation (agroinoculation). These vectors are also compatible with direct plasmid mutagenesis, which drastically increases the variety and range of mutations which can be introduced to the GFLV genome, while also significantly reducing the time-frame for the production of mutant constructs. Chapter 3 describes the development and optimization of this GFLV agroinoculation system for the establishment of systemic GFLV infection in *N. benthamiana*, including its utility for reverse genetics studies (Osterbaan et al. 2018).

Chapter 4 describes the use of the GFLV agroinoculation system for the identification and characterization of a single residue, 802, of GFLV-GHu $1E^{Pol}$ which modulates vein clearing in *N. benthamiana.* This residue is necessary, but not sufficient, to elicit vein clearing in *N. benthamiana* during GFLV infection and is found to be flanked by conserved, stable secondary structures of unknown function (Osterbaan et al. 2019).

While much is known about the molecular processes of GFLV infection (for review, see Schmitt-Keichinger et al. 2017), little is known about the molecular interactions in which $1E^{Pol}$ participates in symptom development. This leaves us with few clues as to how $1E^{Pol}$ residue 802 interacts with viral or host components to modulate vein clearing symptoms. To that end, Chapter 5 follows the identification of the GFLV-GHu symptom determinant with a preliminary study of the protein interactome of GFLV $1E^{Pol}$, which was probed via affinity purification and mass spectrometry. This study lays the groundwork for the dissection and characterization of $1E^{Pol}$ protein interactions, with the aim of identifying the partners involved in and elucidating the mechanism of symptom development.

Finally, in Chapter 6, I provide some perspectives on my research and outline potential future directions of GFLV-host interaction studies to advance our understanding of mechanisms underlying symptom development.

REFERENCES

Andret-Link, P., Schmitt-Keichinger, C., Demangeat, G., Komar, V., and Fuchs, M. 2004. The specific transmission of *Grapevine fanleaf virus* by its nematode vector *Xiphinema index* is solely determined by the viral coat protein. Virology 320:12–22

Candresse, T., Morch, M. D., and Dunez, J. 1990. Multiple alignment and hierachical clustering of conserved amino acid sequences in the replication-associated proteins of plant RNA viruses. Res. Virol. 141:315–329

Chisholm, J., Zhang, G., Wang, A., and Sanfaçon, H. 2007. Peripheral association of a polyprotein precursor form of the RNA-dependent RNA polymerase of *Tomato ringspot virus* with the membrane-bound viral replication complex. Virology 368:133–144

Digiaro, M., Elbeaino, T., and Martelli, G. P. 2017. *Grapevine fanleaf virus* and other Old World nepoviruses. Page 47–82 in: Grapevine Viruses: Molecular Biology, Diagnostics and Management, B. Meng, G.P. Martelli, D.A. Golino, and M. Fuchs, eds. Springer International Publishing, Cham.

Elbeaino, T., Kiyi, H., Boutarfa, R., Minafra, A., Martelli, G. P., and Digiaro, M. 2014. Phylogenetic and recombination analysis of the homing protein domain of grapevine fanleaf virus (GFLV) isolates associated with "yellow mosaic" and "infectious malformation" syndromes in grapevine. Arch. Virol. 159:2757–2764

Fuchs, M., Schmitt-Keichinger, C., and Sanfaçon, H. 2017. A renaissance in nepovirus research provides new insights into their molecular interface with hosts and vectors. Pages 61–105 in: Advances in Virus Research, M. Kielian, T.C. Mettenleiter, and M.J. Roossinck, eds. Academic Press, Cambridge, MA, USA.

Gaire, F., Schmitt, C., Stussi-Garaud, C., Pinck, L., and Ritzenthaler, C. 1999. Protein 2A of grapevine fanleaf nepovirus is implicated in RNA2 replication and colocalizes to the replication site. Virology 264:25–36

Hansen, J. L., Long, A. M., and Schultz, S. C. 1997. Structure of the RNA-dependent RNA polymerase of poliovirus. Structure 5:1109–1122

Hewitt, W., Raski, D., and Goheen, A. 1958. Nematode vector of soil-borne fanleaf virus of grapevines. Phytopathology 48:586–595

Horváth, J., Tóbiás, I., and Hunyadi, K. 1994. New natural herbaceous hosts of grapevine fanleaf nepovirus. Kertészeti Tudomány. 26:31–32

Izadpanah, K., Zaki-Aghl, M., Zhang, Y., Daubert, S., and Rowhani, A. 2003a. Non-vitis hosts of grapevine fanleaf virus and their possible epidemiological significance. Pages 210 in: Extended abstracts of the 14th Meeting of ICVG, Locorotondo-Bari, Italy.

Izadpanah, K., Zaki-Aghl, M., Zhang, Y. P., Daubert, S. D., and Rowhani, A. 2003b. Bermuda grass as a potential reservoir host for *Grapevine fanleaf virus*. Plant Dis. 87:1179–1182

Laporte, C., Vetter, G., Loudes, A.-M., Robinson, D. G., Hillmer, S., Stussi-Garaud, C., and Ritzenthaler, C. 2003. Involvement of the secretory pathway and the cytoskeleton in intracellular targeting and tubule assembly of *Grapevine fanleaf virus* movement protein in tobacco BY-2 cells. Plant Cell 15:2058–2075

Margis, R., Viry, M., Pinck, M., Bardonnet, N., and Pinck, L. 1994. Differential proteolytic activities of precursor and mature forms of the 24K proteinase of Grapevine fanleaf nepovirus. Virology 200:79–86

Margis, R., Viry, M., Pinck, M., and Pinck, L. 1991. Cloning and *in vitro* characterization of the grapevine fanleaf virus proteinase cistron. Virology 185:779–787

Martin, I. R., Vigne, E., Berthold, F., Komar, V., Lemaire, O., Fuchs, M., and Schmitt-Keichinger, C. 2018. The 50 distal amino acids of the $2A^{HP}$ homing protein of *Grapevine fanleaf virus* elicit a hypersensitive reaction on *Nicotiana occidentalis*. Mol. Plant Pathol. 19:731–743

Osterbaan, L. J., Choi, J., Kenney, J., Flasco, M., Vigne, E., Schmitt-Keichinger, C., Rebelo, A. R., Heck, M., and Fuchs, M. 2019. The identity of a single residue of the RNA-dependent RNA polymerase of grapevine fanleaf virus modulates vein clearing in *Nicotiana benthamiana*. Mol. Plant-Microbe Interact. 32:790–801

Osterbaan, L. J., Schmitt-Keichinger, C., Vigne, E., and Fuchs, M. 2018. Optimal systemic grapevine fanleaf virus infection in *Nicotiana benthamiana* following agroinoculation. J. Virol. Methods. 257:16–21

Pinck, M., Reinbolt, J., Loudes, A. M., Le Ret, M., and Pinck, L. 1991. Primary structure and location of the genome-linked protein (VPg) of grapevine fanleaf nepovirus. FEBS Lett. 284:117–119

Quacquarelli, A., Gallitelli, D., Savino, V., and Martelli, G. P. 1976. Properties of Grapevine fanleaf virus. J. Gen. Virol. 32:349–360

Ritzenthaler, C., Viry, M., Pinck, M., Margis, R., Fuchs, M., and Pinck, L. 1991. Complete nucleotide sequence and genetic organization of grapevine fanleaf nepovirus RNA1. J. Gen. Virol. 72:2357–2365

Sanfacon, H. 2013. Investigating the role of viral integral membrane proteins in promoting the assembly of nepovirus and comovirus replication factories. Front. Plant Sci. 3:313

Schellenberger, P., Andret-Link, P., Schmitt-Keichinger, C., Bergdoll, M., Marmonier, A., Vigne, E., Lemaire, O., Fuchs, M., Demangeat, G., and Ritzenthaler, C. 2010. A stretch of 11 amino acids in the βB- βC loop of the coat protein of *Grapevine fanleaf virus* is essential for transmission by the nematode *Xiphinema index*. J. Virol. 84:7924–7933

Schellenberger, P., Sauter, C., Lorber, B., Bron, P., Trapani, S., Bergdoll, M., Marmonier, A., Schmitt-Keichinger, C., Lemaire, O., Demangeat, G., and Ritzenthaler, C. 2011. Structural insights into viral determinants of nematode mediated Grapevine fanleaf virus transmission. PLoS Pathog. 7:e1002034

Schmitt-Keichinger, C., Hemmer, C., Berthold, F., and Ritzenthaler, C. 2017. Molecular, Cellular, and Structural Biology of *Grapevine fanleaf virus*. Pages 83–107 in: Grapevine Viruses: Molecular Biology, Diagnostics and Management, B. Meng, G.P. Martelli, D.A. Golino, and M. Fuchs, eds. Springer International Publishing, Cham.

Steil, B. P., and Barton, D. J. 2009. Cis-active RNA elements (CREs) and picornavirus RNA replication. Virus Res. 139:240–252

Valat, L., Burrus, M., Fuchs, M., and Mauro, M. C. 2003. Review of techniques to inoculate grapevines with Grapevine fanleaf virus: lessons and perspectives. Am. J. Enol. Vitic. 54:279–285

Vigne, E., Gottula, J. W., Schmitt-Keichinger, C., Komar, V., Ackerer, L., Belval, L., Rakotomalala, L., Lemaire, O., Ritzenthaler, C., and Fuchs, M. 2013. A strain-specific segment of the RNA-dependent RNA polymerase of grapevine fanleaf virus determines symptoms in *Nicotiana* species. J. Gen. Virol. 94:2803–2813

Viry, M., Serghini, M. A., Hans, F., Ritzenthaler, C., Pinck, M., and Pinck, L. 1993. Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. J. Gen. Virol. 74:169–174

Wang, A., and Sanfaçon, H. 2000. Proteolytic processing at a novel cleavage site in the N-terminal region of the tomato ringspot nepovirus RNA-1-encoded polyprotein *in vitro*. J. Gen. Virol. 81:2771–2781

Zaki-Aghl, M., Izadpanah, K., Gholampour, Z., Kargar, M., and Mehrvar, M. 2015. Molecular characterization of grapevine fanleaf virus from non *Vitis* hosts. Pages 149–150 in: Extended abstracts of the 18th Meeting of ICVG, Ankara, Turkey.

CHAPTER 3

OPTIMAL SYSTEMIC GRAPEVINE FANLEAF VIRUS INFECTION IN *NICOTIANA BENTHAMIANA* FOLLOWING AGROINOCULATION[2]

ABSTRACT

One of the greatest hindrances to the study of grapevine fanleaf virus (GFLV) is the dearth of robust protocols for reliable, scalable, and cost-effective inoculation of host plants, especially methods which allow for rapid and targeted manipulation of the virus genome. Agroinoculation fulfills these requirements: it is a relatively rapid, inexpensive, and reliable method for establishing infections, and enables genetic manipulation of viral sequences by modifying plasmids. We designed a system of binary plasmids based on the two genomic RNAs [RNA1 (1) and RNA2 (2)] of GFLV strains F13 (F) and GHu (G) and optimized parameters to maximize systemic infection frequency in *Nicotiana benthamiana* via agroinoculation. The genomic make-up of the inoculum (G1-G2 and reassortant F1-G2), the identity of the co-infiltrated silencing suppressor (grapevine leafroll associated virus 2 p24), and temperature at which plants were maintained (25 °C) significantly increased systemic infection, while high optical densities of infiltration cultures (OD_{600nm} of 1.0 or 2.0) increased the consistency of systemic infection frequency in *N. benthamiana.* In contrast, acetosyringone in the bacterial culture media, regardless of concentration, had no effect. Plasmids in this system are amenable to rapid and reliable manipulation by one-step site-directed

[2]This chapter was published as: Osterbaan, L. J., Schmitt-Keichinger, C., Vigne, E., and Fuchs, M. 2018. Optimal systemic grapevine fanleaf virus infection in *Nicotiana benthamiana* following agroinoculation. J. Virol. Methods. 257:16–21. doi:10.1016/j.jviromet.2018.04.006. Section titles not appearing in the published manuscript have been added here for increased readability.

mutagenesis, as shown by the creation of infectious RNA1 chimeras of the GFLV-F13 and GHu strains. The GFLV agroinoculation plasmids described here, together with the optimized protocol for bacterial culturing and plant maintenance, provide a robust system for the establishment of systemic GFLV infection in *N. benthamiana* and the rapid generation of GFLV mutants, granting a much-needed tool for investigations into GFLV-host interactions.

INTRODUCTION

Grapevine fanleaf virus (GFLV), from the genus *Nepovirus* in the family *Secoviridae,* is one of several nepoviruses that cause fanleaf degeneration, one of the most devastating viral diseases of grapevine worldwide (Andret-Link et al. 2004). In addition to yield losses due to poor and uneven berry set as well as premature vine death, GFLV elicits a number of foliar symptoms in grapevine, including vein yellowing, internode shortening, mosaics, and the eponymous fanleaf symptom characterized by abnormally open petiolar sinuses and reduced lobing (Andret-Link et al. 2004; Digiaro et al. 2017; Schmitt-Keichinger et al. 2017).

GFLV is a subgroup A nepovirus and has a bipartite (+) sense RNA genome. Each RNA contains a single open reading frame (ORF) which is translated as a single polyprotein matured into functional products by proteolytic processing (Fuchs et al. 2017; Schmitt-Keichinger et al. 2017). The function of most mature GFLV translation products is well established. RNA1 encodes five proteins (1A-1E) for genome replication and polyprotein processing. No function has yet been assigned to the 1A protein; protein $1B^{Hel}$ is a putative helicase (Ritzenthaler et al. 1991); protein $1C^{VPg}$ is a genome-linked viral protein (VPg) which covalently attaches to the 5' end of both genomic RNAs (Pinck et al. 1991); protein $1D^{Pro}$ is a protease, which processes the polyproteins produced from the ORFs of RNA1 and RNA2 into their individual maturation products (Margis et al. 1991); and protein $1E^{Pol}$ is an RNA-dependent RNA polymerase (Ritzenthaler et al. 1991). RNA2 encodes three proteins for genome encapsidation, viral movement, and replication (Fuchs et al. 2017). Protein $2A^{HP}$ is designated as a putative homing protein since it is necessary for localizing RNA2 to

the cellular site of virus replication and is indispensable for RNA2 replication (Gaire et al., 1999b); protein $2B^{MP}$ is a movement protein of the tubule-forming type (Ritzenthaler et al. 1995); and protein $2C^{CP}$ is the coat protein (Serghini et al. 1990). GFLV genomic RNAs are separately encapsidated by a pseudo T=3 icosahedron made up of 60 identical copies of the coat protein (Quacquarelli et al. 1976; Wood 1998; Schellenberger et al. 2011). While RNA1 is capable of independent replication, plant systemic infection can only occur if both genomic RNAs are present (Viry et al. 1993).

As with many plant viruses, investigations of GFLV-host interactions are hampered by a lack of reliable, scalable, and cost-effective methods for plant inoculation (Valat et al. 2003), especially methods which also allow for the rapid, targeted manipulation of the virus genome. While mechanical inoculation with infectious crude sap or purified virus preparations has a high infection frequency (near 100%), neither process allows for direct, targeted manipulation of the virus genome. Inoculation using viruliferous *Xiphinema index* nematodes, the vector of GFLV, has a certain attraction for its biological relevance (Valat et al. 2003) and is necessary for investigations of vector-virus interactions (Schellenberger et al. 2010). However, inoculation with viruliferous nematodes requires the maintenance of nematode colonies, which is both laborious and costly. Additionally, such colonies may represent a biological hazard and special permits may be required to maintain colonies in regions to which they, or the virus they vector, are not endemic. Rub inoculation with *in vitro* RNA transcripts of cDNAs to establish infections *in planta* was a boon to plant virus research, including GFLV. Viral genomic cDNA borne on plasmids made

directed manipulation of viral sequences possible, enabling a burst of reverse genetic investigations into the function of viral proteins and genetic sequences (Belin et al. 1999; Gaire et al. 1999; Martin et al. 2018; Schellenberger et al. 2010; Vigne et al. 2013). However, *in vitro* RNA transcription remains a delicate and expensive process. Furthermore, this method is less than ideal for studies requiring symptom monitoring as *in vitro* transcription provides sufficient inoculum for only 1-2 plants, requiring a secondary passage to establish a population sufficient for reliable symptom monitoring, further extending the experiment timeline (Vigne et al. 2013).

Agroinoculation -the use of *Agrobacterium tumefaciens* cultures to deliver viral cDNA sequences via T-DNA transfer to establish viral infection- is a commonly used technique for viral inoculation of host plants (Vaghchhipawala et al. 2011). An agroinoculation system for GFLV, which includes a set of plasmids containing the cDNAs of RNA1 and RNA2 of GFLV strains F13 (Ritzenthaler et al. 1991) and GHu (Huss *et al.*, 1989) in the T-DNA region of pGA482G (An 1986), was developed previously (Gottula, pg. 77-121, 2014). While an improvement over the previous *in vitro* RNA transcription protocol in terms of ease-of-use, this first-generation GFLV agroinoculation system suffered from low infection frequency. Furthermore, site-directed mutagenesis of these plasmids still relied on a laborious two-step cloning process (Gottula, pg. 89, 2014). Thus, a second generation GFLV agroinoculation system was developed. Here we report on the engineering of the cDNAs of the two GFLV genomic RNAs of strains F13 and GHu into binary plasmids, and on the optimization of agroinoculation of *Nicotiana benthamiana* for consistently high systemic infection frequency. To demonstrate the versatility of this second-generation

GFLV agroinoculation system for reverse genetics, infectious RNA1 chimeras of GFLV strains GHu and F13, with regions of up to 90 bp mutated with a single mutagenic PCR protocol, were produced.

MATERIALS AND METHODS

Cloning GFLV genomic cDNAs into pCLEAN-G181

Full-length GFLV genomic cDNAs in a cauliflower mosaic virus (CaMV) 35S expression cassette were amplified by PCR from the first-generation pGA482G plasmids (Gottula, pg. 77-119, 2014) using specific primers (Table 3-1) and Phusion polymerase (New England Biolabs). Amplicons were cloned by restriction enzyme digest (Table 3-1) and ligation into the multiple cloning site (MCS) of pCLEAN-G181 (Thole et al. 2007). Plasmid pCLEAN-G181 contains a minimal T-DNA region, which consists of duplicate left border sequences with a 26 nt spacer, a 52 bp MCS site, and a single right border (RB) sequence followed by an RB overdrive sequence. Complete plasmid sequencing of the resulting constructs- pCLEAN-GHu-1, pCLEAN-GHu-2, pCLEAN-F13-1, and pCLEAN-F13-2 -was performed at the Center of Computational and Integrative Biology DNA Core Facility at Massachusetts General Hospital (Cambridge, MA). The plasmid backbone and GFLV cDNA sequences were aligned to GenBank reference sequences available for pCLEAN-G181 (EU186083.1), and GFLV strains GHu (JN391442 [RNA1] and EF426852 [RNA2]) and F13 (NC_003615 [RNA1] and NC_003623 [RNA2]) using Clustal W in MegAlign 14 (DNASTAR Lasergene 14). The plasmid backbone (Thole et al., 2007) sequences of all plasmids were found to be as expected. While some differences were found between the GFLV cDNA sequences and their GenBank reference sequences, they

were identical to those identified by Vigne et al. (2013), confirming that the cDNA sequences were successfully cloned without further mutation.

Agroinoculation of **Nicotiana benthamiana** *with GFLV constructs*

Sequencing-confirmed, purified (Omega Bio-tek E.Z.N.A. plasmid mini kit) plasmids pCLEAN-GHu-1 and pCLEAN-GHu-2 were electroporated separately into electrocompetent *A. tumefaciens* GV-S161 cells: GV3101::pMP90 cells (Hellens et al. 2000) harboring the pCLEAN-S161 helper plasmid (Thole et al. 2007). Transformed cells were prepared for agroinfiltration according to Annamalai and Rao (2005). Cultures of pCLEAN-GHu-1, pCLEAN-GHu-2, and pGA-p24, the viral silencing suppressor (VSR) p24 from grapevine leafroll-associated virus 2 (GLRaV-2) (Chiba et al. 2006), within the binary plasmid pGA482G, were adjusted to OD_{600nm} = 1.0 with infiltration buffer and combined in a 1:1:1 ratio to form the inoculum designated G1-G2. Cultures of pCLEAN-F13-1 and pCLEAN-F13-2 were similarly prepared with pGA-p24 to form GFLV inoculum F1-F2. Cultures were also assorted to form inocula F1-G2 and G1-F2. Each GFLV inoculum was infiltrated into at least five *N. benthamiana* seedlings, planted in 10 cm plastic pots in Cornell potting mix (a mixture of peat, vermiculite, ground limestone, and Uni-mix 0-20-5) at the 6-8 leaf stage with a needle-less syringe. Inoculated plants were maintained in growth chambers (25 °C, 50% relative humidity, 16-hour photoperiod). Plants were fertilized once a week with Osmocote®. At 3-4 weeks post inoculation, apical (uninoculated) leaves were collected and tested for systemic GFLV infection by DAS-ELISA using specific GFLV antibodies (BIOREBA, Reinach, Switzerland). Samples were considered

positive if the absorbance (OD_{405nm}) of both duplicate wells were greater than three times the average absorbance of the negative control (healthy *N. benthamiana* tissue).

Statistical analysis

The MIXED procedure of SAS 9.4 (SAS Institute, Cary, NC) without a link function was used to determine which treatments (virus genomic makeup, temperature at which plants are maintained, identity of co-infiltrated viral silencing suppressor, optical densities of infiltration cultures and concentration of acetosyringone in culture media) had a significant effect on incidence of GFLV systemic infection in *N. benthamiana* following agroinoculation. Each parameter was considered a fixed effect and the independent agroinoculations (experimental replicates) were considered random effects in the model. Differences between treatments were determined by the least squares means statement of the MIXED procedure at the 5% level of significance.

Table 3-1. Templates, primers, and restriction enzymes used to engineer GFLV agroinoculation constructs

Construct	Template	Primers	Primer sequence 5'-3'[a]	Restriction sites
pCLEAN-GHu-1	pGA482G-GHu-1	GHu1topCLEANG181-F	CTTC**TTAATTAA**ACCGTTAACGAGCTCTAG	PacI
		GHu1topCLEANG181-R	CTTCC**GTACGC**GTTGATGAAGCTTCTAG	BsiWI
pCLEAN-GHu-2	pGA482G-GHu-2	GHu2toCLEANG181-F2	TAAGC**GCGGCCGC**GTTGATGAAGCTTCTAGA	NotI
		GHu2toCLEANG181-R3	TAAGC**GCGGCCGC**TACCGTTAACGAGCTCTAG	NotI
pCLEAN-F13-1	pGA482G-F13-1	F131toCLEANG181-F	GTAAGC**GCGGCCGC**GCGCGAGGCAGCAGATCT	NotI
		F131toCLEANG181-R	TAAGCA**TTAATTAA**ATCGATAGATCTTAGTAC	PacI
pCLEAN-F13-2	pGA482G-F13-2	F132toCLEANG181-R	CTTC**CTCGAG**CCGAAACGCGCGAGGCAG	XhoI
		GHu2toCLEANG181-F2	TAAGC**GCGGCCGC**GTTGATGAAGCTTCTAGA	NotI

[a] Underlined nucleotides represent restriction sites

RESULTS AND DISCUSSION

Genomic makeup of GFLV inoculum affects agroinoculation success

In triplicated experiments, the GFLV inoculum F1-G2 consistently provided high rates of systemic infection in *N. benthamiana*, and the inoculum F1-F2 consistently produced low levels of systemic infection via agroinoculation (Figure 3-1). These results were consistent with those reported previously (Gottula, pg. 120-165, 2014). The infection frequency of the G1-G2 inoculum was not statistically different from that of F1-G2, while G1-F2 could not be statistically differentiated from either the high performing (F1-G2 and G1-G2) or low performing (F1-F2) inocula on *N. benthamiana* (Figure 3-1). The fact that G1-F2 inoculum had an intermediate behavior compared to the high-performing and low-performing inocula is most likely due to the high level of variability seen with its infectivity (0%-100%). Furthermore, the small sample sizes in the first two replicates (n = 6, 5, and 8 for G1-G2 and F1-G2; n = 6, 5, and 14 for F1-F2 and G1-F2) may have prevented us from detecting a significant difference between the treatments. Additional experiments are needed to address the amount of variability in the G1-F2 treatment. Reassortment of genomic RNAs is common amongst GFLV isolates due to the high incidence of mixed infections of distinct genetic variants (Vigne et al. 2004; Naraghi-Arani et al. 2001). Thus, the fact that every RNA1-RNA2 reassortant of GFLV strains GHu and F13 can establish systemic infection in *N. benthamiana* was expected. However, the highest performing inocula in terms of systemic infection of *N. benthamiana* (i.e. F1-G2 and G1-G2) both contained GHu RNA2. In contrast, inocula containing F13 RNA2 were poor performers, as shown previously (Gottula, pg. 120-166, 2014). This could reflect an

actual biological difference in the replication efficiencies of the RNA2s or perhaps reveal a form of optimal compatibility for long distance spread between RNA1 and RNA2 of these two strains in *N. benthamiana.* More work is needed to address this issue. The difference in behavior of RNA2 could also reflect unintentional alterations made to the functionality of F13 RNA2 as a result of cloning strategies. As previously documented by Vigne et al. (2013), alignment of the cDNA of pCLEAN-F13-2 to the GenBank reference sequence of F13 RNA2 (NC_003623) revealed several differences, including insertion of sequences into the 5' and 3' UTR, as well as the insertion of an amino acid into the $2A^{HP}$ coding region. The biological significance of these sequence differences is unknown. In any event, based on the intermediate performance of G1-G2 in terms of systemic infection, this GFLV inoculum was further used to optimize agroinoculation conditions.

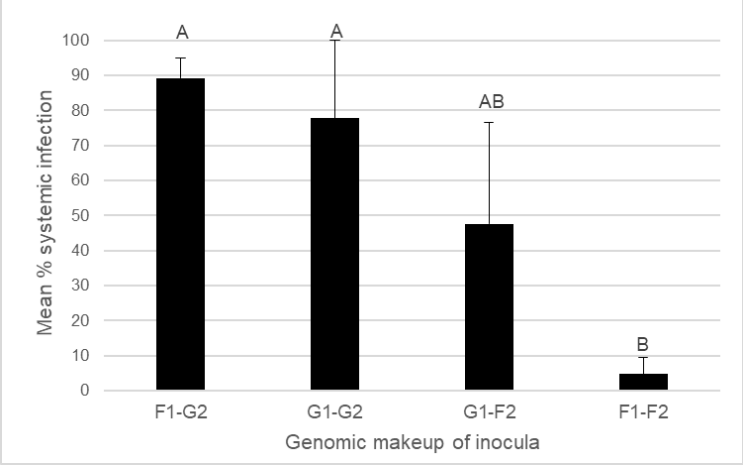


Figure 3-1. Effect of genomic makeup of inocula on incidence of GFLV infection following agroinoculation. F1: *A. tumefaciens* culture harboring cDNA construct of RNA1 of GFLV strain F13. F2: GFLV-F13 RNA2. G1: RNA1 of GFLV strain GHu. G2: GFLV-GHu RNA2. Each inoculum also contained a construct of grapevine leafroll associated virus 2 (GLRaV-2) p24, a silencing suppressor. *A. tumefaciens* cultures were diluted to an OD_{600} of 1.0 and amended with 200 µM acetosyringone prior to infiltration. Each bar represents the mean percent systemic infection of three independent replicates of at least five *N. benthamiana* seedlings (for F1-G2 and G1-G2, n = 6, 5, and 8; for G1-F2 and F1-F2, n = 6, 5, and 14) maintained at 25°C. Bars denoted by the same letter are not significantly different at the 5% level of significance.

A silencing suppressor from grapevine leafroll associated virus 2 increases rate of GFLV infection via agroinoculation

Co-infiltration of VSR constructs during agroinfiltration is known to boost the infectivity of plant viruses in agroinoculation systems (Annamalai and Rao 2005; Chiba et al. 2006). To date, silencing suppressor activity has not been detected for any GFLV protein. To test the effect of VSRs including GLRaV-2 p24, potato virus Y HC-Pro, tomato bushy stunt virus P19 (Canto et al. 2006; Gottula, pg. 120-165, 2014), and cucumber mosaic virus 2b (Choi et al. 2008; Gottula, pg. 120-165, 2014), on the frequency of systemic GFLV infection, cultures of single VSRs were combined with cultures of pCLEAN-GHu-1 and pCLEAN-GHu-2. Among the four VSRs tested in triplicated experiments, only p24 of GLRaV-2 significantly boosted the infection frequency of G1-G2 above the level of infection obtained with an empty vector used as a negative control (Figure 3-2). It is curious that only co-infiltration with GLRaV-2 p24 had a dramatic effect on GFLV systemic infection in *N. benthamiana* following agroinoculation given that the silencing suppression strategy of GLRaV-2 p24 does not appear to differ drastically from the canonical VSR activities of P19, 2b, and HC-Pro with sequestration of siRNAs, prevention of their loading into the RNA-induced silencing complex and interference with the regulation of agronaute 1 (AGO1) (Burgyán and Havelda 2011), though p24 does not physically interact with AGO1 (Li et al. 2017). Additionally, P19 reduces cellular levels of AGO1 (Várallyay et al. 2010) and 2b directly interacts with AGO1 to inhibit its slicing function (Zhang et al. 2006). Notably, p24 is derived from a grapevine virus. However, it is unknown as to why a VSR from GLRaV-2, a phloem-restricted virus, would interfere with RNA silencing

and boost the infection rate of GFLV -which has no known tissue tropism- following agroinoculation.

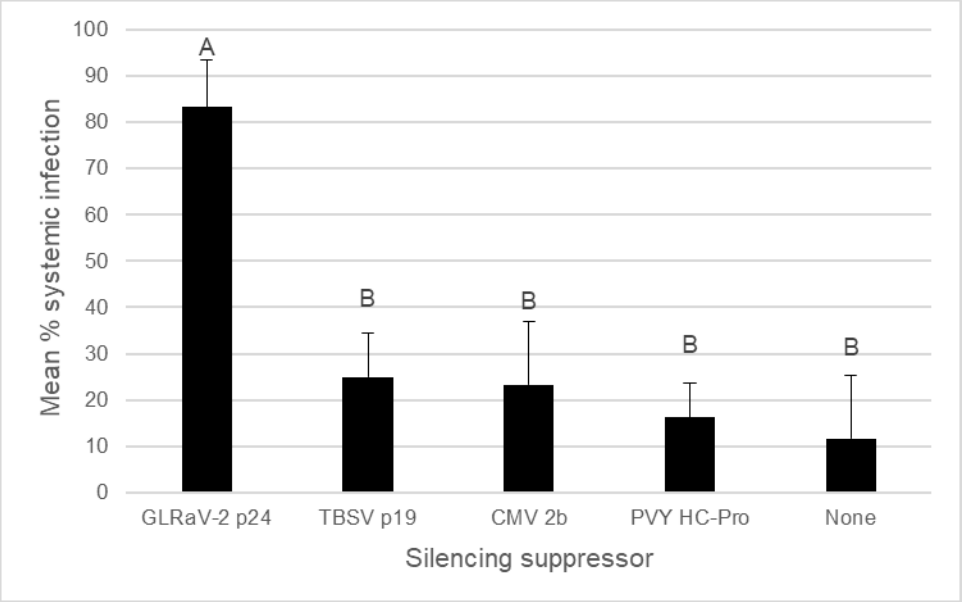


Figure 3-2. Effect of co-infiltration of viral silencing suppressors (VSRs) on incidence of GFLV infection following agroinoculation. *A. tumefaciens* cultures of RNA1 and RNA2 of GFLV strain GHu were combined with cultures of constructs of various VSRs: p24 of GLRaV-2 (grapevine leafroll associated virus-2); P19 of TBSV (tomato bushy stunt virus); 2b of CMV (cucumber mosaic virus); HC-Pro of PVY (potato virus Y); None: culture of the empty binary plasmid pGA482G. *A. tumefaciens* cultures were diluted to an OD_{600nm} of 1.0 and amended with 200 µM acetosyringone prior to infiltration. Each bar represents the mean percent systemic infection of three independent replicates of at least 15 *N. benthamiana* seedlings (for all VSR treatments, n = 15 for all replicates; for the "none" control, n = 60 for all replicates), maintained at 25°C. Bars denoted by the same letter are not significantly different at the 5% level of significance.

GFLV agroinoculation is more efficient at 25°C than 28°C

Standard protocols recommend agroinoculation experiments to be carried out at 25°C (Annamalai and Rao 2005; Vaghchhipawala et al. 2011). To test the effect of temperature on the frequency of systemic GFLV infection, *N. benthamiana* seedlings were maintained in a growth chamber at 25°C or 28°C. In duplicated experiments, the infection frequency was significantly higher ($p<0.005$) at 25 °C than at 28°C (mean percent systemic infection 48.9% and 1.8% for 25°C and 28°C, respectively), with

both treatments presenting little variability (standard error 2.5 and 1.79 for 25°C and 28°C, respectively). It was also noted that vein-clearing symptoms associated with GFLV-GHu (Schmitt-Keichinger et al. 2017; Vigne et al. 2013) were generally more intense in plants maintained at 25°C, with peak symptoms showing at 10 dpi (data not shown). This is consistent with an inverse correlation between virus symptom severity and temperature (Patil and Fauquet 2015).

A. tumefaciens *culture density does not affect efficiency of GFLV agroinoculation*

Some agroinoculation protocols recommend a low culture density, with optical densities (OD_{600nm}) around 0.5, to avoid the tissue yellowing and necrosis typically induced by high culture densities of *A. tumefaciens* (Vaghchhipawala et al. 2011). Other protocols recommend an optical density of 1.0, a number that appears to be chosen for its convenience rather than any empirically supported advantage it confers (Annamalai and Rao 2005). To test the effect of culture densities on the frequency of systemic GFLV infection, suspensions of pCLEAN-GHu-1, pCLEAN-GHu-2 and pGA-p24 were prepared as described above, except each culture was adjusted to OD_{600nm} of 0.1, 0.5, 1.0, and 2.0. Cultures of the same adjusted OD_{600nm} were then combined at a 1:1:1 ratio to form GFLV inocula. In replicated experiments, the final optical density of infiltration culture had no significant effect on GFLV infection frequency in *N. benthamiana* (Figure 3-3). However, it should be noted that use of cultures of a higher OD_{600nm} (1.0 and 2.0) produced more consistent infection frequencies (standard error 1.85 and 3.70, respectively) than those of lower OD_{600nm} (0.1 and 0.5; standard errors 18.19 and 23.81, respectively). This could be because cultures of a higher density are less sensitive to a drop in the effective colony forming

units of the infiltration culture and thus more tolerant of incidental variability between replicates.

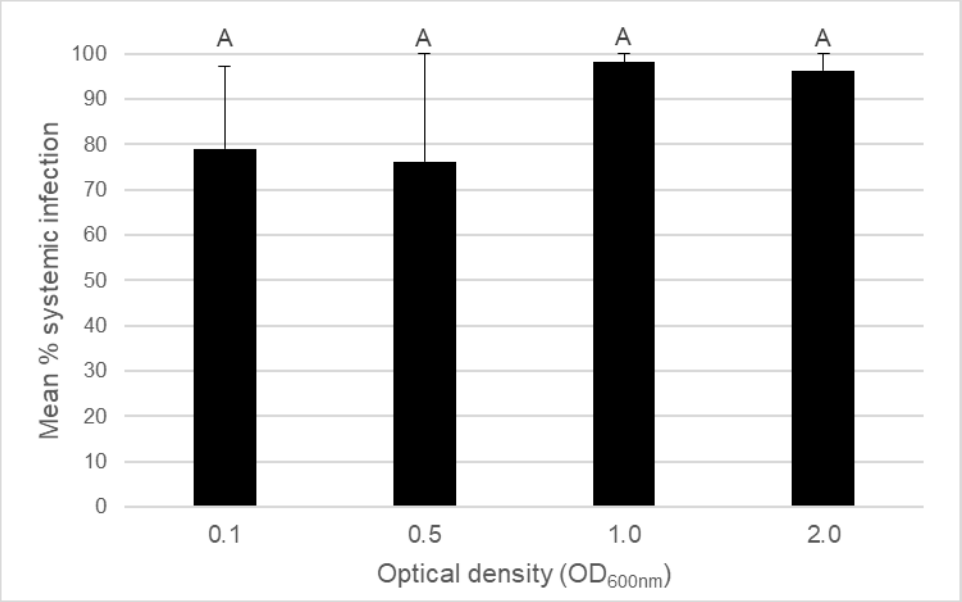


Figure 3-3. Effect of optical density (OD_{600nm}) of infiltration cultures on incidence of GFLV infection following agroinoculation. Cultures of RNA1 and RNA2 of GFLV strain GHu and of GLRaV-2 p24 were individually adjusted to the indicated OD_{600}, combined, and amended with 200 µM acetosyringone prior to infiltration. Each bar represents the mean percent systemic infection of three independent replicates of at least 11 *N. benthamiana* seedlings (for 0.1 and 0.5 treatments, n = 14, 18, and 15; for 1.0 treatment, n = 11, 18, and 15; for 2.0 treatment, n = 12, 18, 15), maintained at 25°C. Bars denoted by the same letter are not significantly different at the 5% level of significance.

Inclusion of a T-DNA transfer inducing-compound does not affect efficiency of GFLV agroinoculation

The inclusion of acetosyringone in *A. tumefaciens* culture media is often recommended to activate T-DNA transfer and thus increase infection frequency (Godwin et al. 1991). Acetosyringone is a phenolic compound released by wounded plants which *A. tumefaciens* uses as a signal molecule to induce expression of Ti-plasmid encoded virulence genes (Stachel et al. 1985). Some work has determined acetosyringone to be essential in some systems (Sheikholeslam and Weeks 1987), while providing little or no benefit in others (Godwin et al. 1991). To test the effect of

acetosyringone on the frequency of systemic GFLV infection following agroinoculation, four separate formulations of induction medium (LB broth amended with 10 mM *N*-morpholinoethanesulfonic acid and 0 µM, 10 µM, 100 µM, or 200 µM acetosyringone) and infiltration medium (sterile water amended with 10 mM each MES and $MgCl_2$ and 0 µM, 20 µM, 200 µM, or 400 µM acetosyringone) were used to prepare the GFLV inoculum G1-G2. In duplicated experiments, no significant effect of the inclusion of acetosyringone in bacterial culture media on GFLV infection frequency in *N. benthamiana* seedlings following agroinoculation was found (Figure 3-4). Godwin et al. (1991) found that *N. tabacum* cells naturally produced sufficient acetosyringone in response to wounding to trigger effective T-DNA transfer. In the context of GFLV agroinoculation, this phenomenon may extend to *N. benthamiana* as the inclusion of acetosyringone in culture media had no significant effect on GFLV infection frequency.

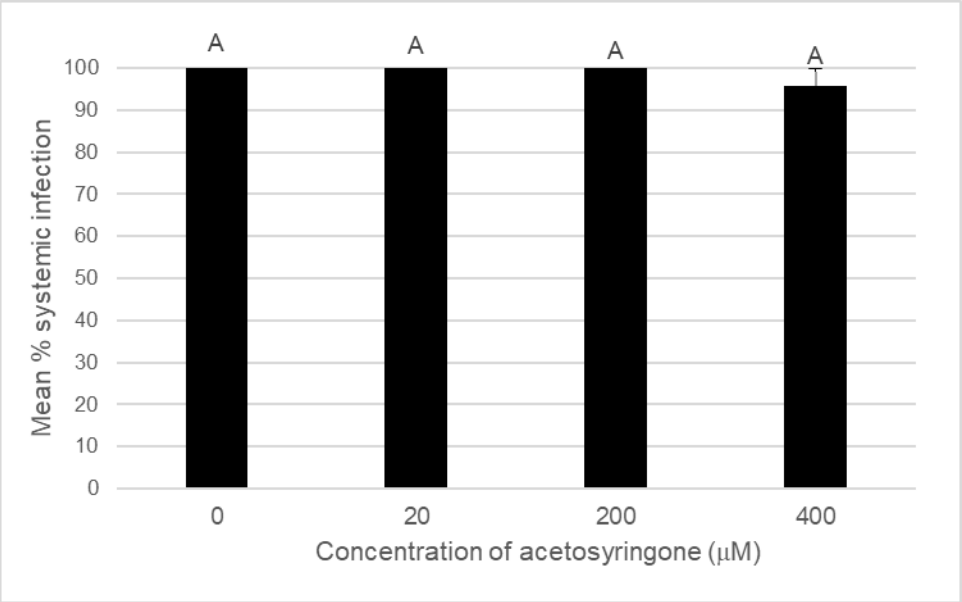


Figure 3-4. Effect of inclusion of acetosyringone in *A. tumefaciens* culture media on incidence of GFLV infection in agroinoculation. Cultures of RNA1 and RNA2 of GFLV strain GHu and of GLRaV-2 p24 were prepared in infiltration media amended with acetosyringone at the indicated concentrations, diluted to an OD_{600} of 1.0, and combined. Note that during culture creation, the bacterial culture media (induction media) was amended with acetosyringone at 0 µM, 10 µM, 100 µM, and 200 µM, respectively. Each bar represents the mean percent systemic infection of two independent replicates of at least 12 *N. benthamiana* seedlings (n = 12 for all treatments for all replicates), maintained at 25 °C. Bars denoted by the same letter are not significantly different at the 5% level of significance.

GFLV mutants can be rapidly produced and characterized using this system

To demonstrate the versatility of our GFLV agroinoculation system for reverse genetic studies, several GFLV chimera constructs were generated to replace segments of the 3' end of the $1E^{Pol}$ coding region of pCLEAN-GHu-1 with the homologous sequences from strain F13. The chimeras were generated by one-step site-directed mutagenesis of plasmid pCLEAN-GHu-1 using the Q5 Site-directed mutagenesis kit (New England Biolabs) as per manufacturer's instructions with specific back-to-back primers (Table 3-2) designed using the NEBaseChanger software (v1.2.6). Three constructs were engineered, each with a chimeric region ranging from 66 to 90 base pairs. The resulting chimeric plasmids, pCLEAN-GHu-1-1EA, pCLEAN-GHu-1-1EB

and pCLEAN-GHu-1-1EC, were screened for the proper mutations by Sanger sequencing (Cornell University Biotechnology Resource Center) with primers specific to the region of interest (GHusymdeterSeq-F: 5'- TGC TAG GAC CAA ATC AGA GGA ACG-3'; GHusymdeterSeq-R: 5'- AAA CTT GGT TAT CCC AGT ACC A-3'). One correctly mutated plasmid clone of each construct was then fully sequenced at the Center of Computational and Integrative Biology DNA Core Facility at Massachusetts General Hospital (Cambridge, MA). Fully sequence-verified plasmids were then analyzed for *in planta* infectivity by agroinoculation using the protocol described above and using pCLEAN-GHu-2 as the RNA2 source. The three chimeric clones established systemic infection in *N. benthamiana* via agroinoculation (Figure 3-5) and produced symptoms consistent with those of GFLV-GHu, i.e. vein clearing symptoms on apical leaves starting 7-8 days post-agroinoculation and fading about 12-15 days post-agroinoculation (Vigne et al. 2013). Characterizing the progeny viruses from the three constructs by sequencing (Cornell University Biotechnology Resource Center) RT-PCR products (QIAGEN OneStep) obtained with primers GHusymdeterSeq-F and GHusymdeterSeq-R and total RNA (E.Z.N.A. Plant RNA kit) extracted from apical leaves of *N. benthamiana* confirmed their fidelity with no indel or substitution observed in any of the sequences analyzed. Furthermore, these chimeras stably retained the intended mutations following up to three serial mechanical passages in *N. benthamiana* (data not shown).

Table 3-2. Primers for mutagenesis of the cDNA of GFLV RNA1 in pCLEAN-GHu-1

Construct	Primers	Primer sequence 5'-3'[a]	Chimeric region[b]
pCLEAN-GHu-1-1E-A	GHu1-1E-A-F	tcttatgaggggggtgaagctttaaaggaaattTTCTCTTTTTGTGAGACTGC	2065..2130
	GHu1-1E-A-R	gtttgttttggcatggactttcttgatatgaatGGGACTTTCAAAATCAGAAAAATAG	
pCLEAN-GHu-1-1E-B	GHu1-1E-B-F	actaggaaagtgatgttgcaaaatcaaccagacgtctatcctagtATCAGTCTTGTTGGGGGAATTTG	2131..2220
	GHu1-1E-B-R	atccatttcggtggcttcatagaaaatagtctcacaaaaagtgaaAACTTCCAGTAAAGCCTCAC	
pCLEAN-GHu-1-1E-C	GHu1-1E-C-F	cctggagccatgtactcggaaacagatgttacgatggctagagagACTCCAGGGGTTTACGTG	2221..2310
	GHu1-1E-C-R	ctctcctccctcatttgggaaacaaaccccccccaacaagacttatACTAGGATATATATCAGACTGGTTTTG	

[a] Nucleotides in lowercase are chimeric (mutagenic) sequences of GFLV-F13 origin; nucleotides in uppercase bind to the template pCLEAN-GHu-1

[b] Nucleotide positions are given for the $1E^{Pol}$ coding region of GFLV strain GHu (JN391442)

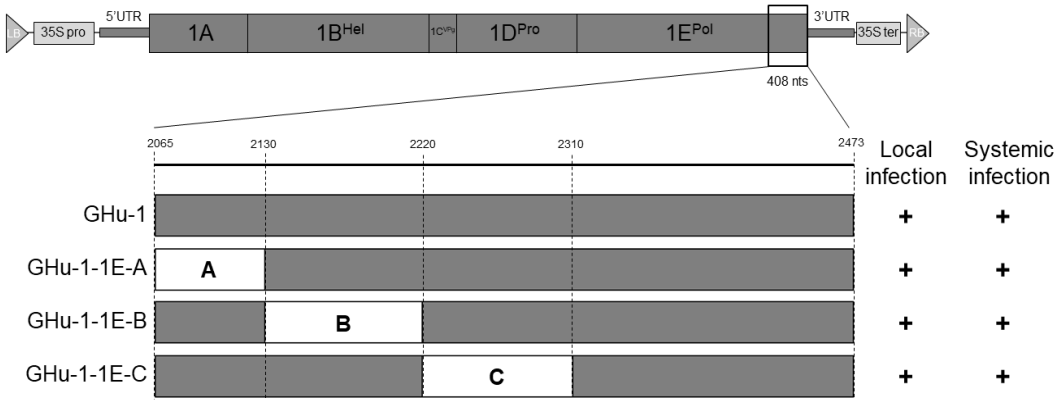


Figure 3-5. Infectivity of grapevine fanleaf virus (GFLV) $1E^{Pol}$ chimeras. *Top* Schematic of the T-DNA region of pCLEAN-GHu-1 (GFLV-GHu RNA1 cDNA in a CaMV 35S expression cassette within the T-DNA region of pCLEAN-G181). LB: left border; 35S pro: promoter of CaMV 35S RNA; UTR: untranslated region; 35S ter: terminator of CaMV 35S RNA; RB: right border. *Bottom* Close up of chimeric regions of $1E^{Pol}$ chimeras. The first column shows the abbreviated names of the constructs. The second column shows graphical representations of the chimera constructs. Each full bar represents the ultimate 408 nt of the $1E^{Pol}$ coding region of GFLV. Sequence segments derived from asymptomatic strain F13 are shown in white; sequences derived from the symptomatic strain GHu are shown in grey. Regions A, B, C represent positions nts 2065-2130, 2131-2220, and 2221-2310 of the $1E^{Pol}$ coding region that were swapped between GFLV strains GHu and F13. The third column indicates if the constructs are locally infectious, as determined by DAS-ELISA of agroinoculated *N. benthamiana* tissue 3 days post inoculation with GFLV-specific antibodies. The fourth column indicates if the constructs are systemically infectious, as determined by GFLV DAS-ELISA of apical, uninoculated *N. benthamiana* tissue at 3 weeks post inoculation.

Functional domains of several GFLV proteins have been characterized by reverse genetics approaches (Belin et al. 1999; Gaire et al. 1999; Martin et al. 2018; Schellenberger et al. 2010; Vigne et al. 2013; Schmitt-Keichinger et al. 2017), which require precise manipulation of viral genomic sequences. Site-directed mutagenesis allows for such manipulations to be made accurately and quickly, without reliance on naturally available or artificially created compatible restriction enzyme sites (Hemsley et al. 1989). The advantage of our pCLEAN-GFLV plasmids over the previous generation (pGA482G-GFLV) is their reduced size (7-11 kb versus 19-22.5 kb), which allows them to be efficiently altered by one-step site-directed mutagenesis. In this

study, regions of up to 90 base pairs of pCLEAN-GHu-1 were mutated. Each mutant construct required only one pair of back-to-back primers and a single PCR run each, representing a notable decrease in time-at-bench compared to the primer extension PCR methods used to create previous GFLV mutants (Belin et al. 1999; Gaire et al. 1999; Martin et al. 2018; Schellenberger et al. 2010), which requires several PCR runs, followed by restriction enzyme digest and ligation. This highlights the versatility of the second-generation GFLV agroinoculation system for rapid and reliable manipulation of the GFLV genome for functional analyses.

CONCLUSIONS

In conclusion, the pCLEAN based GFLV cDNA constructs allow for rapid and accurate manipulation of GFLV genomic sequences. High levels of systemic GFLV infection via agroinoculation (80-100%) were obtained with the F1-G2 inoculum, co-infiltrated with a construct containing the GLRaV-2 p24 silencing suppressor on *N. benthamiana* maintained at 25°C. Agroinfiltration cultures adjusted to an optical density (OD_{600nm}) of 1.0 to 2.0 provided greater consistency among independent inoculations. Adherence to these conditions will provide a method for scalable, cost-effective, and reproducible establishment of GFLV systemic infection in *N. benthamiana* via agroinoculation and enable easy alteration of the GFLV genome for functional analyses.

REFERENCES

An, G. 1986. Development of plant promoter expression vectors and their use for analysis of differential activity of nopaline synthase promoter in transformed tobacco cells. Plant Physiol. 81:86–91

Andret-Link, P., Laporte, C., Valat, L., Ritzenthaler, C., Demangeat, G., Vigne, E., Laval, V., Pfeiffer, P., Stussi-Garaud, C., and Fuchs, M. 2004. Grapevine fanleaf virus: Still a major threat to the grapevine industry. J. Plant Pathol. 86:183–195

Annamalai, P., and Rao, A. L. N. 2005. Delivery and expression of functional viral RNA genomes *in planta* by agroinfiltration. in: Current Protocols in Microbiology, John Wiley & Sons, Inc.

Belin, C., Schmitt, C., Gaire, F., Walter, B., Demangeat, G., and Pinck, L. 1999. The nine C-terminal residues of the grapevine fanleaf nepovirus movement protein are critical for systemic virus spread. J. Gen. Virol. 80:1347–1356

Burgyán, J., and Havelda, Z. 2011. Viral suppressors of RNA silencing. Trends Plant Sci. 16:265–272

Canto, T., Uhrig, J. F., Swanson, M., Wright, K. M., and MacFarlane, S. A. 2006. Translocation of Tomato bushy stunt virus P19 protein into the nucleus by ALY proteins compromises its silencing suppressor activity. J. Virol. 80:9064–9072

Chiba, M., Reed, J. C., Prokhnevsky, A. I., Chapman, E. J., Mawassi, M., Koonin, E. V, Carrington, J. C., and Dolja, V. V. 2006. Diverse suppressors of RNA silencing enhance agroinfection by a viral replicon. Virology 346:7–14

Choi, M. S., Yoon, I.-S., Rhee, Y., Choi, S. K., Lim, S.-H., Won, S.-Y., Lee, Y.-H., Choi, H.-S., Lee, S.-C., Kim, K.-H., Lomonossoff, G. P., and Sohn, S.-H. 2008. The effect of *Cucumber mosaic virus* 2b protein to transient expression and transgene silencing mediated by agro-infiltration. Plant Pathol. J. 24:296–304

Digiaro, M., Elbeaino, T., and Martelli, G. P. 2017. *Grapevine fanleaf virus* and other Old World nepoviruses. Pages 47–82 in: Grapevine Viruses: Molecular Biology, Diagnostics and Management, B. Meng, G.P. Martelli, D.A. Golino, and M. Fuchs, eds. Springer International Publishing, Cham.

Fuchs, M., Schmitt-Keichinger, C., and Sanfaçon, H. 2017. A renaissance in nepovirus research provides new insights into their molecular interface with hosts and vectors. Pages 61–105 in: Advances in Virus Research, M. Kielian, T.C. Mettenleiter, and M.J. Roossinck, eds. Academic Press, Cambridge, MA, USA.

Gaire, F., Schmitt, C., Stussi-Garaud, C., Pinck, L., and Ritzenthaler, C. 1999. Protein 2A of grapevine fanleaf nepovirus is implicated in RNA2 replication and colocalizes to the replication site. Virology 264:25–36

Godwin, I., Todd, G., Ford-Lloyd, B., and Newbury, H. J. 1991. The effects of acetosyringone and pH on *Agrobacterium*-mediated transformation vary according to plant species. Plant Cell Rep. 9:671–675

Gottula, J. W. 2014. Grapevine fanleaf virus: biology, biotechnology and resistance. Ph.D. Dissertation. Cornell University

Hellens, R., Mullineaux, P., and Klee, H. 2000. Technical focus: A guide to *Agrobacterium* binary Ti vectors. Trends Plant Sci. 5:446–451

Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G., and Galas, D. J. 1989. A simple method for site-directed mutagenesis using the polymerase chain reaction. Nucleic Acids Res. 17:6545–6551

Huss, B., Walter, B., and Fuchs, M. 1989. Cross-protection between arabis mosaic virus and grapevine fan leaf virus isolates in *Chenopodium quinoa*. Ann. Appl. Biol. 114:45–60

Li, M., Zhang, J., Feng, M., Wang, X., Luo, C., Wang, Q., and Cheng, Y. 2017. Characterization of silencing suppressor p24 of *Grapevine leafroll-associated virus 2*. Mol. Plant Pathol. 19:355-368

Margis, R., Viry, M., Pinck, M., and Pinck, L. 1991. Cloning and *in vitro* characterization of the grapevine fanleaf virus proteinase cistron. Virology 185:779–787

Martin, I. R., Vigne, E., Berthold, F., Komar, V., Lemaire, O., Fuchs, M., and Schmitt-Keichinger, C. 2018. The 50 distal amino acids of the $2A^{HP}$ homing protein of *Grapevine fanleaf virus* elicit a hypersensitive reaction on *Nicotiana occidentalis*. Mol. Plant Pathol. 19:731–743

Naraghi-Arani, P., Daubert, S., and Rowhani, A. 2001. Quasispecies nature of the genome of *Grapevine fanleaf virus*. J. Gen. Virol. 82:1791–1795

Patil, B. L., and Fauquet, C. M. 2015. Light intensity and temperature affect systemic spread of silencing signal in transient agroinfiltration studies. Mol. Plant Pathol. 16:484–494

Pinck, M., Reinbolt, J., Loudes, A. M., Le Ret, M., and Pinck, L. 1991. Primary structure and location of the genome-linked protein (VPg) of grapevine fanleaf nepovirus. FEBS Lett. 284:117–119

Quacquarelli, A., Gallitelli, D., Savino, V., and Martelli, G. P. 1976. Properties of Grapevine fanleaf virus. J. Gen. Virol. 32:349–360

Ritzenthaler, C., Pinck, M., and Pinck, L. 1995. Grapevine fanleaf nepovirus P38 putative movement protein is not transiently expressed and is a stable final maturation product *in vivo*. J. Gen. Virol. 76:907–915

Ritzenthaler, C., Viry, M., Pinck, M., Margis, R., Fuchs, M., and Pinck, L. 1991. Complete nucleotide sequence and genetic organization of grapevine fanleaf nepovirus RNA1. J. Gen. Virol. 72:2357–2365

Schellenberger, P., Andret-Link, P., Schmitt-Keichinger, C., Bergdoll, M., Marmonier, A., Vigne, E., Lemaire, O., Fuchs, M., Demangeat, G., and Ritzenthaler, C. 2010. A stretch of 11 amino acids in the βB- βC loop of the coat protein of *Grapevine fanleaf virus* is essential for transmission by the nematode *Xiphinema index*. J. Virol. 84:7924–7933

Schellenberger, P., Sauter, C., Lorber, B., Bron, P., Trapani, S., Bergdoll, M., Marmonier, A., Schmitt-Keichinger, C., Lemaire, O., Demangeat, G., and Ritzenthaler, C. 2011. Structural insights into viral determinants of nematode mediated Grapevine fanleaf virus transmission. PLoS Pathog. 7:e1002034

Schmitt-Keichinger, C., Hemmer, C., Berthold, F., and Ritzenthaler, C. 2017. Molecular, Cellular, and Structural Biology of *Grapevine fanleaf virus*. Pages 83–107 in: Grapevine Viruses: Molecular Biology, Diagnostics and Management, B. Meng, G.P. Martelli, D.A. Golino, and M. Fuchs, eds. Springer International Publishing, Cham.

Serghini, M. A., Fuchs, M., Pinck, M., Reinbolt, J., Walter, B., and Pinck, L. 1990. RNA2 of grapevine fanleaf virus: sequence analysis and coat protein cistron location. J. Gen. Virol. 71:1433–1441

Sheikholeslam, S. N., and Weeks, D. P. 1987. Acetosyringone promotes high efficiency transformation of *Arabidopsis thaliana* explants by *Agrobacterium tumefaciens*. Plant Mol. Biol. 8:291–298

Stachel, S. E., Messens, E., Van Montagu, M., and Zambryski, P. 1985. Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. Nature 318:624–629

Thole, V., Worland, B., Snape, J. W., and Vain, P. 2007. The pCLEAN dual binary vector system for *Agrobacterium*-mediated plant transformation. Plant Physiol. 145:1211–1219

Vaghchhipawala, Z., Rojas, C. M., Senthil-Kumar, M., and Mysore, K. S. 2011. Agroinoculation and agroinfiltration: simple tools for complex gene function analyses. Pages 65–76 in: Plant Reverse Genetics: Methods and Protocols, A. Pereira, ed. Humana Press, Totowa, NJ.

Valat, L., Burrus, M., Fuchs, M., and Mauro, M. C. 2003. Review of techniques to inoculate grapevines with Grapevine fanleaf virus: lessons and perspectives. Am. J. Enol. Vitic. 54:279–285

Várallyay, É., Válóczi, A., Ágyi, Á., Burgyán, J., and Havelda, Z. 2010. Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. EMBO J. 29:3507–3519

Vigne, E., Bergdoll, M., Guyader, S., and Fuchs, M. 2004. Population structure and genetic variability within isolates of *Grapevine fanleaf virus* from a naturally infected vineyard in France: evidence for mixed infection and recombination. J. Gen. Virol. 85:2435–2445

Vigne, E., Gottula, J. W., Schmitt-Keichinger, C., Komar, V., Ackerer, L., Belval, L., Rakotomalala, L., Lemaire, O., Ritzenthaler, C., and Fuchs, M. 2013. A strain-specific segment of the RNA-dependent RNA polymerase of grapevine fanleaf virus determines symptoms in *Nicotiana* species. J. Gen. Virol. 94:2803–2813

Viry, M., Serghini, M. A., Hans, F., Ritzenthaler, C., Pinck, M., and Pinck, L. 1993. Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. J. Gen. Virol. 74:169–174

Wood, K. R. 1998. Nepovirus isolation and RNA extraction. Pages 197–204 in: Plant Virology Protocols, G.D. Foster and S.C. Taylor, eds. Humana Press, Totowa, NJ.

Zhang, X., Yuan, Y.-R., Pei, Y., Lin, S.-S., Tuschl, T., Patel, D. J., and Chua, N.-H. 2006. *Cucumber mosaic virus*-encoded 2b suppressor inhibits *Arabidopsis* Argonaute1 cleavage activity to counter plant defense. Genes Dev. 20:3255–3268

CHAPTER 4

THE IDENTITY OF A SINGLE RESIDUE OF THE RNA-DEPENDENT RNA POLYMERASE OF GRAPEVINE FANLEAF VIRUS MODULATES VEIN CLEARING IN *NICOTIANA BENTHAMIANA*[3]

ABSTRACT

The mechanisms underlying host plant symptom development upon infection by viruses of the genus *Nepovirus* in the family *Secoviridae*, including grapevine fanleaf virus (GFLV), are poorly understood. In the systemic host *Nicotiana benthamiana,* GFLV strain GHu produces characteristic symptoms of vein clearing in apical leaves unlike other GFLV strains, such as F13, which cause an asymptomatic infection. In this study, we expanded on earlier findings and used reverse genetics to identify residue 802 (lysine, K) of the GFLV-GHu RNA1-encoded RNA-dependent RNA polymerase ($1E^{Pol}$) as a modulator of vein-clearing development in *N. benthamiana.* Mutations to this site abolished (K to G, A, or Q) or attenuated (K to N or P) symptom expression. Noteworthy, residue 802 is necessary but not sufficient for vein clearing as GFLV-F13 RNA1 carrying K^{802} remained asymptomatic in *N. benthamiana*. No correlation was found between symptom expression and RNA1 accumulation, as shown by reverse transcription quantitative polymerase chain reaction. Additionally, the involvement of RNA silencing in vein clearing was ruled out by virus-induced gene silencing experiments and structure predictions for protein $1E^{Pol}$ suggested that

[3]This chapter was published as: Osterbaan, L.J., Choi, J., Kenney, J., Flasco, M., Vigne, E., Schmitt-Keichinger, C., Rebelo, A.R., Heck, M., and Fuchs, M. 2019. The identity of a single residue of the RNA-dependent RNA polymerase of grapevine fanleaf virus modulates vein clearing in *Nicotiana benthamiana.* Mol Plant Microb Interact. 32(7): 790-801 doi.org/10.1094/MPMI-12-18-0337-R.

residue 802 is flanked by strongly predicted stable secondary structures, including a conserved motif of unknown function (^{805}LLKT/AHLK/RT/ALR^{814}). Together, these results reveal the protein nature of the GFLV-GHu symptom determinant in *N. benthamiana* and provide a solid basis for probing and determining the virus-host proteome network for vein clearing.

INTRODUCTION

Grapevine fanleaf virus (GFLV) is one of several viruses from the genus *Nepovirus* in the family *Secoviridae* which is responsible for fanleaf degeneration of grapevines. This disease is one of the most costly viral diseases of grapevine worldwide, causing massive yield losses due to uneven and poor berry set and decreased vine productive lifespan (Digiaro et al. 2017; Schmitt-Keichinger et al. 2017). The symptoms of GFLV in grapevine include vein yellowing, mosaics, internode shortening, and the virus' namesake "fanleaf" symptom characterized by abnormally open petiolar sinuses and diminished lobing (Schmitt-Keichinger et al. 2017). The molecular mechanisms underlying these symptoms remain unknown.

Like all nepoviruses, GFLV has a bipartite positive-sense RNA genome. Each RNA contains a single open reading frame which is expressed by monocistronic translation followed by proteolytic processing by the viral protease, $1D^{Pro}$ (Fuchs et al. 2017; Schmitt-Keichinger et al. 2017). RNA1 contains the viral proteins necessary for genome replication and polyprotein processing. Therefore, RNA1 is capable of replication independent of RNA2; however, plant systemic infection can only occur if both genomic RNAs are present (Viry et al. 1993). The other proteins of RNA1 include 1A, of unknown function; $1B^{Hel}$, a putative helicase (HEL); $1C^{VPg}$, a genome-linked viral protein; and $1E^{Pol}$, a canonical RNA-dependent RNA polymerase (RdRp) (Schmitt-Keichinger et al. 2017). The proteins encoded by RNA2 include $2A^{HP}$, a putative homing protein; $2B^{MP}$, a tubule-type movement protein; and $2C^{CP}$, the coat protein, which makes up the pseudo T=3 icosahedron of the GFLV capsid and is responsible for virus-vector specificity (Schmitt-Keichinger et al. 2017). To date, no

GFLV protein has shown definitive silencing suppressor activity, although the coat protein of tomato ringspot virus (ToRSV), another nepovirus, was documented as a weak silencing suppressor (Karran and Sanfaçon 2014).

How plant viruses produce symptoms in their hosts is a long-standing question in plant pathology. An early hypothesis was the competitive disease model, which proposed that plant viruses replicate so prodigiously in their hosts as to outcompete host components for access to cellular resources, essentially starving the host at the cellular level (Culver and Padmanabhan 2007). While some plant viruses are able to accumulate to remarkably high levels in their hosts, the fact that the presence and severity of symptoms often cannot be strictly correlated to virus titer (Duff-Farrier et al. 2015; Fujita et al. 2018; Mansilla et al. 2009; Rodríguez-Cerezo et al. 1991) suggests that the competitive disease model is not sufficient to explain the mechanism by which virus infection leads to symptoms.

A more recent and compelling explanation for viral symptom development *in planta* is the interaction disease model, which proposes that alterations of host physiology are the result of specific interactions between viral and host components (Culver and Padmanabhan 2007). Indeed, several cases of interactions between viral proteins with specific host components have been reported in the literature and such interactions have been conclusively linked to symptom development (García and Pallás 2015).

The identity and function of viral symptom determinants (also commonly called “pathogenicity determinants”) are quite varied and include RdRps (Mansilla et al. 2009; Liu et al. 2017; Padmanabhan et al. 2005; Vigne et al. 2013), movement

proteins (Rao and Grantham 1995; Hasiów-Jaroszewska et al. 2011), helper component-proteases (Shiboleth et al. 2007), capsid proteins (Heaton et al. 1991; Hasiów-Jaroszewska et al. 2013; Qiu et al. 2018) and a cysteine-rich protein (Fujita et al. 2018). In addition to their basic functions, many viral protein symptom determinants possess RNA silencing suppression activity (Shiboleth et al. 2007; Fujita et al. 2018). Some studies have even mapped symptom determinants to individual amino acids (Heaton et al. 1991; Kagiwada et al. 2005; Ozeki et al. 2006), and, in one instance, the symptom determinant (modulator) and symptom elicitor are separate viral components (Komatsu et al. 2011). However, there have also been instances where symptom determinants are based on nucleotide sequences, such as satellite RNAs (Smith et al. 2011; Shimura et al. 2011) or non-coding regions of the genome, such as 5' and 3' untranslated regions (Rodríguez-Cerezo et al. 1991).

GFLV isolates may be mechanically transmitted from grapevine tissue to herbaceous plants, including *Nicotiana benthamiana*, a model host plant that sustains systemic virus infection. Most strains of GFLV, including strain F13, are asymptomatic in *N. benthamiana*; however, strain GFLV-GHu elicits distinctive symptoms of vein clearing on the upper, uninoculated leaves of *N. benthamiana* (Vigne et al. 2013). These observations suggested that GFLV-GHu encodes a strain-specific symptom determinant. A previous investigation mapped the GHu symptom determinant to the 408 nts (136 aa) at the 3' end of the coding region of $1E^{Pol}$. However, it was not determined if the underlying mechanism is dependent on the nucleotide or amino acid sequence (Vigne et al. 2013). We recently developed a series of binary plasmids carrying the full-length cDNAs of RNA1 and RNA2 of GFLV

strains F13 and GHu for the rapid generation of mutant GFLV RNA1 constructs (Osterbaan et al. 2018). We used this system to characterize the nature of the GFLV-GHu symptom determinant and examine whether it acts at the nucleotide or amino acid level. Here we report vein clearing symptom development for GFLV-GHu in *N. benthamiana* is dependent on the identity of residue 802 of protein $1E^{Pol}$.

MATERIALS AND METHODS

In silico *characterization of GFLV* $1E^{Pol}$

Nucleotide sequences of GFLV isolates which included sequence data for the 3' ultimate 408 nucleotides of the $1E^{Pol}$ coding region were retrieved from GenBank as of December 2017, along with the predicted translation sequences of $1E^{Pol}$. Sequences were manually trimmed to include only the 408 nucleotides and 136 amino acids of interest. A recent metagenomic study provided 70 new complete GFLV RNA1 sequences to GenBank (Hily et al. 2018). We selected four representative sequences from this set for inclusion in our *in silico* analyses. For the full list and reference information of GFLV isolates used for *in silico* analyses, see Table 4-1.

Table 4-1. Grapevine fanleaf virus isolates and their RNA1 sequence GenBank accession numbers that were used for disorder predictions

Isolate	GenBank accession number	Reference[a]
F13	NC_003615	Ritzenthaler et al., 1991
GHu	JN391442	Vigne et al., 2013
CACSC4	GU972570	Oliver et al., 2010
CACSC3	GU972569	Oliver et al., 2010
CACSC2	GU972568	Oliver et al., 2010
CACSC1	GU972567	Oliver et al., 2010
CACSB5	GU972566	Oliver et al., 2010
CACSB3	GU972565	Oliver et al., 2010
CACSB2	GU972564	Oliver et al., 2010
CACSB1	GU972563	Oliver et al., 2010
CAZINA5	GU972562	Oliver et al., 2010
CAZINA4	GU972561	Oliver et al., 2010
CAZINA3	GU972560	Oliver et al., 2010
CAZINA2	GU972559	Oliver et al., 2010
CAZINA1	GU972558	Oliver et al., 2010
SACH44	KC900162	Lamprecht et al., 2013
SAPCS3	JF968120	Lamprecht et al., 2013
SDHN	KU522584	Zhou et al. 2017, D.S.
1050-02	JX513889	Rott and Belton, D.S.
3138-01	JX513894	Rott and Belton, D.S.
WAPN173	GQ332372	Mekuria et al., 2009
WAPN6132	GQ332373	Mekuria et al., 2009
SWT2-3	KX034858	Hily et al., 2018
SWT2-1	KX034856	Hily et al., 2018
WTR4-2	KX034848	Hily et al., 2018
WTR6-3	KX034851	Hily et al., 2018

[a] D.S. = direct submission

To identify potential disordered regions, the trimmed polyprotein sequences were run through the Predictor of Natural Disordered Regions (PONDR) VL-XT algorithm (Li et al. 1999; Romero et al. 2001) and the IUPred algorithm (Dosztányi et al. 2005). Since both programs treat termini differently than internal sequences, the sequences for all isolates were extended to include eight amino acids upstream of the 136 C-terminal amino acids (i.e. predictions were run on residues 681 to 824, rather

than 689 to 824) such that residues 689 and beyond were treated as internal rather than terminal residues. Residues 689 to 824 of GFLV-GHu $1E^{Pol}$ were used in the Simple Modular Architecture Research Tool (SMART; smart.embl-heidelberg.de) to search for putative functional protein motifs or domains within this region of the $1E^{Pol}$ protein. For protein modeling, the complete amino acid sequence of the $1E^{Pol}$ coding region for GFLV-GHu, along with truncated sequences of residues 1 to 520 and 520 to 824, were run through the iterative threading assembly refinement algorithm (I-TASSER; Yang et al., 2014; zhanglab.ccmb.med.umich.edu/I-TASSER/) using the default settings.

Creation of chimeric and mutant GFLV RNA1 constructs

Each GFLV RNA1 mutant construct used in *A. tumefaciens*-mediated inoculation (agroinoculation) assays was generated using pCLEAN-GHu-1, pCLEAN-F13-1, or derivatives thereof (Osterbaan et al. 2018) as templates for site-directed mutagenesis using the Q5 Site-Directed Mutagenesis kit (New England Biolabs) as per manufacturer's instructions with specific back-to-back primers designed using the NEBaseChanger software (v1.2.6). Plasmids pCLEAN-GHu-1 and pCLEAN-F13-1 contain the full-length cDNA of GFLV -GHu and -F13 RNA1, respectively. Expression of the GFLV RNA1 cDNAs in pCLEAN is driven by the cauliflower mosaic virus 35S promoter and terminator sequences. Chimeric plasmids from single colonies were screened for proper mutations by Sanger sequencing with primers specific to the region of interest at Cornell University Biotechnology Resource Center. At least one successfully mutated plasmid clone of each construct was selected for *in planta* characterization by agroinoculation. See Table 4-2 for all primers used for

construct and viral progeny sequencing confirmation, including those for TRV and pG_R constructs (see below).

Constructs used in mechanical inoculation with *in vitro* RNA transcripts were generated by site-directed mutagenesis with overlap PCR (Hemsley et al. 1989) using $pF1_S$ (Vigne et al. 2013) as a template and mutagenic primers. Plasmid $pF1_S$ contains the full-length cDNA of GFLV-F13 RNA1 slightly modified by the insertion of a *Sna*BI site (Vigne et al, 2013). Expression of the GFLV-F13 RNA1 is driven by the T7 promoter. For each construct, three PCRs were performed with Phusion High-Fidelity DNA Polymerase (New England Biolabs). In the first PCR, the forward mutagenic primer was paired with primer EVP1R7854; in the second PCR, the reverse mutagenic primer was paired with primer EV1SGFsn4628; in the third PCR, the products of the first two reactions were combined and amplified by the primer pair EV1SFGsn4628/EVP1R7854. The resulting amplicon was then cloned into the $pF1_S$ backbone by *Sna*BI/*Not*I digestion and ligation. See Table 4-3 for all primers used for mutagenesis and cloning, including those used for cloning into TRV and pG_R VIGS constructs.

Table 4-2. Primers for sequencing progeny of grapevine fanleaf virus (GFLV) RNA1 and RNA2 mutant and chimeric constructs; strain-specific detection of GFLV-F13 and GFLV-GHu RNA1 in mixed infections; and sequencing confirmation of TRV RNA2 clones.

Construct type	Forward primer name	Forward primer sequence (5'-3')	Reverse primer name	Reverse primer sequence (5'-3') [a]
pCLEAN-GHu-1 derivatives	GHusymdeterSeq-F	TGCTAGGACCAAATCAGAGGAACG	GHusymdeterSeq-R	AAACTTGGTTATCCCAGTACCA
pG_R derivatives	pROXMCS-F	CAGTCCTGTGCTAGGGCTT	pROXMCS-R	GCTTCCTGGGCAATCGTTT
$pF1_S$ derivatives	G9	CAGAAAGAAGAGCTGGCGGG	G10	GGATCAGGATATGGAAAGCAC
	G11	CCCAAAAGTCATCGCAATGCT	G12	GAGYACTACTGTGAGAGCACA
	G13	GCATCAATCCTTATAGTCGCGA	G14	ACATATGCAAGCAACTGCCCA
	G15	TACTCAACAAGCTGGTACAGA	G16	AAATTTGCAWAACAGTAAAAAGAAA
F13 RNA1 (mixed infection)	GFLV-F13-1-F	TGCTGCAAAGCAACGTCTTC	GFLV-F13-1-R	CACTGGCAAGATGCTGGAGA
GHu RNA1 (mixed infection)	GFLV-GHu-1-F	CGTGACCACTGCAAACACTG	GFLV-GHu-1-R	GGGATGTGGTGGTTCGGATT
TRV RNA2	p207F	TCGCGTTAACGCTAGCATGGATCTC	p207R	GTAACATCAGAGATTTTGAGACAC
	pQ11F	GAGTGGAGGTCCGATACGTC	pQ11R	AGACAATGAGTCGGCCAAAC

[a] IUPAC degenerate base codes: W = A or T; Y = C or T

Table 4-3. Mutagenic and cloning primers used to generate grapevine fanleaf virus (GFLV) RNA1 mutants, chimeras, and pG_R and TRV constructs.

Table 4-3

Construct name	Template	Forward primer name	Forward primer sequence (5'-3')[a]	Reverse primer name	Reverse primer sequence (5'-3')[a]
F13 2131-2220$_{GHu}$	pCLEAN-F13-1	F131-Ghusd-B-F	actaggaaggagatgcttcaaaaccagtctgatatatatcctagtATAAGTCTTGTTGGGGGG	F131-Ghusd-B-R	atctgagctggtggcttcataaaaagcagtctcacaaaaagagaaAATTTCCTTTAAAGCTTCACC
F13 2221-2310$_{GHu}$	pCLEAN-F13-1	F131-Ghusd-C-F	cctggagcgttgttttccgatgaagacgtgcttaaggctgaacaaGTTCAAGGAGTCTATGTAAG	F131-Ghusd-C-R	tttttccccctctttagggaaacaaattcccccaacaagactgatACTAGGATAGACGTCTGG
F13 2311-2400$_{GHu}$	pCLEAN-F13-1	F131-Ghusd-D-F	agatgtttaggtgttgtcgccaaaactgtgagtagtactgcgcacTTTGGCAACAATCTTTTAAAGAC	F131-Ghusd-D-R	gtggcaacatttcaaacacacctcactcacgtaaacccctggagtCTGTCTAGCCATCGTAAC
F13 2401-2472$_{GHu}$	pCLEAN-F13-1	F131-Ghusd-E-F	ctgaggcgtgttcaggctcactcatgcccgaggaagTAAGCCTTCCAATTCTTGG	F131-Ghusd-E-R	agtcttaagatgcgtcttcaaaaggctattcttaaaAAGTTGCGTATCAGTCAC
GHu 2311-2400$_{F13}$	pCLEAN-GHu-1	GHu1-F13sd-D-F	cgttgtgtaggagtagcaaccagggttgtgactgatacgcaacttTTTAAGAATAGCCTTTTGAAGAC	GHu1-F13sd-D-R	cctgcaacatttcacgcacgcttcacttacatagactccttgaacTTGTTCAGCCTTAAGCAC
GHu 2311-2472$_{F13}$	pCLEAN-GHu1-F13sd-D	GHu1-F13sd-E-F	ttgaggaaaattcagaatcatacatgccttaggaaaTAATTCTTCCAACCCTTGG	GHu1-F13sd-ABCDE-R	agccttaagatgagtctttaaaagattgttgccaaaAAGTTGCGTATCAGTCAC
GHu 2401-2472$_{F13}$	pCLEAN-GHu-1	GHu1-F13sd-E-F	ttgaggaaaattcagaatcatacatgccttaggaaaTAATTCTTCCAACCCTTGG	GHu1-F13sd-E-R	agccttaagatgagtctttaaaagattgttgccaaaGTGCGCAGTACTACTCAC
GHu 2065-2130$_{F13}$	pCLEAN-GHu-1	GHu1-F13sd-A-F2	tcttatgaggggggtgaagctttaaaggaaattTTCTCTTTTTGTGAGACTGC	GHu1-F13sd-A-R2	gtttgttttggcatggactttcttgatatgaatGGGACTTTCAAAATCAGAAAAATAG
GHu 2131-2220$_{F13}$	pCLEAN-GHu-1	GHu1-F13sd-B-F	actaggaaagtgatgttgcaaaatcaaccagacgtctatcctagtATCAGTCTTGTTGGGGGAATTTG	GHu1-F13sd-B-R	atccatttcggtggcttcatagaaaatagtctcacaaaaagtgaaAACTTCCAGTAAAGCCTCAC
GHu 2221-2310$_{F13}$	pCLEAN-GHu-1	GHu1-F13sd-C-F	cctggagccatgtactcggaaacagatgttacgatggctagagagACTCCAGGGGTTTACGTG	GHu1-F13sd-C-R	ctctcctccctcatttgggaaacaaaccccccccaacaagacttatATACTAGGATATATATCAGACTGGTTTTG
GHu A818A	pCLEAN-GHu-1	GHu1-1E-A818A-F	GCGTGTTCAGgccCACTCATGCC	GHu1-1E-A818-R	CTCAGAGTCTTAAGATGCGTCTTCAAAAG
GHu A818N	pCLEAN-GHu-1	GHu1-1E-A818N-F	GCGTGTTCAGaatCACTCATGCC	GHu1-1E-A818N-R	CTCAGAGTCTTAAGATGC
GHu H800H	pCLEAN-GHu-1	GHu1-1E-H800H-F	TAGTACTGCGcatTTTAAGAATAGC	GHu1E-H800L-H-R	CTCACAGTTTTGGCGACAAC

Table 4-3, cont.					
Construct name	Template	Forward primer name	Forward primer sequence (5'-3')[a]	Reverse primer name	Reverse primer sequence (5'-3')[a]
GHu H800H+P822P	pCLEAN-GHu-1E-H800H	GHu-1E-P822P-F	TCACTCATGCcctAGGAAGTAATTC	GHu-1E-P822L-P-R	GCCTGAACACGCCTCAGA
GHu H800L	pCLEAN-GHu-1	GHu-1E-H800L-F	TAGTACTGCGcttTTTAAGAATAGCCTTTTG	GHu-1E-H800L-H-R	CTCACAGTTTTGGCGACAAC
GHu H800L+P822L	pCLEAN-GHu-1E-H800L	GHu-1E-P822L-F	TCACTCATGCcttAGGAAGTAATTCTTCCAAC	GHu-1E-P822L-P-R	GCCTGAACACGCCTCAGA
GHu P822L	pCLEAN-GHu-1	GHu-1E-P822L-F	CACTCATGCcttAGGAAGTAATTCTTCCAAC	GHu-1E-P822L-P-R	GCCTGAACACGCCTCAGA
GHu P822P	pCLEAN-GHu-1	GHu-1E-P822P-F	TCACTCATGCcctAGGAAGTAATTC	GHu-1E-P822L-P-R	GCCTGAACACGCCTCAGA
GHu V779A	pCLEAN-GHu-1	GHu-1E-V779A-F2	CGTGAGTGAGgcgTGTTTGAAATG	GHu1-1E-V77A-V-R	TAAACCCCTGGAGTTTGTTC
GHu V779V	pCLEAN-GHu-1	GHu-1E-V779V-F	CGTGAGTGAGgttTGTTTGAAATG	GHu-1E-V779A-V-R	TAAACCCCTGGAGTTTGTTC
GHu V791A	pCLEAN-GHu-1	GHu-1E-V791A-F	TTTAGGTGTTgcaGCCAAAACTGTG	GHu-1E-V791A-V-R	CATCTGTGGCAACATTTC
GHu V791V	pCLEAN-GHu-1	GHu-1E-V791V-F	TTTAGGTGTTgttGCCAAAACTG	GHu-1E-V791A-V-R	CATCTGTGGCAACATTTC
GHu 2311-2341$_{F13}$	pCLEAN-GHu-1	GHu1E-F13-2311-41-F	taagtgaagcgtgcgTGAAATGTTGCCACAGATG	GHu1E-F13-2311-41-R	catagactccttgaacTTGTTCAGCCTTAAGCAC
GHu 2372-2400$_{F13}$	pCLEAN-GHu-1	GHu1E-F13-2372-2400-F	ctgatacgcaacttTTTAAGAATAGCCTTTTGAAG	GHu1E-F13-2372-2400-R	tcacaaccctggttgCAACACCTAAACATCTGTG
GHu 2434-2458$_{F13}$	pCLEAN-GHu-1	GHu1E-F13-2434-58-F	ttcagaatcataCATGCCCGAGGAAGTAATTC	GHu1E-F13-2434-58-R	ttttcctcaaagcCTTAAGATGCGTCTTCAAAAG
GHu 2305-2315$_{F13}$	pCLEAN-GHu-1	GHu1E-F13-2305-15-F	gttcaAGGGGTTTACGTGAGTGA	GHu1E-F13-2305-15-R	ctctctAGCCTTAAGCACGTCTTC

Table 4-3, cont.

Construct name	Template	Forward primer name	Forward primer sequence (5'-3')[a]	Reverse primer name	Reverse primer sequence (5'-3')[a]
GHu 2404-2424$_{F13}$	pCLEAN-GHu-1	GHu1E-F13-2404-24-F	tttaaagactCATCTTAAGACTCTGAGGC	GHu1E-F13-2404-24-R	agattgttgccAAAGTGCGCAGTACTACTC
GHu 2452-2472$_{F13}$	pCLEAN-GHu-1	GHu1E-F13-2454-72-F	ccttaggaaaTAATTCTTCCAACCCTTGG	GHu1E-F13-2452-72-R	catgtatgattCTGAACACGCCTCAGAGT
F13 2404-2424$_{GHu}$	pCLEAN-F13-1	F13-1E-2404-24-F	tttgaagacgCATCTTAAGGCTTTGAGG	F13-1E-2404-24-R	aggctattcttAAAAAGTTGCGTATCAGTC
F13 G802G	pCLEAN-F13-1	F13-1E-G802G-F2	GCAACTTTTTgggAACAATCTTTTAAAG	F13-1E-G802K-R	GTATCAGTCACAACCCTG
F13 G802K	pCLEAN-F13-1	F13-1E-G802K-F	GCAACTTTTTaagAACAATCTTTTAAAGACTC	F13-1E-G802K-R	GTATCAGTCACAACCCTG
GHu K802K	pCLEAN-GHu-1	GHu-1E-K802K-F	TGCGCACTTTaaaAATAGCCTTT	GHu-1E-K802K-R	GTACTACTCACAGTTTTGGC
GHu S804N	pCLEAN-GHu-1	GHu-1E-S804N-F2	CTTTAAGAATaacCTTTTGAAGACGC	GHu-1E-S804N-R	TGCGCAGTACTACTCACA
GHu S804S	pCLEAN-GHu-1	GHu-1E-S804S-F	CTTTAAGAATtctCTTTTGAAGACGCATCTTAAGAC	GHu-1E-S804N-R	TGCGCAGTACTACTCACA
F13 N804N	pCLEAN-F13-1	F13-1E-N804N-F	TTTTGGCAACaacCTTTTAAAGACTC	F13-1E-N804S-R	AGTTGCGTATCAGTCACAAC
F13 N804S	pCLEAN-F13-1	F13-1E-N804S-F2	TTTTGGCAACagtCTTTTAAAGACTC	F13-1E-N804S-R	AGTTGCGTATCAGTCACAAC
F13 G802G+N804N	pCLEAN-F13-1	F13-1E-G802G-N804N-F	caacCTTTTAAAGACTCATCTTAAGG	F13-1E-G802G-N804N-R	ttaccAAAAAGTTGCGTATCAGTC
F13 G802K+N804S	pCLEAN-F13-1	F13-1E-G802K-N804S-F	cagcCTTTTAAAGACTCATCTTAAGG	F13-1E-G802K-N804S-R	ttcttAAAAAGTTGCGTATCAGTC
GHu K802G+S804N	pCLEAN-GHu-1	GHu-1E-K802G-S804N-F3	aacCTTTTGAAGACGCATCTTAAGAC	GHu-1E-K802G-S804N-R2	ATTGCCAAAGTGCGCAGT

Table 4-3, cont.					
Construct name	Template	Forward primer name	Forward primer sequence (5'-3')[a]	Reverse primer name	Reverse primer sequence (5'-3')[a]
GHu K802K+S804S	pCLEAN-GHu-1	GHu-1E-K802K-S804S-F	ttctCTTTTGAAGACGCATCTTAAG	GHu-1E-K802K-S804S-R	tttttAAAGTGCGCAGTACTACTC
F13 G802P	pCLEAN-F13-1	F131E-G802P-F	GCAACTTTTTcccAACAATCTTTTAAAGAC	F13-1E-G802K-R	GTATCAGTCACAACCCTG
GHu K802R	pCLEAN-GHu-1	GHu-1E-K802R-F	TGCGCACTTTcgtAATAGCCTTTT	GHu-1E-K802G-R	GTACTACTCACAGTTTTGG
GHu K802Y	pCLEAN-GHu-1	GHu-1E-K802Y-F	TGCGCACTTTtacAATAGCCTTTTG	GHu-1E-K802G-R	GTACTACTCACAGTTTTGG
GHu ΔS804	pCLEAN-GHu-1	GHu1E-deltaS804S-F	CTTTTGAAGACGCATCTTAAG	GHu1E-deltaS804-R	ATTCTTAAAGTGCGCAGTAC
GHu K802A [GCC]	pCLEAN-GHu-1	GHu1E-K802A-GCC-F	TGCGCACTTTgccAATAGCCTTTTG	GHu-1E-K802G-R	GTACTACTCACAGTTTTGG
GHu K802A [GCG]	pCLEAN-GHu-1	GHu1E-K802A-GCG-F	TGCGCACTTTgcgAATAGCCTTTTG	GHu-1E-K802G-R	GTACTACTCACAGTTTTGG
GHu K802E	pCLEAN-GHu-1	GHu-1E-K802E-F	TGCGCACTTTgagAATAGCCTTT	GHu-1E-K802K-R	GTACTACTCACAGTTTTGGC
GHu K802G	pCLEAN-GHu-1	GHu-1E-G802G-F	TGCGCACTTTggcAATAGCCTTTTG	GHu-1E-G802G-R	GTACTACTCACAGTTTTGG
GHu K802N	pCLEAN-GHu-1	GHu-1E-K802N-F	TGCGCACTTTaacAATAGCCTTTTG	GHu-1E-K802K-R	GTACTACTCACAGTTTTGGC
GHu K802P	pCLEAN-GHu-1	GHu-1E-K802P-F	TGCGCACTTTccgAATAGCCTTTTG	GHu-1E-K802K-R	GTACTACTCACAGTTTTGGC
GHu K802Q	pCLEAN-GHu-1	GHu-1E-K802Q-F	TGCGCACTTTcagAATAGCCTTTTG	GHu-1E-K802K-R	GTACTACTCACAGTTTTGGC
F13 A779V	pF1s	EV1SF6946V	ATGTAAGTGAAGtGTGCGTGAAATGTTGC	EV1RF6970V	CATTTCACGCACaCTTCACTTACATAGAC
F13 A791V	pF1s	EV1SF6984V	GCAGGAGTAGtcACCAGGGTTGTGACTG	EV1RF7007V	CACAACCCTGGTgaCTACTCCTGCACAACG

Table 4-3, cont.

Construct name	Template	Forward primer name	Forward primer sequence (5'-3')[a]	Reverse primer name	Reverse primer sequence (5'-3')[a]
F13 L800H	pF1s	EV1SF7011H	GATACGCAAcacTTTGGCAACAATCTTTTA	EV1RF7035H	GATTGTTGCCAAAgtgTTGCGTATCAGTCACAAC
pF1s derivatives	pF1s amplicons	EV1SGFsn4628	AGTGCCATACGTACCAGAAGATG	EVP1R7854	ACCATGATTACGAATTCGAGCTC
pG_R-2147-2248-fs	pCLEAN-GHu-1	GHusd102fsFor	TGCTGTGGATCCCGTACGCTGCTTTTTATGAAGCCACCA	GHusd102fsRev	GCATAAGTCCATGGACGCGTGGAAACAAATTCCCCCAACAA
pG_R-2042-2146-fs	pCLEAN-GHu-1	GHu sd amp frsht For	TGCTGTCGTACGATTTTTCTGATTTTGAAAGTCCCGTGT	GHusd5-105fsRev	ATAAGTACGCGTTCTCACAAAAAGAGAAAACTTCCAGTA
pG_R-2248-2352-rc	pCLEAN-GHu-1	GHusd105invFor	GCATAAGTCCATGGACGCGTCCTAAAGAGGGGGAAAAACCT	GHusd105invRev	TGCTGTGGATCCCGTACGGCAACATTTCAAACACACCTCA
pG_R-2313-2447-rc	pCLEAN-GHu-1	GHusd3-siRNAsplitFor	ATAAGTCCATGGACGCGTTCCAGGGGTTAACGTGAGTGA	GHusd3-siRNAsplitRev	TGCTGTGGATCCCGTACGACACGCCTCAGAGTCTTAAGA
pG_R-2147-2314-fs	pCLEAN-GHu-1	GHusd102fsFor	TGCTGTGGATCCCGTACGCTGCTTTTTATGAAGCCACCA	GHusd5-siRNAsplitRev	ATAAGTCCATGGACGCGTGAGTTTGTTCAGCCTTAAGCA
pG_R-2147-2326-fs	pCLEAN-GHu-1	GHusd102fsFor	TGCTGTGGATCCCGTACGCTGCTTTTTATGAAGCCACCA	GHu sd amp frsht Rev	ATAAGTACGCGTCGTAAACCCCTGGAGTTTGTTCAG
pG_R-2042-2248-fs	pCLEAN-GHu-1	Ghu sd amp frsht For	TGCTGTCGTACGATTTTTCTGATTTTGAAAGTCCCGTGT	GHusd102fsRev	GCATAAGTCCATGGACGCGTGGAAACAAATTCCCCCAACAA
pG_R-2248-2472-if	pCLEAN-GHu-1	GHusd105invFor	GCATAAGTCCATGGACGCGTCCTAAAGAGGGGGAAAAACCT	Ghu sd amp full Rev	ATAAGTCCATGGACGCGTCTTCCTCGGGCATGAGTGAG
pG_R-2042-2314-fs	pCLEAN-GHu-1	GHu sd amp frsht For	TGCTGTCGTACGATTTTTCTGATTTTGAAAGTCCCGTGT	GHusd5-siRNAsplitRev	ATAAGTCCATGGACGCGTGAGTTTGTTCAGCCTTAAGCA

Table 4-3, cont.					
Construct name	Template	Forward primer name	Forward primer sequence (5'-3')[a]	Reverse primer name	Reverse primer sequence (5'-3')[a]
pG_R-2041-2352-if	pCLEAN-GHu-1	GHu sd amp full For	TGCTGTGGATCCCGTACGTATTTTTC TGATTTTGAAAGTCCCGTG	GHusd105inv Rev	TGCTGTGGATCCCGTACGGCAACATT TCAAACACACCTCA
pG_R-2041-2472-if	pCLEAN-GHu-1	GHu sd amp full For	TGCTGTGGATCCCGTACGTATTTTTC TGATTTTGAAAGTCCCGTG	GHu sd amp full Rev	ATAAGTCCATGGACGCGTCTTCCTCG GGCATGAGTGAG
TRV:$1E_{F13}$	pCLEAN-F13-1	GFLV_VIGS_F13_attB1_F	*GGGGACAAGTTTGTACAAAAAAGCAGG CTTA*AGTGAAGCGTGCGTGAAATGTT GC	GFLV_VIGS_attB2_R	*GGGGACCACTTTGTACAAGAAAGCTGG GTG*GGGTTATTTAAACTTGGTTATCCC AG
TRV:$1E_{GHu}$	pCLEAN-GHu-1	GFLV_VIGS_GHu_attB1_F	*GGGGACAAGTTTGTACAAAAAAGCAGG CTTA*AGTGAGGTGTGTTTGAAATGTT GC	GFLV_VIGS_attB2_R	*GGGGACCACTTTGTACAAGAAAGCTGG GTG*GGGTTATTTAAACTTGGTTATCCC AG

[a] Lower-case type indicates mutagenic nucleotides; underlined type indicates a restriction enzyme cut site (*Mlu*I or *Bsi*WI for pG_R constructs; *Sna*BI for $pF1_S$ constructs); italics indicates attB1 sites

Creation of TRV and GFLV RNA2 VIGS constructs

For VIGS experiments, all pG_R constructs were generated by amplification of regions of the GFLV-GHu symptom determinant region from pCLEAN-GHu-1 using specific primers. The pG_R plasmid contains a *Bsi*WI/*Mlu*I cloning site between the $2A^{HP}$ and $2B^{MP}$ coding regions and a short peptide sequence (TYSSVIPNYSSSFIR/GEVPYVPEDG) immediately downstream of the cloning site that is recognized by the GFLV $1D^{Pro}$ protease, allowing any inserted segments to be cleaved from the mature $2B^{MP}$ protein and to exist as a C-terminal fusion of the $2A^{HP}$ protein (Gottula, pg. 80, 2014). For cloning purposes, the symptom determinant region was expanded to include 24 nts upstream of the 3' ultimate 408 nts of the $1E^{Pol}$ coding region (thus fragments were selected from the 3' ultimate 432 nts). Amplicons were digested with *Bsi*WI and/or *Mlu*I and ligated into pG_R. Plasmids from single colonies were screened for successful insertion by sequencing with specific primers flanking the insertion site.

To generate constructs for TRV-based VIGS, primers with Gateway attB sites were designed to amplify a 191 bp fragment from the full-length cDNAs of GFLV-F13 and -GHu RNA1 clones at nucleotide position 6951 to 7141 (specifically, nts 2329 to 2475 of the $1E^{Pol}$ coding region plus the first 44 nts of the 3'UTR). The attB2 reverse primer was used for both the TRV:$1E_{F13}$ and TRV:$1E_{GHu}$ clones. Fragments were cloned into a Gateway compatible version of the pTRV2 expression vector (Liu et al. 2002a, 2002b) using Gateway cloning technology (Invitrogen). Clones were confirmed by sequencing and mobilized into the *Agrobacterium tumefaciens* strain

GV3101. The negative control vector TRV:*GFP* was generated by cloning a GFP fragment into the pTRV2 expression vector.

Agrobacterium tumefaciens-*mediated inoculation and characterization of GFLV chimeras and mutants in* N. benthamiana

GFLV RNA1 chimeras and mutants were paired with a wild type GFLV-GHu RNA2 and used for *A. tumefaciens*-mediated inoculation as previously described (Osterbaan et al. 2018). Wild type GHu RNA1 and F13 RNA1 constructs were inoculated in parallel with GFLV-GHu RNA2 for each agroinoculation experiment.

Chimeras and mutants were considered infectious if the majority of the plants used in at least two independent *A. tumefaciens*-mediated inoculation (agroinoculation) experiments of at least five *N. benthamiana* plants each became systemically infected with GFLV by three weeks post inoculation, as determined by DAS-ELISA using specific antibodies (BIOREBA, Reinach, Switzerland). In the rare occurrence of less than half of the agroinoculated *N. benthamiana* plants becoming systemically infected by three weeks post inoculation, as shown by DAS-ELISA, the virus was mechanically passaged from the initially infected *N. benthamiana* onto leaves of four week old *N. benthamiana* plants lightly dusted with Carborundum by rubbing crude sap of infected tissue ground 1:5 (wt:vol) in phosphate buffer (35 mM Na_2HPO_4, 15 mM KH_2PO_4, pH 7.0). Chimeras and mutants were considered unable to systemically infect *N. benthamiana* if none of the plants used in at least two independent agroinoculation experiments of at least five *N. benthamiana* plants each became systemically infected with GFLV by three weeks post inoculation, as determined by DAS-ELISA.

Mutants were considered symptomatic if all infected *N. benthamiana* plants displayed symptoms of vein clearing visually indistinguishable to those produced by a parallel wild type GFLV-GHu control. Mutants were considered weakly symptomatic if, in at least two independent inoculations (agroinoculation or mechanical inoculation) all infected *N. benthamiana* plants displayed less distinct symptoms of vein clearing which were visually similar to, but faded more quickly than, those of a parallel wild type GHu control. Mutants were considered inconsistently symptomatic if there was a mixture of symptomatic and asymptomatic plants among those plants which became infected.

For VIGS assays, all pG_R constructs were paired with wild type GFLV-F13 RNA1 for agroinoculation. Control inoculations were the wild type GFLV-GHu RNA1 paired with the empty pG_R plasmid (positive control for development of vein-clearing symptoms) and F13 RNA1 paired with pG_R-*Nb*-pds (positive control for virus-induced gene silencing which produces photobleaching symptoms due to silencing of the *N. benthamiana* phytoene desaturase gene). *N. benthamiana* plants were agroinoculated with the TRV:1E_{F13} or TRV:1E_{GHu} by mixing cultures (OD_{600nm} = 1.0) of chimeric pTRV2 1:1 with pTRV1 (Liu et al. 2002a).

At least two independent *A. tumefaciens*-mediated inoculation experiments of at least five plants each were conducted for chimeric and mutant GFLV RNA1 construct in this study in the presence of GFLV GHu RNA2. Agroinoculated *N. benthamiana* were maintained at 25°C in growth chambers (16 hours light cycle) and monitored for symptom development. Plants were photographed (Nikon D40, NIKKOR AF-S DX Micro 85mm lens, growth chamber lighting) once typical

symptoms of vein clearing developed. In the case of asymptomatic constructs, plants were photographed once symptoms of vein clearing developed in the wild-type GHu or photobleaching symptoms developed in pG_R-*Nb*-pds control treatments. Apical leaves were collected at 3 weeks post inoculation and used to detect systemic GFLV infection by DAS-ELISA using specific antibodies (BIOREBA, Reinach, Switzerland). Samples were recorded as positive if the absorbances values at 405 nm of duplicate wells were greater than three times the average absorbance values of the negative control samples (healthy *N. benthamiana* tissue of commensurate age). To confirm the fidelity of the virus progeny, total RNA was extracted from apical leaves at four weeks post inoculation (E.Z.N.A. Plant RNA kit) and used for RT-PCR (QIAGEN OneStep) with specific primers. DNA products of RT-PCR were sequenced at Cornell University Biotechnology Resource Center and matched to the original constructs using the SeqManPro program of the DNASTAR/Lasergene (v. 15) molecular biology software package.

***Mechanical inoculation with* in vitro *RNA transcripts and characterization of virus progeny in* N. benthamiana**

In vitro RNA transcripts of $pF1_S$ derivatives were prepared, along with transcripts of GFLV-F13 RNA2, as described in Vigne et al. (2013) and used to mechanically inoculate *Chenopodium quinoa* plants. Tissue from systemically infected *C. quinoa* plants was used as inoculum for mechanical passage to *N. benthamiana.* Infection was determined by DAS-ELISA using specific antibodies (BIOREBA) and symptoms of vein clearing were monitored on infected *N. benthamiana.* Fidelity of

progeny viruses was confirmed by sequencing IC-RT-PCR products of crude extracts of apical leaves using primer pairs G9/G10, G11/G12, G13/G14 and G15/16.

Confirmation of TRV infection by RT-qPCR

Apical tissue of *N. benthamiana* plants inoculated via *A. tumefaciens* was collected 17 dpi, flash frozen in liquid nitrogen and stored at -80°C until RNA extraction. RNA extraction, cDNA synthesis, and primer sets were as previously described (DeBlasio et al. 2018). Quantitative PCR for assessing TRV infection was performed with *Power*SYBR Green Master Mix on a QuantStudio 6 Flex 384 well System. Thermocycling conditions were 95 °C for 20 sec; 95 °C for 1 sec; and 60 °C for 20 sec (40 cycles).

Determination of GFLV titer by RT-qPCR and statistical analysis

Four-week old *N. benthamiana* plants were mechanically inoculated with GFLV wild type and mutant strains. Tissue from apical leaves was collected 13 dpi for total RNA extraction (E.Z.N.A. Plant RNA kit). This time point was chosen to align with that used by Vigne et al. (2013) for the quantification of GFLV by semi-quantitative DAS-ELISA. Total RNA was treated with amplification grade DNAse I (Invitrogen) and converted to cDNA using the Superscript IV VILO MasterMix (Invitrogen). Quantitative PCR was performed with iTaq Universal SYBR Green Supermix (Bio-Rad) and primers specific to GFLV RNA1 and three plant reference genes (actin, cyclin-dependent kinase, and elongation factor 1α) with three technical replicates per sample per target (Table 4-4) on a Bio-Rad CFX96 Touch thermocycler. Thermocycling conditions were 95°C for 3 min; 95°C for 10 sec; and 60°C for 30 sec (40 cycles). Calibrated normalized relative quantity (CNRQ) of GFLV RNA1 was

calculated in qbase+ (Biogazelle) by comparison of GFLV ΔCqs normalized against the three reference genes (Vandesompele et al. 2002) and calibrated against an inter-run calibrator sample which was included in every run (Hellemans et al. 2007). The qbase+ program uses the geNorm algorithm to calculate the stability of the reference genes (M value). The M values for actin, cyclin-dependent kinase, and elongation factor 1α in this study were 0.69, 0.87, and 0.69, respectively. CNRQ values were imported into R and one-way ANOVA (for comparison of F13-based mutants and GHu-based mutants to wild type strains) or t-test (comparison of wild type strains F13 and GHu) were performed on untransformed CNRQ values.

Table 4-4. Virus and host sequence targets and primers used in RT-qPCR to determine relative grapevine fanleaf virus (GFLV) titer.

Target	**GenBank acc. No.**	**Forward primer (5'-3')**	**Reverse primer (5'-3')**
N. benthamiana actin	JQ256516.1	AAATTACTGCACTTGCTCCTAGC	CAATCCAGACACTGTATTTCCTCTC
N. tabacum cyclin-dependent kinase	D50738.1	GTGTAGTGTACAAGGCTCGTGA	TAGCTGTGCTTGGTACTCCCTC
N. benthamiana elongation factor 1α	AY206004.1	GTACTGTCCCTGTTGGTCGT	GTAGGTCCAAAGGTCACAACCAT
GFLV-GHu RNA1	JN391442.1	CCCAAAAGTCATCGCAATGCT	GGATCAGGATATGGAAAGCAC

RESULTS

The symptom determinant lies within the ultimate 162 nucleotides (54 residues) of the GFLV-GHu 1E^{Pol} coding region.

Vigne et al. (2013) had previously narrowed down the symptom determinant to the 3' ultimate 408 nts (136 aa) of the GFLV-GHu 1E^{Pol} coding region (Figure 4-1A). We hypothesized that the minimal GFLV-GHu symptom determinant is a subset of these 408 nts/136 aa and sought to identify the minimal region of the GFLV-GHu 1E^{Pol} coding region responsible for symptom development by creating series of chimeras swapping short nucleotide sequences between GFLV strains GHu and F13 to characterize the nature of the GFLV-GHu vein clearing determinant.

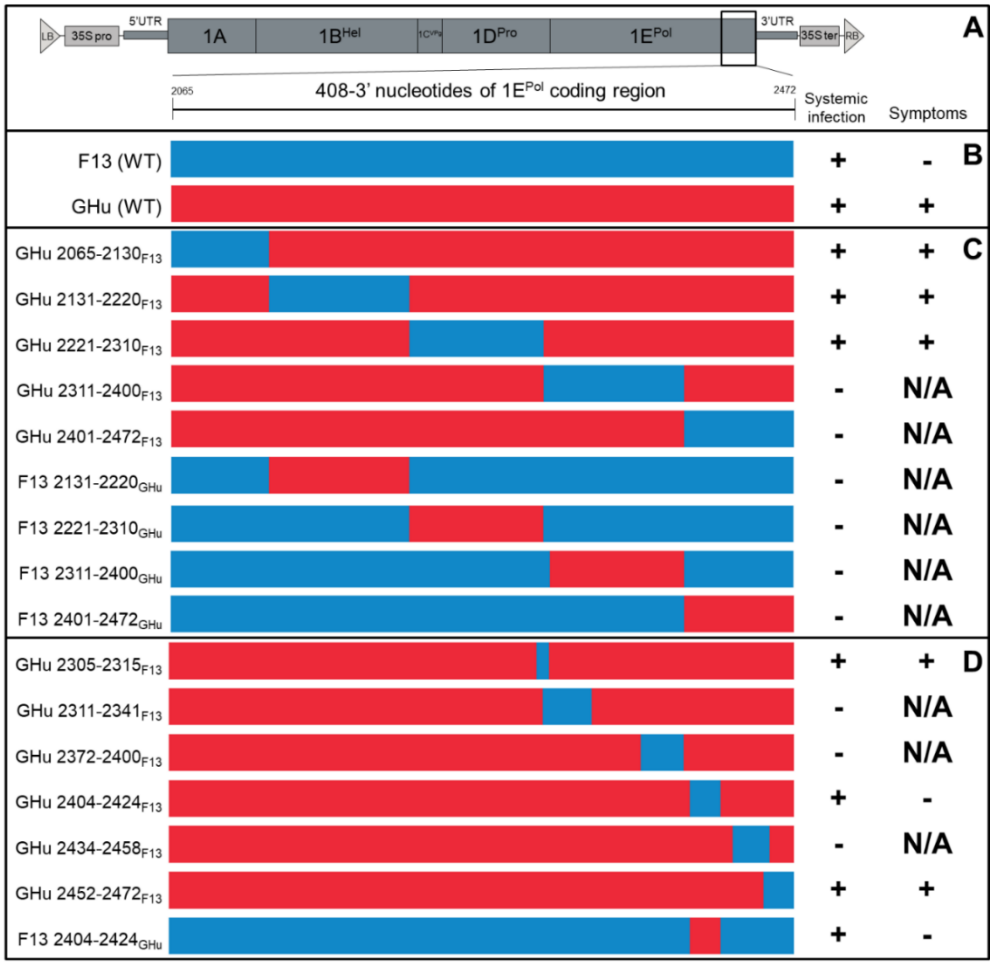


Figure 4-1. Grapevine fanleaf virus (GFLV) 1E^{Pol} chimeras and their behavior in *Nicotiana benthamiana* following agroinoculation. **A.** Schematic of GFLV RNA1 cDNA within the T-DNA region of pCLEAN-G181, a binary vector. The unshaded box at the 3' end of the 1E^{Pol} coding region indicates the 408 nucleotide stretch that was targeted for mutagenesis in the creation of the chimeras and is equivalent to the full-length boxes of panels B-D. **B.** Both RNA1 of GFLV strains F13 and GHu are systemically infectious in *N. benthamiana* via agroinoculation in the presence of GFLV-GHu RNA2, with F13 causing an asymptomatic infection and GHu causing vein clearing on systemically infected leaves. Blue boxes indicate sequences of GFLV-F13 origin; red boxes represent sequences of GFLV-GHu origin. **C.** Chimeras in which regions of 66 to 90 nts of the 408 nt symptom determinant region were exchanged between strains GHu and F13. **D.** Chimeras in which regions of 11 to 31 nts lying in the 3' ultimate 162 nts of the 1E^{Pol} coding region were exchanged between strains GHu and F13. In panels B-D: the first column shows the construct name, with the first word designating the strain origin of the RNA1 (F13 or GHu) and the numbers indicating which stretch of nucleotides of the 1E^{Pol} coding region were mutated to the cognate sequence of the other strain with the subscript indicating the source of the chimeric region. The second column is a cartoon representation of the chimeric region of the construct. The third column indicates whether the construct was able to establish systemic infection in *N. benthamiana* via agroinoculation, as determined by DAS-ELISA using specific GFLV antibodies. The fourth column indicates if the construct produced typical GFLV-GHu vein clearing.

Nine GFLV RNA1 chimeric constructs were obtained and used with a GFLV-GHu RNA2 construct in *Agrobacterium tumefaciens*-mediated inoculation (agroinoculation) of *N. benthamiana*. In replicated experiments of at least five plants each, three constructs (GHu 2065-2130_{F13}, GHu 2131-2220_{F13}, and GHu 2221-2310_{F13}) were able to establish systemic infection in *N. benthamiana*, as shown by DAS-ELISA, and produced symptoms of vein clearing typical of wild type GHu (Figure 4-1C; Figure 4-2M-O) while six constructs (four F13-based constructs, plus GHu 2311-2400_{F13} and GHu 2401-2472_{F13}) failed to establish systemic infection (Figure 4-1C). Sequencing of the 3' end of the $1E^{Pol}$ coding region of virus progeny showed that the three systemically infectious constructs retained the engineered chimeric regions, with no indels or substitutions detected within the 3' ultimate 500 nts of the $1E^{Pol}$ coding region. The chimeric regions of the three infectious and symptomatic constructs together span nts 2065 to 2310 (aa 689 to 770) of the $1E^{Pol}$ coding region (Figure 4-1C). Since alteration of this region had no effect on the ability of GFLV-GHu to produce vein clearing in *N. benthamiana*, we concluded that it is not directly involved in symptom development. Since chimeras targeting the remaining sequence (nts 2311 to 2472/ aa 771 to 824) did not cause systemic infection in *N. benthamiana* we were not able to determine what role these remaining 162 nts (54 aa) play in vein clearing.

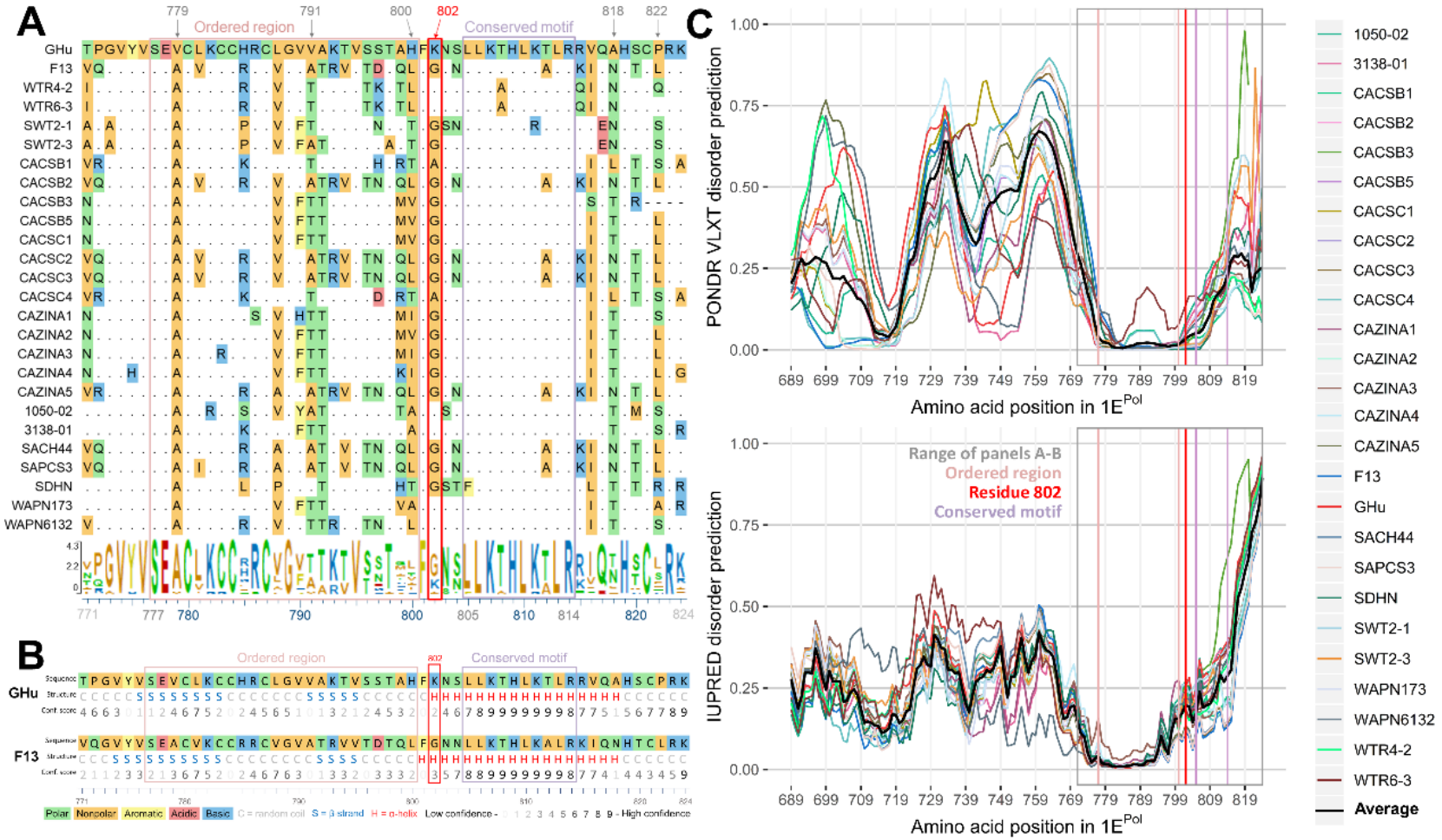
Figure 4-2. Symptoms of vein clearing or absence thereof on *Nicotiana benthamiana* plants systemically infected with select grapevine fanleaf virus (GFLV) RNA1 constructs and GFLV-GHu RNA2 following agroinoculations. **A to J.** Asymptomatic chimeric and mutant GFLV RNA1 constructs. **K to L.** GFLV RNA1 constructs generating weak symptoms. **M to AA.** Symptomatic chimeric and mutant GFLV RNA1 constructs. **BB.** Single instance of vein clearing produced by GHu K802E. **CC.** Wild type GFLV-GHu. **DD.** Wild type GFLV-F13. Apical, systemically infected leaves were photographed 8-14 days post inoculation.

Unique amino acids within the 54 ultimate residues of GFLV-GHu 1E^{Pol} are not necessary for vein-clearing symptoms.

We hypothesized that the symptom determinant may be a component unique to the GHu strain because it is the only GFLV strain known to cause vein clearing in *N.*

benthamiana (E. Vigne and M. Fuchs, *personal observation*). Therefore, unique GFLV-GHu $1E^{Pol}$ amino acids were identified within the ultimate 54 residues following sequence alignments and point mutants were created to further pinpoint the symptom determinant. A search of GenBank as of December 2017 uncovered a total of 22 GFLV entries containing sequences of the ultimate 162 nucleotides (54 residues) of $1E^{Pol}$. By aligning the 22 sequences, five amino acids unique to the GHu strain were identified within this region: V^{779}, V^{791}, H^{800}, A^{818}, and P^{822} (Figure 4-3A).

Figure 4-3. *In silico* analyses of the $1E^{Pol}$ C-terminus of grapevine fanleaf virus (GFLV) isolates. **A.** Alignment of the $1E^{Pol}$ C-terminus amino acid sequence (residues 771 to 824) of 26 GFLV isolates. Only residues differing from the GHu strain sequence (top line) are shown (a dot indicates a match to the GHu sequence). Residues unique to GFLV-GHu, identified as of December 2017, are highlighted by a grey arrow and labelled with the residue number along the top. Residue 802, which is involved in symptom expression, is highlighted in red in panels A-C. A sequence logo running along the bottom of the alignment provides an overview of the consensus sequence. In A-C, the open pink box from residues 777 to 800 denotes the region predicted to be ordered by the PONDR VL-XT and IUPRED programs for disorder prediction (see panel C). In A-C, the open lilac box from residue 805 to 814 indicates a highly conserved motif. **B.** PSSpred predictions of secondary structure of the $1E^{Pol}$ C-terminus of GFLV-GHu and -F13. Color code for amino acid chemistry is the same in panels A and B. **C.** Predictions of disorder for the C-terminus of $1E^{Pol}$ of 26 GFLV isolates. Output of the Predictor of Natural Disordered Regions (PONDR) VL-XT algorithm (top panel) and Prediction of Intrinsically Unstructured Proteins (IUPred) short order algorithm (bottom panel) for residues 681 to 824 of protein $1E^{Pol}$. In both plots, values >0.5 indicate disordered regions and values <0.5 indicate ordered (structured) regions. The thick black line in both panels shows the average prediction score across all shown isolates. The open grey box from residue 771 to 824 in both plots corresponds to the residues shown in panels A-B.


A
779
791
800
802
818
822
Ordered region
Conserved motif
GHu
F13
WTR4-2
WTR6-3
SWT2-1
SWT2-3
CACSB1
CACSB2
CACSB3
CACSB5
CACSC1
CACSC2
CACSC3
CACSC4
CAZINA1
CAZINA2
CAZINA3
CAZINA4
CAZINA5
1050-02
3138-01
SACH44
SAPCS3
SDHN
WAPN173
WAPN6132
771
777
780
790
800
805
810
814
820
824
B
Ordered region
Conserved motif
GHu
F13
Polar
Nonpolar
Aromatic
Acidic
Basic
C = random coil
S = β strand
H = α-helix
Low confidence -
- High confidence
C
PONDR VLXT disorder prediction
IUPRED disorder prediction
1.00
0.75
0.50
0.25
0.00
689
699
709
719
729
739
749
759
769
779
789
799
809
819
Amino acid position in 1E^Pol
Range of panels A-B
Ordered region
Residue 802
Conserved motif
1050-02
3138-01
CACSB1
CACSB2
CACSB3
CACSB5
CACSC1
CACSC2
CACSC3
CACSC4
CAZINA1
CAZINA2
CAZINA3
CAZINA4
CAZINA5
F13
GHu
SACH44
SAPCS3
SDHN
SWT2-1
SWT2-3
WAPN173
WAPN6132
WTR4-2
WTR6-3
Average

In a gain-of-function assay, we made a set of point mutations targeting residues 779, 791, and 800 for which the GFLV-F13 amino acid was mutated to the GHu residue using a GFLV-F13 RNA1 *in vitro* RNA transcription construct. Following mechanical inoculation of *in vitro* transcripts, each of these mutants was systemically infectious in the presence of functional GFLV-F13 RNA2, as shown by DAS-ELISA, but asymptomatic in *N. benthamiana*. Sequencing of RNA1 of viral progeny confirmed retention of the engineered mutations. In a parallel loss-of-function assay, we made point mutations targeting residues 779, 791, 800, 818, and 822 in a GFLV-GHu RNA1 binary plasmid. We also produced constructs bearing silent mutations at each of these sites to determine if changes to symptom expression, if any, were dependent on changes to the amino acid or nucleotide sequence. All mutants were used in *A. tumefaciens*-mediated inoculation of *N. benthamiana* in the presence of GFLV-GHu RNA2. All ten of the constructs were systemically infectious in *N. benthamiana*, as shown by DAS-ELISA, and produced typical GFLV-GHu vein clearing (Figure 4-2R-V). Additionally, simultaneous mutation of H^{800} and P^{822} had no effect on symptom development (Figure 4-2W-X). As before, sequencing of viral progeny confirmed retention of the engineered mutations. These findings revealed that amino acids V^{779}, V^{791}, H^{800}, A^{818}, and P^{822}, which had been identified as unique to GFLV-GHu within the 54 ultimate residues of protein $1E^{Pol}$, are not necessary for vein clearing in *N. benthamiana* (Table 4-5).

Table 4-5. Grapevine fanleaf virus **(**GFLV) single and double amino acid mutants targeting $1E^{Pol}$ residues 779, 791, 800, 818, and 822 and their behavior in *Nicotiana benthamiana*. Bold type indicates mutated codons.

$1E^{Pol}$ codon[a] >	779		791		800		818		822		
Construct name	NT	AA	NT	AA	NT	AA	NT	AA	NT	AA	Symptoms[b]
F13 (WT)	GCG	A	GCA	A	CTT	L	AAT	N	CTT	L	-
GHu (WT)	GTG	V	GTC	V	CAC	H	GCT	A	CCG	P	+
GHu V779A	**GCG**	**A**	GTC	V	CAC	H	GCT	A	CCG	P	+
GHu V779V	**GTT**	**V**	GTC	V	CAC	H	GCT	A	CCG	P	+
GHu V791A	GTG	V	**GCA**	**A**	CAC	H	GCT	A	CCG	P	+
GHu V791V	GTG	V	**GTT**	**V**	CAC	H	GCT	A	CCG	P	+
GHu H800L	GTG	V	GTC	V	**CTT**	**L**	GCT	A	CCG	P	+
GHu H800H	GTG	V	GTC	V	**CAT**	**H**	GCT	A	CCG	P	+
GHu A818N	GTG	V	GTC	V	CAC	H	**AAT**	**N**	CCG	P	+
GHu A818A	GTG	V	GTC	V	CAC	H	**GC C**	**A**	CCG	P	+
GHu P822L	GTG	V	GTC	V	CAC	H	GCT	A	**CTT**	**L**	+
GHu P822P	GTG	V	GTC	V	CAC	H	GCT	A	**CCT**	**P**	+
GHu H800L + P822L	GTG	V	GTC	V	**CTT**	**L**	GCT	A	**CTT**	**L**	+
GHu H800H + P822P	GTG	V	GTC	V	**CAT**	**H**	GCT	A	**CCT**	**P**	+
F13 A779V	**GTG**	**V**	GCA	A	CTT	L	AAT	N	CTT	L	-
F13 A791V	GCG	A	**GTC**	**V**	CTT	L	AAT	N	CTT	L	-
F13 L800H	GCG	A	GCA	A	**CAC**	**H**	AAT	N	CTT	L	-

[a] Nucleotide (NT) and amino acid (AA) composition is indicated for each $1E^{Pol}$ codon
[b] Symptoms (+) refer to typical GFLV-GHu vein clearing; (-) denotes absence of symptoms

GFLV-GHu symptoms of vein clearing are abolished by mutation of nucleotides 2404 to 2424 (residues 801 to 808) of the $1E^{Pol}$ coding region.

Given that RNA1 chimeras targeting large regions of the $1E^{Pol}$ protein often did not cause systemic infection (Figure 4-1C), we further investigated the symptom determinant region by creating chimeras targeting smaller regions. We produced six such chimeras based on GHu RNA1. In replicated experiments of at least five plants each, three constructs (GHu 2311-2341_{F13}, GHu 2372-2400_{F13}, and GHu 2434-2458_{F13}) failed to establish systemic infection in *N. benthamiana* via agroinoculation,

once again highlighting the sensitivity of this region of $1E^{Pol}$ to modifications which abolish the ability of the virus to systemically infect plants. Of the three systemically infectious constructs, one failed to produce typical GHu symptoms (GHu 2404-2424_{F13}; Figure 4-1D; Figure 4-2E). This was our first incidence of a construct with a GHu RNA1 backbone which was unable to produce vein clearing. To determine if the cognate chimera could produce symptoms with a F13 RNA1 background, we produced F13 2404-2424_{GHu}. Though this construct was systemically infectious, as shown by DAS-ELISA, it remained asymptomatic in *N. benthamiana* (Figure 4-1D; Figure 4-2A). Sequencing of RT-PCR products showed that RNA1 of virus progeny of systemically infectious constructs maintained the targeted mutations. This work revealed that GFLV-GHu vein clearing can be abolished by mutating nucleotides 2404-2424 (residues 801-808) of the $1E^{Pol}$coding region. This region was further dissected to determine the nature of the vein clearing determinant.

GFLV-GHu vein clearing is dependent on the identity of residue 802 of the $1E^{Pol}$ protein.

We hypothesized that a single amino acid within residues 801 to 808 of the $1E^{Pol}$ protein is causing vein clearing in *N. benthamiana*. Of these seven amino acids, only two are not conserved between GFLV strains GHu and F13: K^{802}/G^{802} and S^{804}/N^{804} (GHu/F13, respectively). We created single and double residue mutants (as well as silent mutations) targeting these sites in RNA1 of both GFLV-GHu and GFLV-F13 (Table 4-6) and inoculated them to *N. benthamiana* via *A. tumefaciens* in the presence of GFLV-GHu RNA2. Notably, mutation of GFLV-GHu $1E^{Pol}$ lysine 802 to glycine (construct GHu K802G) abolished symptoms, while silent mutation at this

site (GHu K802K [G2406A]) had no effect on systemic infection and development of vein clearing (Table 4-6; Figure 4-2H and Y). Also, mutation of GFLV-GHu $1E^{Pol}$ serine 804 to asparagine (GHu S804N), as well as silent mutation at this site (GHu S804S [AGC to TCT]), had no effect on systemic infection and development of vein clearing (Table 4-6; Figure 4-2AA). Furthermore, reciprocal mutations to GFLV-F13 RNA1 (F13 G802K, F13 G802G [C2406G], F13 N804S, and F13 N804N [T2412C]) all produced asymptomatic systemic infections typical of wild type GFLV-F13 (Table 4-6; Figure 4-2B and D). Sequencing of RT-PCR products from RNA1 of virus progeny of single and dual mutants showed that the targeted mutations were maintained. These data confirmed our hypothesis by indicating that GFLV-GHu vein clearing is dependent on the identity of residue 802 of the $1E^{Pol}$ protein.

Table 4-6. Grapevine fanleaf virus **(**GFLV) single and double amino acid mutants targeting the RNA1-encoded $1E^{Pol}$ residues 802 and 804 and their behavior in *Nicotiana benthamiana*.

		$1E^{Pol}$ codon 802		$1E^{Pol}$ codon 804			
Wild type (WT) virus or construct name	RNA 1	NT	AA	NT	AA	Systemic infection [a]	Sym [b]
F13 (WT)	F13	GGC	G	AAT	N	+	-
GHu (WT)	GHu	AAG	K	AGC	S	+	+
F13 G802K	F13	AAG	K	AAT	N	+	-
F13 G802G [C2406G]	F13	GGG	G	AAT	N	+	-
F13 G802K+N804S	F13	AAG	K	AGC	S	+	-
F13 G802G+N804N	F13	GGT	G	AAC	N	+	-
F13 N804S	F13	GGC	G	AGT	S	+	-
F13 N804N [T2412C]	F13	GGC	G	AAC	N	+	-
GHu K802G	GHu	GGC	G	AGC	S	+	-
GHu K802K [G2406A]	GHu	AAA	K	AGC	S	+	+
GHu K802G+S804N	GHu	GGC	G	AAC	N	+	-
GHu K802K+S804S	GHu	AAA	K	TCT	S	+	+
GHu S804N	GHu	AAG	K	AAC	N	+	+
GHu S804S [AGC to TCT]	GHu	AAG	K	TCT	S	+	+
GHu K802A [GCC]	GHu	GCC	A	AGC	S	+	-
GHu K802A [GCG]	GHu	GCG	A	AGC	S	+	-
GHu K802Q	GHu	CAG	Q	AGC	S	+	-
GHu K802E	GHu	GAG	E	AGC	S	+	Incon. [c]
GHu K802N	GHu	AAC	N	AGC	S	+	Weak [d]
GHu K802P	GHu	CCG	P	AGC	S	+	Weak [d]
GHu K802R	GHu	CGT	R	AGC	S	-	N/A
GHu K802Y	GHu	TAC	Y	AGC	S	-	N/A
GHu ΔS804	GHu	AAG	K	---	-	-	N/A
F13 G802P	F13	CCC	P	AAT	N	-	N/A

[a]Systemic infection (+) or lack thereof (-) was determined in apical leaves by DAS-ELISA using GFLV specific antibodies three weeks post *Agrobacterium tumefaciens*-mediated inoculation.

[b]Presence of vein clearing characteristic of wild type GFLV-GHu (+) or lack thereof (-) was assessed 1 to 2 weeks post-inoculation.

[c]This construct generally produced asymptomatic systemic infections. However, in a single mechanical inoculation experiment, one of the four systemically infected plants produced symptoms of vein clearing. Sequencing confirmed retention of the K802E mutation in both symptomatic and asymptomatic plants.

[d]These constructs produced weak symptoms of vein clearing, which was fainter than that produced by wild type GHu and faded more quickly (fading within 2 to 3 days of symptom development, compared to 5 to 8 days for wild type).

***The nature of residue 802 of the 1E^{Pol} protein of GFLV-GHu modulates symptom expression in* N. benthamiana.**

To investigate whether the nature of amino acid 802 of GFLV-GHu 1E^{Pol} modulates development of vein clearing in *N. benthamiana*, we mutated this residue to several other amino acids covering a range of chemistries and structures. Mutation of GFLV-GHu 1E^{Pol} lysine 802 to alanine (GHu K802A [GCC] and GHu K802A [GCG]) or glutamine (GHu K802Q) abolished symptoms of vein clearing in *N. benthamiana* (Table 4-6; Figure 4-2F-G and J). Mutation of this residue to glutamic acid, asparagine, or proline (GHu K802E, N, and P, respectively) produced symptoms which did not perfectly mimic those of wild type GFLV-GHu. Mutant GHu K802E generally produced asymptomatic infections; however, in one inoculation experiment, one out of four plants that became systemically infected, as shown by DAS-ELISA, displayed distinct symptoms of vein clearing (Figure 4-2BB). Sequencing of the 3' end of the 1E^{Pol} coding region of the virus progeny confirmed that the K802E mutation was stably maintained in both the symptomatic plant and asymptomatic plants. Meanwhile, GHu K802N and GHu K802P (Figure 4-2K-L) produced symptoms of vein clearing which was less severe than wild type GHu and faded more rapidly (within 2 days of symptom development, compared to 5 to 8 days for wild type GHu). Mutation of this residue to arginine (R) or tyrosine (Y), as well as deletion of residue S^{804}, rendered the virus unable to systemically infect *N. benthamiana* following agroinoculation (Table 4-6). For all systemically infectious constructs, sequencing of virus progeny confirmed their fidelity, with no reversions to wild type

or other substitutions detected. This work showed the nature of residue 802 of the $1E^{Pol}$ protein of GFLV-GHu modulates expression of vein clearing in *N. benthamiana*.

Residue 802 of the GFLV-GHu protein $1E^{Pol}$ is flanked by strongly predicted secondary structures.

Due to the difficulties in obtaining functional mutants in the C-terminal region of the $1E^{Pol}$ protein, we hypothesized that this region is highly structured, particularly the region flanking residue 802. While tertiary structure prediction for $1E^{Pol}$ via the Iterative Threading ASSembly Refinement algorithm for protein structure prediction (I-TASSER) (Yang et al. 2014) produces a model with a good confidence score for the RdRp "core" of this protein by threading the GFLV sequence onto solved RdRp structures, prediction of the C-terminus that includes the symptom determinant has a poor confidence score due to lack of homology to solved structures (data not shown). This leaves us with little insight into the actual tertiary structure of this region. Applying the Protein Secondary Structure prediction algorithm (PSSpred), part of the I-TASSER suite, to predict the secondary structure of the C-terminus of protein $1E^{Pol}$ of both GFLV strains GHu and F13, we showed that residues 807 to 812 of both proteins are predicted to be in a helix with 95% confidence (Figure 4-3B). This putative helix lies within a 10-residue motif (^{805}LLKT/AHLK/RT/ALR^{814}) of unknown function which is highly conserved among all GFLV isolates for which sequence information is available in the NCBI GenBank (Figure 4-3A). Additionally, the Predictor Of Natural Disordered Regions (PONDR) VL-XT (Li et al. 1999) and the IUPred (Prediction of Intrinsically Unstructured Proteins) (Dosztányi et al. 2005) algorithms showed residues 777 to 800 to be ordered in GFLV strains GHu and F13,

as well as 24 other GFLV isolates, with high confidence, i.e. very low disorder prediction values. In contrast, no disordered regions were predicted with confidence (Figure 4-3C). These *in silico* analyses indicated that residue 802 of the $1E^{Pol}$ protein is flanked by strongly predicted stable secondary structures.

***Virus titer is not associated with GFLV-GHu symptomatology in* N. benthamiana.** We determined the relative GFLV titer of at least three plants per mutant by RT-qPCR targeting GFLV RNA1 using specific primers (Figure 4-4; Table 4-7). In comparing the relative titers of mutants derived from GFLV-GHu RNA1 to the wild type strains GHu and F13, we detected no significant differences between any mutant pairs (one-way ANOVA, $p=0.20$) or between the wild type strain F13 and GHu (unpaired t-test, $p = 0.36$). We also compared several F13-based mutants to the wild type strains and found no significant differences in relative GFLV RNA1 accumulation (one-way ANOVA, $p=0.10$). Altogether, RT-qPCR did not reveal any significant difference in virus titer across mutants, as measured by RNA1 accumulation.

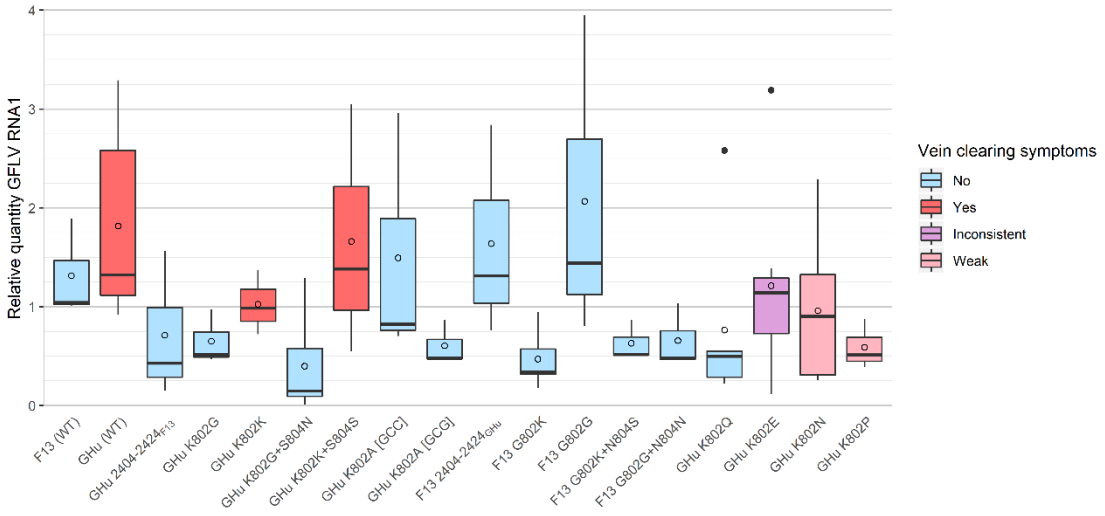


Figure 4-4. Boxplot of relative titer of grapevine fanleaf virus (GFLV) wild type (WT) strains F13 and GHu and GFLV 1E^{Pol} mutants as determined by RT-qPCR. Each box summarizes at least three biological replicates (plants) systemically infected with the same GFLV mutant or strain. The lower and upper hinges represent the 25th and 75th percentiles, respectively. Lower and upper whiskers show the smallest and largest values, respectively, with outliers (values more than 1.5 times the interquartile range beyond the hinge) shown as dots. The middle bar marks the median. Circles (O) superimposed on each box denote the mean.

Table 4-7. Average relative titer of grapevine fanleaf virus (GFLV) wild type (WT) and mutants as determined by RT-qPCR.

Mutant	Mean CNRQ[a]	SD[b]
F13 (WT)	1.31	0.50
GHu (WT)	1.82	1.04
GHu 2404-2424$_{F13}$	0.71	0.75
GHu K802G	0.65	0.28
GHu K802K	1.02	0.33
GHu K802G+S804N	0.40	0.51
GHu K802K+S804S	1.66	1.27
GHu K802A [GCC]	1.49	1.27
GHu K802A [GCG]	0.60	0.22
F13 2404-2424$_{GHu}$	1.64	1.08
F13 G802K	0.47	0.30
F13 G802G	2.07	1.66
F13 G802K+N804S	0.63	0.21
F13 G802G+N804N	0.66	0.32
GHu K802E	1.21	0.98
GHu K802Q	0.76	0.90
GHu K802N	0.96	0.78
GHu K802P	0.59	0.25

[a] Calibrated normalized relative quantity (CNRQ); GFLV RNA1 RT-qPCR ΔCq normalized against three host reference genes and calibrated to an interrun calibrator sample.
[b] Standard deviation

Virus-induced gene silencing (VIGS) is not involved in GFLV-GHu 1E^{Pol} symptom expression.

While our results strongly suggested that the GFLV-GHu symptom determinant acts on a protein level, the silent mutation of residue 802 (GHu K802K [G2406A]) introduced only a single base change to the nucleotide sequence of RNA1, a change possibly insufficient to abolish nucleotide sequence complementarity to putative target host sequences. So, given mounting evidence that silencing of host genes via vsiRNAs plays a pivotal role in plant virus pathogenesis (Zhang et al. 2015), we used VIGS to investigate the involvement of RNA silencing in GFLV-GHu vein clearing in *N. benthamiana*. We first introduced the full-length (408 nts) or small fragments (102 to 312 nts) of the GFLV-GHu 1E^{Pol} symptom determinant region into a VIGS vector based on GFLV-GHu RNA2 known as pG$_R$ (Gottula, pg. 125, 2014). The pG$_R$ plasmid contains a *Bsi*WI/*Mlu*I cloning site between the 2A^{HP} and 2B^{MP} coding regions and a short polypeptide sequence (TYSSVIPNYSSSFIR/GEVPYVPEDG) immediately downstream of the cloning site for recognition by the GFLV 1D^{Pro} protease, allowing any inserted segments to be cleaved from the mature 2B^{MP} protein and to exist as a C-terminal fusion of the 2A^{HP} protein. Fragments of the GHu 1E^{Pol} coding region were cloned into this site as in-frame, reverse complement, or frame-shift insertions (Figure 4-5). Results showed that none of the chimeric pG$_R$ constructs, regardless of their size and orientation, produced vein-clearing symptoms in *N. benthamiana* when paired with a GFLV-F13 RNA1 construct in replicated agroinoculation experiments of at least five plants each, although all of them were systemically infectious in *N. benthamiana*, as shown by

DAS-ELISA. Sequencing of RNA2 of progeny viruses showed that all pG_R constructs retained the inserted sequences, with no indels or substitutions. These findings suggested that the 408 nt symptom determinant previously identified (Vigne et al. 2013) does not cause vein clearing when expressed from GFLV-GHu RNA2.

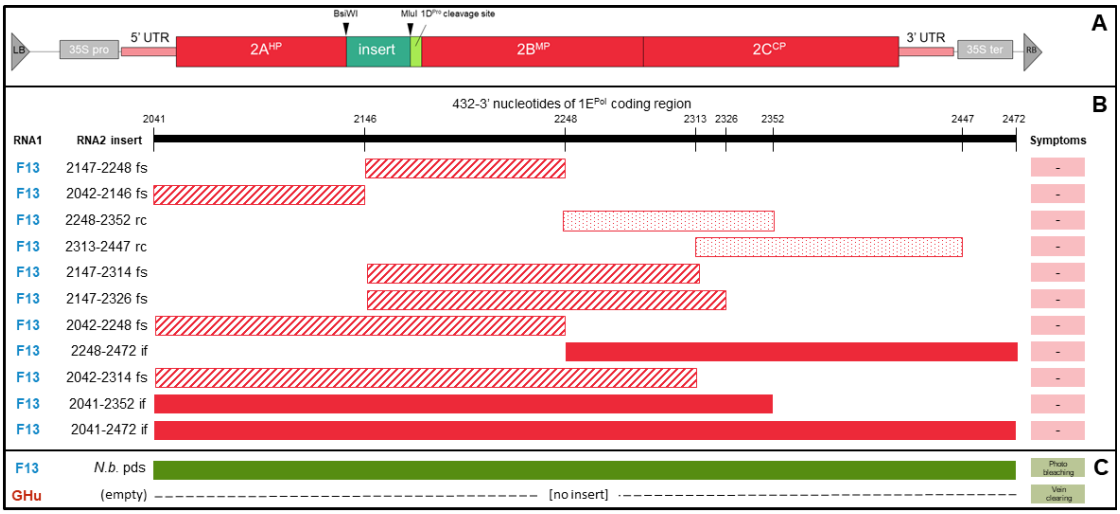


Figure 4-5. Grapevine fanleaf virus (GFLV) strain GHu RNA2 constructs harboring fragments of the GFLV-GHu symptom determinant do not produce vein-clearing symptoms in *Nicotiana benthamiana*. **A.** Schematic of the T-DNA region of pG_R, a binary plasmid containing the cDNA sequence of GFLV-GHu RNA2 within a cauliflower mosaic virus (CaMV) 35S expression cassette. The RNA2 sequence has been modified to allow insertion of fully-translatable exogenous sequences between the coding regions of protein $2A^{HP}$ and $2B^{MP}$ via *Bsi*WI and *Mlu*I. A 75 nt (25 aa) sequence at the 3' end of the insert is recognized by the native GFLV protease, $1D^{Pro}$, resulting in a C-terminal fusion of the insert to the $2A^{HP}$ protein. LB: left border; 35S pro: CaMV 35S promoter; 5'UTR: 5' untranslated region; $2A^{HP}$, helper protein; $2B^{MP}$, movement protein; $2C^{CP}$, coat protein; 3'UTR: 3' untranslated region; 35S ter: CaMV 35S terminator; RB: right border. **B.** pG_R constructs containing fragments of the 3' ultimate 432 nucleotides of the $1E^{Pol}$ coding region of GFLV-GHu and their behavior in *N. benthamiana*. The long hashed, dotted, or filled boxes are schematic representations of the fragments which were cloned into pG_R. The scale for this schematics is shown by the black ruler at the top of the panel. Hashed bars indicate the fragment was inserted frame-shifted relative to the native sequence. Dotted bars indicate the reverse complement of the fragment was inserted. Solid bars indicate the fragment was inserted in-frame relative to the native sequence. **C.** pG_R control constructs. *N.b.* pds contains a 501 nt fragment of the *N. benthamiana* phytoene desaturase coding region and produces photo-bleaching symptoms in *N. benthamiana* via virus-induced gene silencing. GFLV-GHu RNA1 produces typical vein-clearing symptoms when paired with empty pG_R. In **B to C,** the first column indicates which RNA1 the construct was co-agroinoculated with. The second column indicates the identity of the pG_R insert. fs: frame-shift; rc: reverse complement; if: in-frame. The third column shows a cartoon of the insert. Note that B and C have different scales. The fourth column indicates if the RNA1 + RNA2 combination produced symptoms in *N. benthamiana* following agroinoculation.

To rule out the possibility that infection with GFLV-F13 RNA1 may somehow suppress or mask symptoms of vein clearing in *N. benthamiana*, we co-inoculated *N. benthamiana* with F13 RNA1 and GHu RNA1, in the presence of GHu RNA2, via

agroinoculation. These plants displayed typical symptoms of vein clearing on apical leaves. Both GFLV-F13 and GHu RNA1 were detected by RT-PCR with strain-specific primers in apical leaves followed by sequencing confirmation. We also tested whether pG_R could be suppressive of vein clearing by agroinoculation of the empty pG_R construct with GFLV-GHu RNA1. This combination produced systemic infections displaying typical GHu vein clearing, indicating that pG_R does not interfere with the development of vein clearing in *N. benthamina*. Furthermore, we confirmed that pG_R is capable of inducing RNA silencing by using pG_R-*Nb*-pds, which contains a 501 nt segment of *N. benthamiana* phytoene desaturase (*PDS*) gene, with GFLV-F13 RNA1 in *A. tumefaciens*-mediated inoculations. The resulting systemically infected plants displayed photobleaching symptoms typical of silencing of *PDS*, confirming the pG_R-based constructs are competent to induce gene silencing.

To further rule out the potential role of RNA silencing in *N. benthamiana* vein clearing, the GFLV $1E^{Pol}$ symptom determinant sequence was introduced into a tobacco rattle virus (TRV) RNA2 based VIGS vector. While the GFLV VIGS assays utilized the full symptom determinant region, the TRV experiments focused on nucleotides 2404 to 2424 of $1E^{Pol}$, which are involved in symptom development. Plants did not show GFLV-GHu symptoms of vein clearing but typical TRV symptoms (stunting, mild mottling) and tested positive for TRV by RT-qPCR following *A. tumefaciens*-mediated inoculation in *N. benthamiana*. TRV:*N.b.*pds displayed photobleaching symptoms, as expected. To test whether GFLV-GHu symptoms of vein clearing are dominant and distinguishable from TRV symptoms, TRV and GFLV-GHu were co-inoculated via *A. tumefaciens* into *N. benthamiana*.

Plants showed typical GHu symptoms starting 5 to 6 days post inoculation, then developed typical TRV symptoms (mottling and stunting) at 7 to 8 days post inoculation. Plants were confirmed to be systemically co-infected with GFLV and TRV by DAS-ELISA and RT-qPCR, respectively. These findings showed that the introduction of the symptom determinant region of interest into a TRV vector did not cause vein clearing in *N. benthamiana* (Figure 4-6).

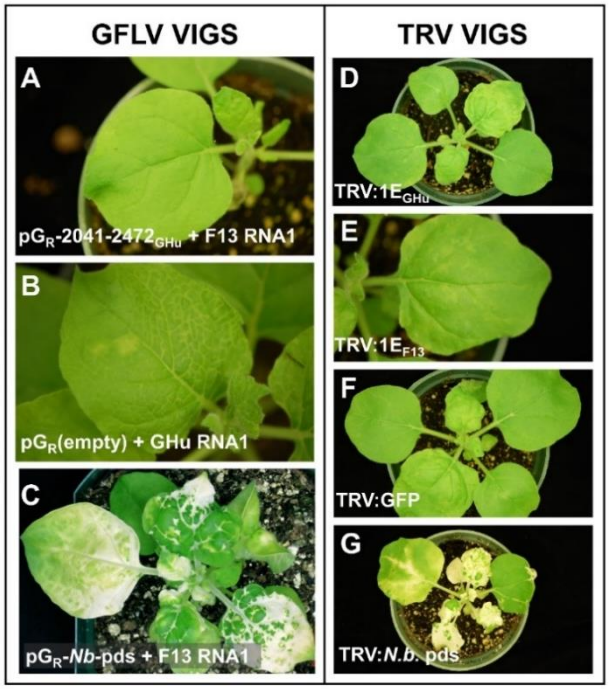


Figure 4-6. The grapevine fanleaf virus (GFLV) strain GHu 1E^{Pol} symptom determinant region does not induce vein clearing in *N. benthamiana* when introduced via virus-induced gene silencing (VIGS) vectors. **A.** Asymptomatic plant agroinoculated with GFLV-F13 RNA1 and pG$_R$, a VIGS construct based on GFLV-GHu RNA2, containing the 3' ultimate 432 nts of the 1E^{Pol} coding region from GFLV-GHu. **B.** Plant displaying typical GFLV-GHu vein clearing on apical, uninoculated leaves following agroinoculation with the empty pG$_R$ construct and GFLV-GHu RNA1. **C.** Plant displaying photobleaching symptoms, indicative of silencing of the phytoene desaturase (*PDS*) gene, following agroinoculation with GFLV F13 RNA1 and pG$_R$-*Nb*-pds. **D-E.** Plants agroinoculated with a TRV RNA2 VIGS construct containing nucleotides 6851 to 7141 of RNA1 of GFLV-GHu or F13 (last 147 nts of 1E^{Pol} plus 44 nts of 3'UTR) produce typical TRV symptoms (mottling, stunting), but do not produce vein clearing. **F.** Plant displaying typical mottling and stunting symptoms following agroinoculation with TRV:GFP. **G.** Plant displaying photobleaching symptoms, indicative of silencing of the phytoene desaturase (*PDS*) gene, following agroinoculation with TRV:*N.b.*pds.

DISCUSSION

In investigating the nature of a symptom determinant responsible for vein clearing of GFLV strain GHu in *N. benthamiana*, we found that symptom development is dependent on the identity of the amino acid at position 802 of protein $1E^{Pol}$. Symptoms of vein clearing develop when the native lysine (K) is present at this position (Table 4-6; Figure 4-2M-Y, and CC), while substitution of this residue to glycine (G, the native residue of GFLV-F13), alanine (A, the native residue of several GFLV isolates), or glutamine (Q) abolishes vein clearing (Table 4-6; Figure 4-2F-H, and J). Glutamic acid (E) in position 802 of protein $1E^{Pol}$ produced inconsistent symptoms (Figure 4-2BB, Table 4-6) while asparagine (N) and proline (P) produced weak symptoms (Table 4-6; Figure 4-2K-L). These results, along with several other lines of evidence, suggest the involvement of a protein interaction between GFLV-GHu $1E^{Pol}$ and *N. benthamiana* in the development of symptoms of vein clearing. Alternatively, the K residue may be involved in intramolecular interactions regulating $1E^{Pol}$ structure.

Mutational analysis of the amino acid at position 802 yielded insights into the function of the K residue in that position for GFLV symptom expression. Substitution of proline at position 802 produced weak symptoms of vein clearing in *N. benthamiana.* However, proline is a non-polar amino acid, like alanine and glycine, and has a unique cyclic structure in which the side chain forms a ring with the peptide backbone. This drastically reduces the rotational freedom of the backbone and this reduction in free entropy means proline is often found at “turns” of a protein (Betts and Russell 2007). That the proline substitution mutant is able to produce symptoms,

though weakly, favors the hypothesis that modulation of symptom development by this domain of $1E^{Pol}$ is due to conformational changes to the $1E^{Pol}$ protein influenced by the residue in this position rather than direct interaction between this residue and host plant proteins. This model contrasts with a study on cucumber green mottle mosaic virus in which single amino acid mutations to the methyltransferase domain of the viral RdRp produced symptom attenuation. Protein modeling of the methyltransferase domain showed the mutated residue to be surface exposed and amino acid replacements at this site appeared to have no effect on the domain's organization, suggesting the mutated residue may be modulating protein interactions necessary for symptom development without changes to the viral protein structure (Liu et al. 2017).

The seven canonical motifs of plant virus RdRps lie within the N-terminal 433 residues of protein $1E^{Pol}$ (Ritzenthaler et al. 1991; Vigne et al. 2013), while the C-terminal region of $1E^{Pol}$ bears no significant sequence homology to known sequences, apart from the RdRps of other nepoviruses (Vigne et al., 2013; this study). This illustrates the modular structures and multifunctionality of viral proteins (Leisner and Schoelz 2018; Valli et al. 2018). This dearth of homology also explains why tertiary structure prediction for $1E^{Pol}$ via I-TASSER (Yang et al. 2014) produces a model with a good confidence score for the RdRp "core" of this protein, while structure prediction of the C-terminus that includes the symptom determinant has a poor confidence score (data not shown). This leaves us with little insight into the actual tertiary structure of this region. Nonetheless, our study confirms earlier findings (Vigne et al., 2013) on the

multifunctionality of protein $1E^{Pol}$ with an implicated role in replication and a documented role in symptom development.

While *de novo* prediction of protein tertiary structure remains difficult, prediction of secondary structure from amino acid sequence alone can be done with a fair amount of accuracy. Thus, we were able to glean a few insights into the nature of the structure(s) surrounding residue 802 in GFLV-GHu $1E^{Pol}$ by looking for evidence of stable secondary structures surrounding this amino acid. Both the IUPred and PONDR VL-XT algorithms strongly predict a highly ordered region from approximately residue 777 to 800 of $1E^{Pol}$ in the 26 representative GFLV isolate sequences used for these analyses. Additionally, the PSSPred algorithm for secondary structure prediction, part of the I-TASSER suite (Yang et al. 2014), strongly predicts a helical structure for residues 807-812 of protein $1E^{Pol}$ for both GFLV-F13 and GFLV-GHu (Figure 4-3B). This region lies within a 10-residue motif (^{805}LLKT/AHLK/RT/ALR^{814}) of unknown function which is highly conserved among GFLV isolates for which sequence information is available in GenBank (Figure 4-3A). The precise function of either of these predicted structured regions remains unknown. It is unsurprising to find symptom determinants within regions of virus proteins which lack functional annotation, such as the case of the highly-divergent N-terminus of carlavirus cysteine-rich proteins (Fujita et al. 2018). The existence of two well-conserved motifs flanking residue 802, along with the knowledge that genetic manipulation of the C-terminus of GFLV $1E^{Pol}$ often renders the virus unable to systemically infect *N. benthamiana* (Vigne et al., 2013; this study), provides further

credence to the hypothesis that the identity of residue 802 determines the stability of a yet to be identified functional domain in the C-terminus of $1E^{Pol}$.

Our work confirmed that symptoms of vein clearing in *N. benthamiana* are not correlated with overall virus titer or accumulation of $1E^{Pol}$ (Vigne et al. 2013). While we acknowledge that RNA1 accumulation may be an incomplete proxy for $1E^{Pol}$ titer given evidence in ToRSV, another nepovirus, that host responses to viral infection may repress translation without reducing RNA titer (Ghoshal and Sanfaçon 2014; Paudel et al. 2018), the earlier findings of Vigne et al. (2013) appear to hold true for the mutants produced for this study. There was no consistent pattern between GFLV RNA1 titer and mutant symptomatology (Figure 4-4). While great strides have been recently made to advance our understanding of the mechanistic underpinnings of GFLV symptom expression in model host plants by documenting the role of proteins $1E^{Pol}$ (Vigne et al., 2013; this study) and $2A^{HP}$ (Martin et al., 2018), limited information is available on virus-host interactions in grapevine, the natural host. There have been attempts to correlate GFLV sequences and symptomatology in grapevine (Elbeaino et al. 2014) but no conclusions could be drawn on a specific symptom determinant. It is anticipated that reverse genetics approaches applied to model hosts will target grapevine in the future. Indeed, work is underway to transfer some of the mutants developed in this study to grapevine. It will be interesting to analyze in the natural host of GFLV the phenotype of these mutants and of the wild type strains GHu and F13, which were originally isolated from diseased grapevines (Huss et al. 1989; Pinck et al. 1988) on which they cause a mild mosaic.

Lysine 802 of GFLV-GHu $1E^{Pol}$ is indispensable but not sufficient for vein clearing in *N. benthamiana*. Indeed, substituting glycine 802 with lysine in protein $1E^{Pol}$ of the asymptomatic GFLV strain F13 (F13 G802K) is not sufficient to produce vein clearing in this strain (Table 4-6; Figure 4-2B). It appears we do not yet have the full story of the GFLV-GHu symptom determinant: some component other than $1E^{Pol}$ residue 802 must play a role in symptom development in *N. benthamiana.* While VIGS experiments ruled out a potential role of RNA silencing in GFLV-GHu symptom development, it is interesting to note that GFLV VIGS constructs bearing translatable portions of the $1E^{Pol}$ C-terminus, including the region surrounding residue 802, failed to elicit vein clearing in *N. benthamiana.* Thus, it appears that the GFLV-GHu symptom determinant is dependent on its native RNA1 context for its role in symptom development. Furthermore, there is precedent for a viral symptom determinant (modulator) and a viral symptom elicitor to be separate components. For example, Komatsu et al. (2011) found that both symptomatic and asymptomatic strains of plantago asiatica mosaic virus (PlAMV) encode a necrosis-inducing form of the HEL domain of the RdRp; however, symptom development by PlAMV in *N. benthamiana* is dependent on RdRp accumulation, which is controlled by the POL domain of RdRp. More work is needed to address whether a similar situation applies to the GFLV-GHu $1E^{Pol}$ protein.

Little is known of the molecular context of $1E^{Pol}$ during GFLV infection, apart from its association with endoplasmic reticulum membranes that are the site of nepovirus replication (Ritzenthaler et al. 2002; Fuchs et al. 2017). Further work, such as structural determination using NMR spectroscopy or X-ray crystallography and

perhaps co-immunoprecipitation of $1E^{Pol}$ and mass spectroscopy identification of protein interactors, are needed to determine what other viral or host components are involved in GFLV development of vein clearing in *N. benthamiana*.

REFERENCES

Betts, M. J., and Russell, R. B. 2007. Amino-acid properties and consequences of substitutions. Pages 311–342 in: Bioinformatics for Geneticists, M.R. Barnes, ed. John Wiley & Sons, Ltd, Chichester, UK.

Culver, J. N., and Padmanabhan, M. S. 2007. Virus-induced disease: altering host physiology one interaction at a time. Annu. Rev. Phytopathol. 45:221–243

DeBlasio, S. L., Rebelo, A. R., Parks, K., Gray, S. M., and Heck, M. C. 2018. Disruption of chloroplast function through downregulation of phytoene desaturase enhances the systemic accumulation of an aphid-borne, phloem-restricted virus. Mol. Plant-Microbe Interact. 31:1095–1110

Digiaro, M., Elbeaino, T., and Martelli, G. P. 2017. *Grapevine fanleaf virus* and other Old World nepoviruses. Pages 47–82 in: Grapevine Viruses: Molecular Biology, Diagnostics and Management, B. Meng, G.P. Martelli, D.A. Golino, and M. Fuchs, eds. Springer International Publishing, Cham.

Dosztányi, Z., Csizmok, V., Tompa, P., and Simon, I. 2005. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics 21:3433–3434

Duff-Farrier, C. R. A., Bailey, A. M., Boonham, N., and Foster, G. D. 2015. A pathogenicity determinant maps to the N-terminal coat protein region of the Pepino mosaic virus genome. Mol. Plant Pathol. 16:308–315

Elbeaino, T., Kiyi, H., Boutarfa, R., Minafra, A., Martelli, G. P., and Digiaro, M. 2014. Phylogenetic and recombination analysis of the homing protein domain of grapevine fanleaf virus (GFLV) isolates associated with “yellow mosaic” and “infectious malformation” syndromes in grapevine. Arch. Virol. 159:2757–2764

Fuchs, M., Schmitt-Keichinger, C., and Sanfaçon, H. 2017. A renaissance in nepovirus research provides new insights into their molecular interface with hosts and vectors. Pages 61–105 in: Advances in Virus Research, M. Kielian, T.C. Mettenleiter, and M.J. Roossinck, eds. Academic Press, Cambridge, MA, USA.

Fujita, N., Komatsu, K., Ayukawa, Y., Matsuo, Y., Hashimoto, M., Netsu, O., Teraoka, T., Yamaji, Y., Namba, S., and Arie, T. 2018. N-terminal region of cysteine-rich protein (CRP) in carlaviruses is involved in the determination of symptom types. Mol. Plant Pathol. 19:180–190

García, J. A., and Pallás, V. 2015. Viral factors involved in plant pathogenesis. Curr. Opin. Virol. 11:21–30

Ghoshal, B., and Sanfaçon, H. 2014. Temperature-dependent symptom recovery in *Nicotiana benthamiana* plants infected with tomato ringspot virus is associated with reduced translation of viral RNA2 and requires ARGONAUTE 1. Virology 456–457:188–197

Gottula, J. W. 2014. Grapevine fanleaf virus: biology, biotechnology and resistance. Ph.D. Dissertation. Cornell University

Hasiów-Jaroszewska, B., Borodynko, N., Jackowiak, P., Figlerowicz, M., and Pospieszny, H. 2011. Single mutation converts mild pathotype of the *Pepino mosaic virus* into necrotic one. Virus Res. 159:57–61

Hasiów-Jaroszewska, B., Paeleman, A., Ortega-Parra, N., Borodynko, N., Minicka, J., Czerwoniec, A., Thomma, B. P. H. J., and Hanssen, I. M. 2013. Ratio of mutated versus wild-type coat protein sequences in *Pepino mosaic virus* determines the nature and severity of yellowing symptoms on tomato plants. Mol. Plant Pathol. 14:923–933

Heaton, L. A., Lee, T. C., Wei, N., and Morris, T. J. 1991. Point mutations in the turnip crinkle virus capsid protein affect the symptoms expressed by *Nicotiana benthamiana*. Virology 183:143–150

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., and Vandesompele, J. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol. 8:R19

Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G., and Galas, D. J. 1989. A simple method for site-directed mutagenesis using the polymerase chain reaction. Nucleic Acids Res. 17:6545–6551

Hily, J.-M., Demanèche, S., Poulicard, N., Tannières, M., Djennane, S., Beuve, M., Vigne, E., Demangeat, G., Komar, V., Gertz, C., Marmonier, A., Hemmer, C., Vigneron, S., Marais, A., Candresse, T., Simonet, P., and Lemaire, O. 2018. Metagenomic-based impact study of transgenic grapevine rootstock on its associated virome and soil bacteriome. Plant Biotechnol. J. 16:208–220

Huss, B., Walter, B., and Fuchs, M. 1989. Cross-protection between arabis mosaic virus and grapevine fan leaf virus isolates in *Chenopodium quinoa*. Ann. Appl. Biol. 114:45–60

Kagiwada, S., Yamaji, Y., Komatsu, K., Takahashi, S., Mori, T., Hirata, H., Suzuki, M., Ugaki, M., and Namba, S. 2005. A single amino acid residue of RNA-dependent RNA polymerase in the *Potato virus X* genome determines the symptoms in *Nicotiana* plants. Vir. Res. 110:117-182

Karran, R. A., and Sanfaçon, H. 2014. Tomato ringspot virus coat protein binds to ARGONAUTE 1 and suppresses the translation repression of a reporter gene. Mol. Plant-Microbe Interact. 27:1–32

Komatsu, K., Hashimoto, M., Maejima, K., Shiraishi, T., Neriya, Y., Miura, C., Minato, N., Okano, Y., Sugawara, K., Yamaji, Y., and Namba, S. 2011. A necrosis-inducing elicitor domain encoded by both symptomatic and asymptomatic *Plantago asiatica mosaic virus* isolates, whose expression is modulated by virus replication. Mol. Plant-Microbe Interact. 24:408–420

Lamprecht, R. L., Spaltman, M., Stephan, D., Wetzel, T., and Burger, J. T. 2013. Complete nucleotide sequence of a South African isolate of Grapevine fanleaf virus. Viruses. 5:1815–1823

Leisner, S. M., and Schoelz, J. E. 2018. Joining the Crowd: Integrating Plant Virus Proteins into the Larger World of Pathogen Effectors. Annu. Rev. Phytopathol. 56:89–110

Li, X., Romero, P., Rani, M., Dunker, A. K., and Obradovic, Z. 1999. Predicting protein disorder for N-, C-, and internal regions. Genome Inform. Ser. Workshop Genome Inform. 10:30–40

Liu, L., Peng, B., Zhang, Z., Wu, Y., Miras, M., Aranda, M. A., and Gu, Q. 2017. Exploring different mutations at a single amino acid position of *Cucumber green mottle mosaic virus* replicase to attain stable symptom attenuation. Phytopathology 107:1080–1086

Liu, Y., Schiff, M., and Dinesh-Kumar, S. P. 2002a. Virus-induced gene silencing in tomato. Plant J. 31:777–786

Liu, Y., Schiff, M., Marathe, R., and Dinesh-Kumar, S. P. 2002b. Tobacco Rar1, EDS1 and NPR1/NIM1 like genes are required for N-mediated resistance to tobacco mosaic virus. Plant J. 30:415–429

Mansilla, C., Sanchez, F., Padget, H. S., Pogue, G. P., and Ponz, F. 2009. Chimeras between *Oilseed rape mosaic virus* and *Tobacco mosaic virus* highlight the relevant role of the tobamoviral *RdRp* as pathogenicity determinant in several hosts. Mol. Plant Pathol. 10:59–68

Martin, I. R., Vigne, E., Berthold, F., Komar, V., Lemaire, O., Fuchs, M., and Schmitt-Keichinger, C. 2018. The 50 distal amino acids of the $2A^{HP}$ homing protein of *Grapevine fanleaf virus* elicit a hypersensitive reaction on *Nicotiana occidentalis*. Mol. Plant Pathol. 19:731–743

Mekuria, T. A., Gutha, L. R., Martin, R. R., and Naidu, R. A. 2009. Genome diversity and intra- and interspecies recombination events in *Grapevine fanleaf virus*. Phytopathology 99:1394–1402

Oliver, J. E., Vigne, E., and Fuchs, M. 2010. Genetic structure and molecular variability of *Grapevine fanleaf virus* populations. Virus Res. 152:30–40

Osterbaan, L. J., Schmitt-Keichinger, C., Vigne, E., and Fuchs, M. 2018. Optimal systemic grapevine fanleaf virus infection in *Nicotiana benthamiana* following agroinoculation. J. Virol. Methods. 257:16–21

Ozeki, J., Takahashi, S., Komatsu, K., Kagiwada, S., Yamashita, K., Mori, T., Hirata, H., Yamaji, Y., Ugaki, M., and Namba, S. 2006. A single amino acid in the RNA-dependent RNA polymerase of Plantago asiatica mosaic virus contributes to systemic necrosis. Arch. Virol. 151:2067–2075

Padmanabhan, M. S., Goregaoker, S. P., Golem, S., Shiferaw, H., and Culver, J. N. 2005. Interaction of the Tobacco mosaic virus replicase protein with the Aux/IAA protein PAP1/IAA26 is associated with disease development. J. Virol. 79:2549–2558

Paudel, D. B., Ghoshal, B., Jossey, S., Ludman, M., Fatyol, K., and Sanfaçon, H. 2018. Expression and antiviral function of ARGONAUTE 2 in *Nicotiana benthamiana* plants infected with two isolates of tomato ringspot virus with varying degrees of virulence. Virology 524:127–139

Pinck, L., Fuchs, M., Pinck, M., Ravelonandro, M., and Walter, B. 1988. A satellite RNA in grapevine fanleaf virus strain F13. J. Gen. Virol. 69:233–239

Qiu, Y., Zhang, Y., Wang, C., Lei, R., Wu, Y., Li, X., and Zhu, S. 2018. *Cucumber mosaic virus* coat protein induces the development of chlorotic symptoms through interacting with the chloroplast ferredoxin I protein. Sci. Rep. 8:1205

Rao, A. L. N., and Grantham, G. L. 1995. A spontaneous mutation in the movement protein gene of brome mosaic virus modulates symptom phenotype in *Nicotiana benthamiana*. J. Virol. 69:2689–2691

Ritzenthaler, C., Laporte, C., Gaire, F., Dunoyer, P., Schmitt, C., Duval, S., Piéquet, A., Loudes, A.-M., Rohfritsch, O., Stussi-Garaud, C., and Pfeiffer, P. 2002. Grapevine fanleaf virus replication occurs on endoplasmic reticulum-derived membranes. J. Virol. 76:8808–8819

Ritzenthaler, C., Viry, M., Pinck, M., Margis, R., Fuchs, M., and Pinck, L. 1991. Complete nucleotide sequence and genetic organization of grapevine fanleaf nepovirus RNA1. J. Gen. Virol. 72:2357–2365

Rodríguez-Cerezo, E., Klein, P. G., and Shaw, J. G. 1991. A determinant of disease symptom severity is located in the 3’-terminal noncoding region of the RNA of a plant virus. Proc. Natl. Acad. Sci. 88:9863–9867

Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J., and Dunker, A. K. 2001. Sequence complexity of disordered protein. Proteins Struct. Funct. Bioinforma. 42:38–48

Schmitt-Keichinger, C., Hemmer, C., Berthold, F., and Ritzenthaler, C. 2017. Molecular, Cellular, and Structural Biology of *Grapevine fanleaf virus*. Pages 83–107 in: Grapevine Viruses: Molecular Biology, Diagnostics and Management, B. Meng, G.P. Martelli, D.A. Golino, and M. Fuchs, eds. Springer International Publishing, Cham.

Shiboleth, Y. M., Haronsky, E., Leibman, D., Arazi, T., Wassenegger, M., Whitham, S. A., Gaba, V., and Gal-On, A. 2007. The conserved FRNK box in HC-Pro, a plant viral suppressor of gene silencing, is required for small RNA binding and mediates symptom development. J. Virol. 81:13135–13148

Shimura, H., Pantaleo, V., Ishihara, T., Myojo, N., Inaba, J., Sueda, K., Burgyán, J., and Masuta, C. 2011. A viral satellite RNA induces yellow symptoms on tobacco by targeting a gene involved in chlorophyll biosynthesis using the RNA silencing machinery. PLoS Pathog. 7:1–12

Smith, N. A., Eamens, A. L., and Wang, M.-B. 2011. Viral small interfering RNAs target host genes to mediate disease symptoms in plants. PLoS Pathog. 7:e1002022

Valli, A. A., Gallo, A., Rodamilans, B., López-Moya, J. J., and García, J. A. 2018. The HCPro from the *Potyviridae* family: an enviable multitasking Helper Component that every virus would like to have. Mol. Plant Pathol. 19:744–763

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 3:research0034.1

Vigne, E., Gottula, J. W., Schmitt-Keichinger, C., Komar, V., Ackerer, L., Belval, L., Rakotomalala, L., Lemaire, O., Ritzenthaler, C., and Fuchs, M. 2013. A strain-specific segment of the RNA-dependent RNA polymerase of grapevine fanleaf virus determines symptoms in *Nicotiana* species. J. Gen. Virol. 94:2803–2813

Viry, M., Serghini, M. A., Hans, F., Ritzenthaler, C., Pinck, M., and Pinck, L. 1993. Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. J. Gen. Virol. 74:169–174

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. 2014. The I-TASSER suite: protein structure and function prediction. Nat. Methods. 12:7

Zhang, C., Wu, Z., Li, Y., and Wu, J. 2015. Biogenesis, function, and applications of virus-derived small RNAs in plants. Front. Microbiol. 6:1237

CHAPTER 5

PROTEIN INTERACTION PARTNERS OF GRAPEVINE FANLEAF VIRUS RNA-DEPENDENT RNA POLYMERASE DURING INFECTION OF *NICOTIANA BENTHAMIANA* DETERMINED BY AFFINITY PURIFICATION AND TANDEM MASS SPECTROMETRY

INTRODUCTION

Grapevine fanleaf virus (GFLV) is one of the most adverse grapevine-infecting viruses worldwide. This epithet is warranted from its worldwide distribution, especially in some of the premiere viticulture regions, and the devastating effects it has on vine productivity. The disease which it causes, fanleaf degeneration, is marked by symptoms including yellow mosaic, vein banding, foliar deformations (including the namesake "fanleaf", the result of abnormal vein geometry and decreased lobing leading to fan-shaped leaves), shortened internodes, poor and uneven berry set, and an overall decline in the vigor of the vine which shortens its productive lifespan (Digiaro et al. 2017). While some efforts have been made to identify genetic elements of GFLV associated with particular syndromes (Elbeaino et al. 2014), the molecular mechanisms underpinning GFLV symptom development remain unknown.

GFLV is a subgroup A nepovirus (a genus of the family *Secoviridae* in the order *Picornavirales*). Its two positive-sense single-stranded genomic RNAs are separately encapsidated in isometric particles of pseudo T=3 symmetry assembled from 60 copies of the coat protein ($2C^{CP}$). Expression of the genome is by monocistronic translation followed by proteolytic processing by the viral protease ($1D^{Pro}$). RNA1 encodes five proteins for genome replication and polyprotein

maturation. A definitive function for protein 1A has not been identified. Protein $1B^{Hel}$ is a putative helicase (Candresse et al. 1990). The $1C^{VPg}$ protein is a genome-linked viral protein (VPg) which becomes covalently attached to the 5' end of the genomic RNAs, serving as an analog of the 5' cap of mRNAs and possibly participating in RNA priming during genome replication (Pinck et al. 1991; Steil and Barton 2009). $1D^{Pro}$ is the aforementioned cysteine protease which cleaves viral polyproteins into individual maturation product (Margis et al. 1991). $1E^{Pol}$ contains the seven conserved motifs of a plant virus RNA-dependent RNA polymerase (Chisholm et al. 2007; Vigne et al. 2013). RNA2 encodes proteins for local and systemic movement of the virus. This includes $2A^{HP}$, a putative homing protein that appears to play a role in localizing RNA2 to the site of viral replication (Gaire et al. 1999). Protein $2B^{MP}$ is a tubule-forming movement protein which aids the movement of viral progeny into adjacent cells via plasmodesmata (Laporte et al. 2003). Protein $2C^{CP}$ is the aforementioned coat protein. Motifs on $2C^{CP}$ determine the specificity between GFLV and its vector, the ectoparasitic dagger nematode *Xiphinema index* (Belin et al. 2001; Andret-Link et al. 2004; Schellenberger et al. 2010).

While the molecular mechanisms underpinning GFLV symptom development in grapevine are unknown, recent advances have been made in understanding symptom development in herbaceous hosts. GFLV can be mechanically passaged to herbaceous hosts such as *Nicotiana occidentalis* and *N. benthamiana.* In *N. occidentalis,* GFLV strain F13 produces necrotic lesions on inoculated leaves. It was recently shown that this is a genuine hypersensitive response (HR), triggered by the 50-distal amino acids of the $2A^{HP}$ protein of GFLV-F13 (Martin et al. 2018).

In contrast, GFLV-F13 produces an asymptomatic infection in *N. benthamiana*, while GFLV strain GHu produces distinct vein clearing on upper, uninoculated leaves of this species (Vigne et al. 2013). The vein clearing symptom determinant of GFLV-GHu was mapped to residue 802 of the $1E^{Pol}$ protein. This residue, which is a lysine in GFLV-GHu, was found to be necessary but not sufficient for vein clearing. Furthermore, this residue is capable of modulating symptom severity as certain amino acid substitutions (proline, asparagine) produced faint vein clearing which faded more rapidly than those produced by wild type GFLV-GHu (Osterbaan et al. 2019a).

While much work has been done to describe the molecular events of GFLV infection (for review, see Schmitt-Keichinger et al. 2017), little is known about the molecular context of $1E^{Pol}$ during infection, leaving us with few clues as to the mechanistic role GFLV-GHu $1E^{Pol}$ residue 802 plays in symptom development. Thus we sought to create a preliminary map of the $1E^{Pol}$ protein interactome via affinity purification coupled to tandem mass spectrometry (AP-MS). By adding an epitope tag (V5) derived from simian virus 5 (Southern et al. 1991) to the C-terminus of $1E^{Pol}$ in infectious clones of GFLV-GHu and –F13 (Osterbaan et al. 2018), as well as an asymptomatic GFLV-GHu mutant (GHu-1E-K802G, Osterbaan et al. 2019a), we were able to affinity purify tagged $1E^{Pol}$ using a commercially-available anti-V5 antibody from total protein extracts of systemically infected *N. benthamiana.* Tryptic peptides of AP products were analyzed by tandem mass spectrometry (MS/MS), allowing for the identification of viral and host proteins which were co-purified with $1E^{Pol}$, giving

us the first snap shot of the protein interaction network of $1E^{Pol}$ during systemic GFLV infection.

MATERIALS AND METHODS

Creating V5 tagged GFLV cDNA constructs

GFLV RNA1 mutants were generated using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) according to the manufacturer's instructions, using specific primers (designed with the NEBaseChanger software, v1.2.8, www.nebasechanger.neb.com/) and plasmids pCLEAN-F131-35S and pCLEAN-GHu1-35S (Osterbaan et al. 2018) and pCLEAN-GHu-1E-K802G (Osterbaan et al. 2019a) as templates (Table 5-1). Mutated plasmids were transformed into *E. coli* DH5α cells (New England Biolabs) and isolated from single colonies using the E.Z.N.A. Plasmid Mini Kit II (Omega Bio-tek). Sanger sequencing of single colony plasmids were conducted at Cornell Biotechnology Resource Center using the primers GHuSymdeterSeq-F (5'-TGCTAGGACCAAATCAGAGGAACG-3') and GHuSymdeterSeq-R (5'-AAACTTGGTTATCCCAGTACCA-3').

Table 5-1. Mutagenic primers used to insert epitope tag sequences into the coding region of protein $1E^{Pol}$ in GFLV RNA1 cDNA constructs.

Table 5-1

Construct	Template	Forward primer name	Forward primer sequence (5'-3')[a]	Reverse primer name	Reverse primer sequence (5'-3')[a]
pCLEAN-GHu-1E:3XFLAG	pCLEAN-GHu-1	GHu-1E-3XFLAG-F	catgacatcgattacaaggatgacgatgacaagTAATTCTTCCAACCCTTGGTAC	GHu-1E-3XFLAG-R	atctttataatcaccgtcatggtctttgtagtcCTTCCTCGGGCATGAGTG
pCLEAN-GHu-1E:FLAG	pCLEAN-GHu-1	GHu-1E-FLAG-F	gatgatgataaaTAATTCTTCCAACCCTTGGTAC	GHu-1E-FLAG-R	atctttataatcCTTCCTCGGGCATGAGTG
pCLEAN-GHu-1E:HA	pCLEAN-GHu-1	GHu-1E-HA-F	gccggattatgcgTAATTCTTCCAACCCTTGGTAC	GHu-1E-HA-R	acatcatacggataCTTCCTCGGGCATGAGTG
pCLEAN-GHu-1E:myc	pCLEAN-GHu-1	GHu-1E-myc-F	agcgaagaagatctgTAATTCTTCCAACCCTTGGTAC	GHu-1E-myc-R	aatcagtttctgttcCTTCCTCGGGCATGAGTG
pCLEAN-GHu-1E:V5	pCLEAN-GHu-1	GHu-1E-V5-F	ctgctgggcctggatagcaccTAATTCTTCCAACCCTTGGTAC	GHu-1E-V5-R	cgggttcggaatcggtttgccCTTCCTCGGGCATGAGTG
pCLEAN-GHu-K802G-1E:V5	pCLEAN-GHu-1E-K802G	GHu-1E-V5-F	ctgctgggcctggatagcaccTAATTCTTCCAACCCTTGGTAC	GHu-1E-V5-R	cgggttcggaatcggtttgccCTTCCTCGGGCATGAGTG
pCLEAN-F13-1E:3XFLAG	pCLEAN-F13-1	F13-1E-3XFLAG-F	catgacatcgattacaaggatgacgatgacaagTAAGCCTTCCAATTCTTG	F13-1E-3XFLAG-R	atctttataatcaccgtcatggtctttgtagtcTTTCCTAAGGCATGTATG

Table 5-1, cont.

Construct	Template	Forward primer name	Forward primer sequence (5'-3')[a]	Reverse primer name	Reverse primer sequence (5'-3')[a]
pCLEAN-F13-1E:FLAG	pCLEAN-F13-1	F13-1E-FLAG-F	gatgatgataaaTAAGCCTTCCAATTCTTG	F13-1E-FLAG-R	atctttataatcTTTCCTAAGGCATGTATG
pCLEAN-F13-1E:HA	pCLEAN-F13-1	F13-1E-HA-F	gccggattatgcgTAAGCCTTCCAATTCTTG	F13-1E-HA-R	acatcatacggataTTTCCTAAGGCATGTATG
pCLEAN-F13-1E:myc	pCLEAN-F13-1	F13-1E-myc-F	agcgaagaagatctgTAAGCCTTCCAATTCTTG	F13-1E-myc-R	aatcagtttctgttcTTTCCTAAGGCATGTATG
pCLEAN-F13-1E:V5	pCLEAN-F13-1	F13-1E-V5-F	ctgctgggcctggatagcaccTAAGCCTTCCAATTCTTG	F13-1E-V5-R	cgggttcggaatcggtttgccTTTCCTAAGGCATGTATG
pCLEAN-F13-1E:$V5_{C-1}$	pCLEAN-F13-1	F13-1E-V5-F	ctgctgggcctggatagcaccTAAGCCTTCCAATTCTTG	F13-1E-V5:C-1_Rev	cgggttcggaatcggtttgccCCTAAGGCATGTATGATTC

[a]Lower case type indicates mutagenic (insertion) sequences. Upper case type indicates residues which anneal to the original template.

Infection of N. benthamiana *with GFLV variants*

Agrobacterium tumefaciens-mediated inoculations (agroinoculations) were performed as detailed in Osterbaan et al. (2018). Systemic GFLV infection *in planta* requires both RNA1 and RNA2 (Viry et al. 1993), and GFLV-GHu RNA2 was used as the RNA2 construct for all agroinoculations. Each construct was used in at least two independent agroinoculation trials of at least five *N. benthamiana* plants per construct per trial.

Mechanical inoculation of *N. benthamiana* was conducted using archived systemically infected tissue (i.e. whole leaves collected 3 to 4 weeks post agroinoculation and stored at -80°C). Briefly, archived tissue was ground 1:5 in mechanical inoculation buffer (15 mM KH_2PO_4, 35 mM Na_2HPO_4, pH 7.4). Leaves of *N. benthamiana* plants at the 8 to 12 leaf stage were dusted with Carborundum powder, rubbed with a pestle dipped in crude sap, and rinsed with distilled water. Plants were monitored for symptom development and tested for GFLV infection at 1 and 2 weeks post inoculation by double antibody sandwich (DAS) enzyme-linked immunosorbent assay (ELISA) using specific GFLV antibodies (BIOREBA, Reinach, Switzerland).

Time course monitoring of GFLV RNA1 relative titer by reverse transcriptase quantitative polymerase chain reaction and 1E:V5 titer by western blotting

Sixteen plants of *N. benthamiana* were mechanically inoculated with GHu-1E:V5; in parallel, a set of sixteen plants were mock inoculated. Punches from apical leaves were taken each day from two plants starting at three days post inoculation through thirteen days post inoculation and flash frozen on liquid nitrogen. Separate

punches were taken for RNA extraction and western blotting. Frozen leaf punches were stored at -80°C until processing. RNA extraction and reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) of GFLV RNA1 were carried out according to Osterbaan et al. (2019). At 13 days post inoculation, all plants were tested for GFLV infection by DAS-ELISA using GFLV specific antibodies (BIOREBA). All plants inoculated with GHu-1E:V5 tested positive for GFLV and GFLV was not detected in any of the mock-inoculated plants.

Leaf punches for protein extraction were ground on a Retsch MM400 ball mill for 2 minutes at 30 Hz with two steel BBs. Total proteins from leaf punches were extracted using a denaturing lysis buffer (50 mM Tris-HCl, pH 6.8, 10% glycerol, 2.5% SDS, 5% β-mercaptoethanol) and the small-scale protein extraction protocol described below. SDS-PAGE and western blotting were performed as described below.

Small scale protein extractions for optimization of GFLV 1E^{Pol} extraction conditions

Plants of *N. benthamiana* plants were mechanically inoculated with either GHu-1E:V5, F13-1E:V5$_{C\text{-}1}$, GHu-1E-K802G:V5, or untagged GFLV-GHu, -F13, or –GHu-1E-K802G. Additional plants were mock-inoculated and used as controls. Plants were tested for systemic GFLV infection via DAS-ELISA at 6 days post inoculation and apical leaves of GFLV-positive plants were collected at 7 days post inoculation, flash frozen in liquid nitrogen, and stored at -80°C.

Within one week of collection, leaves were cryomilled. Briefly, whole frozen leaves were roughly macerated in chilled 50 mL conical tubes using a chilled metal

scoopula. Roughly macerated tissue was cryomilled for 3 sets of 3 minutes at 30 Hz (with 5 minutes of chilling in liquid nitrogen between grinds) on a Retsch MM400 ball mill using stainless steel 10 mL grinding jars and balls that had been pre-chilled in liquid nitrogen for 30 minutes. The resulting tissue powder was stored at -80°C until extractions.

For small-scale extractions, small aliquots of tissue powder were weighed and transferred to 2 mL microfuge tubes. To each tube, 5 volumes of extraction buffer (variable, see Table 5-2 and figure captions) were added. Samples were incubated on ice for 10 minutes with occasional vortexing. Homogenates were clarified by centrifugation at 10,000 x g for 10 minutes at 4°C. The supernatants were transferred to clean microfuge tubes and immediately analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and western blotting or were stored at -80°C until analysis.

Table 5-2. Formulations for buffers used in trial protein extractions of GHu-1E:V5 from systemically infected *N. benthamiana* tissue

Trial Buffer No.	Formulation[a]	Original application	Reference
1	50 mM Tris-HCl (pH 8.0); 100 mM NaCl; 10 mM DTT; 0.1% Tween-20	Extraction of total soluble proteins from transgenic *N. benthamiana* expressing a GFLV-specific antibody	Nölke et al. (2009)
2	100 mM Tris-HCl (pH 8.1); 10 mM KCl, 5 mM $MgCl_2$; 400 mM sucrose; 10% glycerol; 1 mM β-ME	Extraction of ToRSV CP from *N. benthamiana* tissue	Jovel et al. (2007)
3	50 mM Tris-HCl (pH 7.6); 150 mM NaCl; 10% glycerol; 1 mM DTT; 0.1% IGEPAL	AP of TuMV RdRp from *Arabidopsis thaliana*	Dufresne et al. (2008)
4	50 mM HEPES-KOH (pH 7.4); 110 mM KOAc; 2 mM $MgCl_2$; 0.4% TritonX-100	AP of PLRV RTP from agroinfiltrated *N. benthamiana* tissue	DeBlasio et al. (2014)

[a]All buffers also included 1X Halt protease inhibitor cocktail (Thermo Fisher Scientific)

Antibody preparation for affinity purification

In preparation for AP assays, 2 mg of polyclonal anti-V5 antibody (Invitrogen PA1-993) were cross-absorbed by sodium sulfate precipitation against total soluble proteins extracted from a blend of *N. benthamiana* tissue infected with untagged GFLV-GHu, -F13, and –GHu-1E-K802G. The antibody fraction was recovered by overnight dialysis with 0.1 M Na_2PO_4 (pH 8.0) in a 10K molecular-weight cut-off Slide-A-Lyzer cassette (Thermo Fisher Scientific).

Affinity purification of GFLV 1E^{Pol}-host complexes

AP of GFLV 1E^{Pol}:V5-host complexes was carried out according to DeBlasio et al. (2014) with the following modifications. Polyclonal anti-V5 was bound to Dynabeads Protein A (Life Technologies, Invitrogen) at a concentration of 10 µg antibody per 1 mg of beads. For extraction of proteins from cell lysates, the lysis buffer was amended with 2.5 mM dithiothreitol (DTT) and phenyl methylsulfonyl fluoride (PMSF) was not included. Lysis was allowed to proceed for 15 minutes. Lysate was diluted 1:5 in fresh lysis buffer which had not been amended with DTT and beads were added at 1 mg/mL. Washes of beads were performed with DTT-free lysis buffer and the full bead volume was subjected to on-bead trypsin digestion. AP was performed for GHu-1E:V5, F13-1E:V5$_{C\text{-}1}$, and GHu-1E-K802G:V5, as well as negative controls of untagged GFLV-GHu and –F13. Three biological replicates were performed for each treatment. Trial APs to test the specificity and ability of anti-V5 to purify the V5-tagged 1E^{Pol} were run as above, but were scaled down for 0.6 mg of beads.

SDS-PAGE and western analysis of V5-tagged GFLV 1E^{Pol}

For analysis of small-scale protein extraction products, supernatants were mixed 1:1 with 2X Laemmli sample buffer (Bio-rad; 65.8 mM Tris-HCl, pH 6.8, 26.3% (w/v) glycerol, 2.1% SDS, 0.01% bromophenol blue) amended with 5% β-mercaptoethanol (β-ME) and incubated at 70°C for 10 minutes. For analysis of trial AP products, proteins were eluted by resuspending beads in 50 µL of 2X Laemmli sample buffer amended with 5% β-ME, incubating at 70°C for 10 minutes, then retrieving the supernatant. Aliquots of 20 µL (protein extractions) or 40 µL (trial APs) of prepared samples were separated on 10% or 4-20% gradient Mini-PROTEAN TGX precast gels (Bio-rad) for 60 minutes at 50 V followed by 180 minutes at 100 V at room temperature. Proteins were transferred to 0.45 µm nitrocellulose membranes (Bio-rad) via wet transfer (20% methanol) for 90 minutes at 80 V at 4°C. Sample loading was checked by Ponceau staining of transferred proteins. Membranes were blocked in 5% milk for 1 hour at room temperature. Blotted proteins were probed with 1:1000 rabbit anti-V5 polyclonal antibody (Invitrogen PA1-993) overnight at 4°C, followed by 1:5000 goat anti-rabbit AP-conjugated IgG (Invitrogen 31340) for 1 hour at room temperature. Between steps, membranes were washed three times in phosphate-buffered saline with 0.5% Tween-20 for 10 minutes. Proteins bands were visualized with 1-Step NBT/BCIP substrate solution (Thermo Scientific).

Sample preparation for mass spectrometry

Disulfide bonds of proteins bound to Dynabeads were reduced using 50 µL of 0.2% sodium 3-(4-(1,1-bis(hexyloxy)ethyl)pyridinium-1-yl)propane-1-sulfonate (PPS surfactant from Expedeon, San Diego, CA) in 100 mM ammonium bicarbonate with 1

µL 500 mM Tris(2-carboxyethyl)phosphine hydrochloride (Thermo-Fisher Scientific) for one hour at 60°C. The sample was cooled to room temperature and the cysteine thiols alkylated by the addition of 1.1 µL of 500 mM iodoacetamide for 20 minutes in the dark. The beads were then digested with the addition of 1 µg of modified sequencing-grade trypsin (Promega, catalog number V5111) overnight at 37°C with constant agitation. The supernatant containing the tryptic peptides were removed, and combined with a methanol wash (50 µL) of the beads. The combined supernatants were dried using vacuum centrifugation.

Prior to liquid chromatography-mass spectrometry (LC-MS) analysis, samples were reconstituted in 10 µL 0.2% formic acid (FA) at room temperature and left to sit at room temperature for one hour. The samples were then diluted in an additional 90 uL of 0.2% FA and adjusted to 1% trifluoracetic acid. The following clean-up steps were applied to separate aliquots of each sample. In the first, 90 µL aliquots were treated with OMIX C18 100 µL tips (Agilent, catalog number A57003100) and the peptides were eluted in 100 µL 0.1% FA. In the second, 10 µL aliquots were treated with Millipore C18 10µL ZipTips (Millipore Sigma, catalog number ZTC18S960). Tips were wetted in 50% acetonitrile (ACN) twice, equilibrated in 0.1% trifluoroacetic acid (TFA) twice, the sample was aspirated and dispensed for ten cycles, rinsed in 0.1% TFA twice and the peptides were eluted in 10 uL 0.1% FA.

Mass spectrometry workflow

An Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific), equipped with a nano-ion spray source coupled to an EASY-nLC 1200 system (Thermo Scientific) was used at Cold Spring Harbor Laboratory as an expansion of an ongoing

collaboration with Michelle Heck's laboratory. The LC system was configured with a self-pack PicoFrit™ 75-µm analytical column with an 8-µm emitter (New Objective, Woburn, MA) packed to 25cm with ReproSil-Pur C18-AQ, 1.9 uM material (Dr. Maish HPLC, GmbH). Mobile phase A consisted of 2% acetonitrile and 0.1% FA and mobile phase B consisted of 90% acetonitrile and 0.1% FA. Peptides were then separated using the following steps: at a flow rate of 200 nL/min: 2% B to 6% B over 1 min, 6% B to 30% B over 84 min, 30% B to 60% B over 9 min, 60% B to 90% B over 1 min, held at 90% B for 5 min, 90% B to 50% B over 1 min and then the flow rate was increased to 500 nL/min as 50% B was held for 9 min.

Eluted peptides were directly electrosprayed into the Orbitrap Fusion Lumos mass spectrometer with the application of a distal 2.3 kV spray voltage and a capillary temperature of 300°C. Full-scan mass spectrum (Res=60,000; 400-1600 *m/z*) were followed by MS/MS using the "Top Speed" method for selection. High-energy collisional dissociation (HCD) was used with the normalized collision energy set to 35 for fragmentation, the isolation width set to 1.2 and a duration of 10 seconds was set for the dynamic exclusion with an exclusion mass width of 10ppm. We used monoisotopic precursor selection for charge states 2+ and greater, and all data were acquired in profile mode.

Analysis of mass spectrometry data: database searching and label-free quantification

Mass spectrometry raw files were converted to Mascot generic files (.mgf) using ProteoWizard (Kessner et al. 2008) and searched using Mascot, version 2.5.1 (Perkins et al. 1999). The following modifications were considered: carbamidomethyl

of cysteines (fixed), deamidation of asparagine and glutamine (variable), and oxidation of methionine (variable). Mass accuracy was set to 20 ppm for precursor ions and 0.5 Da for fragment ions. One missed tryptic cleavage was allowed. The database used for protein identifications was the same as used in DeBlasio et al. (2015), except that *Luteoviridae* proteins were substituted with an in-house curated list of nepovirus proteins (Table 5-3). The list of nepovirus proteins was chosen from among nepovirus isolates for which sequence information is available in the NCBI GenBank for the complete genomic RNA1 and/or RNA2 (thus containing full coverage of the P1 and/or P2 polyproteins). Wherever possible, delineation of P1 and P2 sequences into individual mature peptides followed annotations provided in GenBank entries. Where this information was not available (i.e. some GFLV isolates and the arabis mosaic virus, ArMV, isolate Lv), mature peptides were selected by sequence alignment to GFLV-F13 to locate the conserved $1D^{Pro}$ cleavage recognition sites. Mature peptides for P2 of grapevine deformation virus (GDeV) and tomato ringspot virus (ToRSV) were not included due to lack of annotated cleavage sites; these isolates are sufficiently divergent from GFLV-F13 such that sequence alignment could not confidently identify conserved cleavage sites. Search results were exported as Mascot data files (.dat) and imported into Scaffold 4.8.9 for label-free quantification by spectral counting. The protein-clustering feature was used to streamline analysis and spectra were filtered with a peptide and protein false discovery rate (FDR) of 1%. Analysis of variance (ANOVA) of spectral counts (normalized to total spectra) between treatments was performed in Scaffold. Technical replicates from untagged GFLV-GHu and –F13 were pooled into a single negative control treatment.

Table 5-3. List of nepovirus protein sequences included in the protein database used for Mascot database searching of mass spectra.

		GenBank accession numbers[a]	
Species[b]	Isolate	P1	P2[c]
GFLV	1050-02	AFV34731.1	AFV3472.1
GFLV	3138-01	AFV34755.1	AFV34756.1
GFLV	F13	NP_619689.1	NP_619706.1
GFLV	GHu	AFM91094.1	ABR09917.1
GFLV	GHu-K802G	Manual, $1E^{Pol}$ only[d]	N/A
GFLV	SACH44	AGT42200.1	AGT42201.1
GFLV	SAPCS3	AFI71813.1	AFI71814.1
GFLV	SDHN	APD26306.1	APD26307.1
GFLV	SWT2-1	ARO69856.1	ARO69918.1
GFLV	WAPN173	ACZ58632.1	ACZ58628.1
GFLV	WAPN6132	ACZ58633.1	ACZ58626.1
GFLV	WTR4-2	ARO69845.1	ARO69913.1
ArMV	Lv	ACF32434.1	ACF32435.1
GDeV	N66	YP_006347591.1	N/A
ToRSV	raspberry	NP_734003.1	N/A

[a]Mature peptide sequences were curated by following annotations within the GenBank entry where available. Otherwise, polyprotein sequences were aligned to GFLV-F13 sequences to locate the conserved $1D^{Pro}$ cleavage sites.
[b]GFLV: grapevine fanleaf virus; ArMV: arabis mosaic virus; GDeV: grapevine deformation virus; ToRSV: tomato ringspot virus.
[c]N/A indicates that sequences for polyprotein P2 were not included either due to redundancy (GHu-K802G) or due to the lack of mature peptide annotations (GDeV and ToRSV).
[d]Only the protein sequence for $1E^{Pol}$ for this isolate was included in the database since this isolate is otherwise identical to GFLV-GHu. The $1E^{Pol}$ sequence was curated manually by changing residue 802 (lysine, K) to glycine (G).

RESULTS

Selection of epitope tag for tagging of GFLV protein $1E^{Pol}$

We sought to tag the $1E^{Pol}$ protein of GFLV strains GHu and F13 by the insertion of sequences coding for several common epitope tags between the C-terminal codon and stop codon of the $1E^{Pol}$ coding region in GFLV cDNA constructs using a Q5 mutagenesis kit (New England Biolabs) and specific primers (Table 5-1) designed using NEBaseChanger software. When strains F13 and GHu were tagged with

epitopes 3XFLAG (DYKDHDGDYKDHDIDYKDDDDK), FLAG (DYKDDDDK), HA (YPYDVPDYA), or myc (EQKLISEEDL), the resulting constructs were unable to establish systemic infection in *N. benthamiana* following *Agrobacterium tumefaciens*-mediated inoculation, despite several attempts (at least three independent agroinoculations of at least 6 plants per construct). In contrast, tagging of GFLV-GHu $1E^{Pol}$ with epitope V5 (GKPIPNPLLGLDST, derived from the P and V proteins of simian virus 5) produced a construct (GHu-1E:V5) which was systemically infectious in *N. benthamiana* and which produced vein clearing symptoms indistinguishable from wild type GFLV-GHu. Since we were interested in comparing the $1E^{Pol}$ protein interactomes of symptomatic and asymptomatic variants of GFLV, we also added the V5 tag to GHu-1E-K802G. This is an asymptomatic mutant of GFLV-GHu in which the $1E^{Pol}$ residue which modulates vein clearing, lysine 802, is mutated to glycine, the native 802 residue of GFLV-F13 (see Chapter 4). The resulting tagged construct, GHu-1E-K802G:V5 was systemically infectious and asymptomatic in *N. benthamiana*.

Direct C-terminal fusion of the V5 tag to $1E^{Pol}$ of the F13 strain rendered the virus unable to establish systemic infection of *N. benthamiana* via agroinoculation. To rescue the systemic infectivity of the V5-tagged GFLV-F13 clone, we created truncation mutants in which deletions of 1 to 5 amino acids of the C-terminus of the $1E^{Pol}$ protein were introduced in addition to the C-terminal V5 epitope tag. Of these, a construct bearing a single amino acid truncation of the $1E^{Pol}$ protein followed by a C-terminal V5 tag, F13-1E:$V5_{C-1}$, was systemically infectious and asymptomatic in *N. benthamiana*.

Time course monitoring of 1E^{Pol} titer shows that 1E^{Pol} expression peaks between 6 and 9 days post inoculation

Prior to attempting AP experiments, we determined the optimal timing of tissue collection and tissue lysis conditions which would provide maximal 1E^{Pol} yield. To that end, we performed a time course study to monitor 1E^{Pol} titer throughout GFLV infection.

GFLV was first detectable at six days post inoculation in apical leaves by RT-qPCR. GFLV RNA1 titer rose through ten days post inoculation, then dropped slightly (Figure 5-1). In parallel, we performed small-scale protein extractions using a highly denaturing lysis buffer followed by western blotting to determine when 1E^{Pol}:V5 was expressed at detectable levels. We found that 1E^{Pol}:V5 was detectable as early as six days post inoculation and as late as nine days post inoculation (Figure 5-2). Interestingly, both GFLV RNA1 and 1E^{Pol}:V5 titer appear to correlate with symptom expression of GFLV-GHu. Indeed, vein clearing induced by GFLV-GHu is generally first noticeable 4 to 6 days post mechanical inoculation and these symptoms start to fade at 9 to 10 days post inoculation (Figure 5-3).

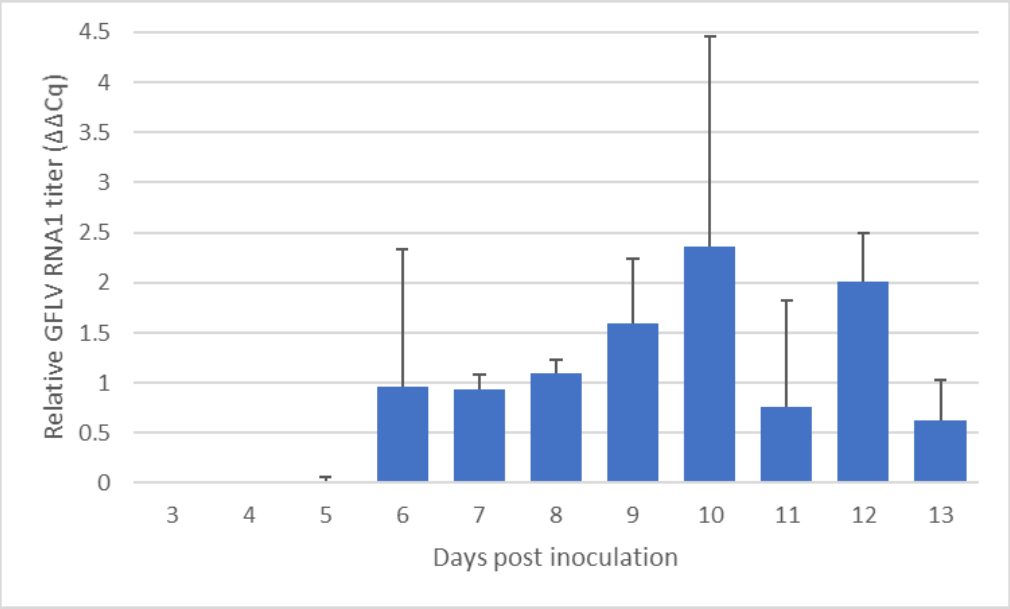


Figure 5-1. Relative titer of grapevine fanleaf virus (GFLV) RNA1 in *Nicotiana benthamiana* as determined by reverse transcriptase quantitative polymerase chain reaction following mechanical inoculation with V5-tagged GFLV-GHu. Bars represent the mean relative titer of two biological samples. Error bars show the standard error of the mean. GFLV RNA1 is first detectable by RT-qPCR at 6 days post inoculation.

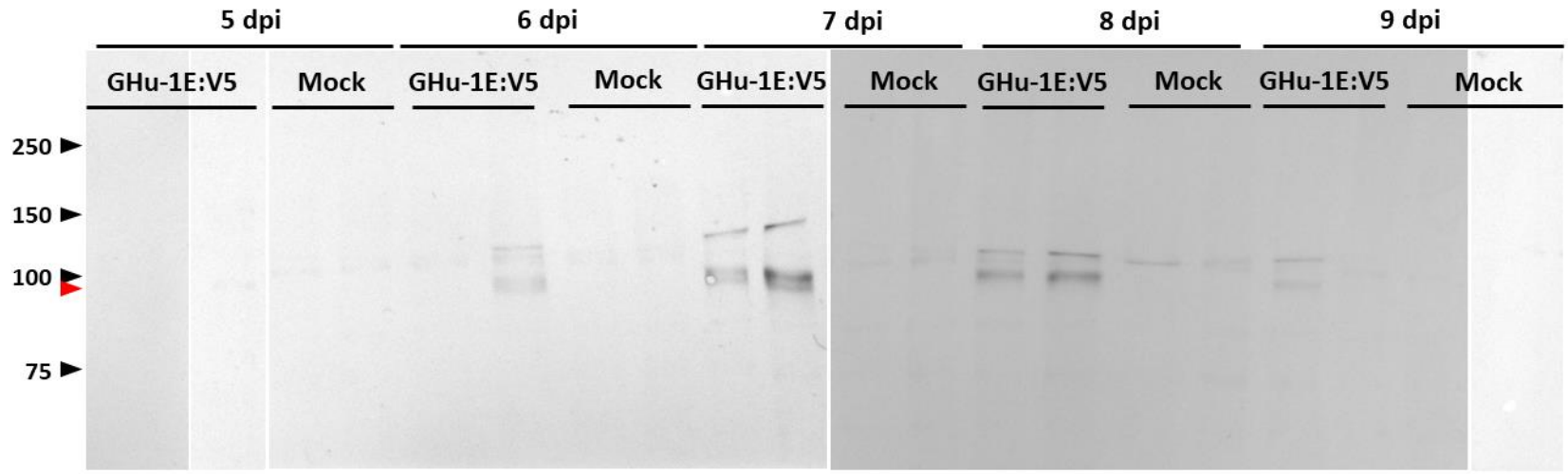
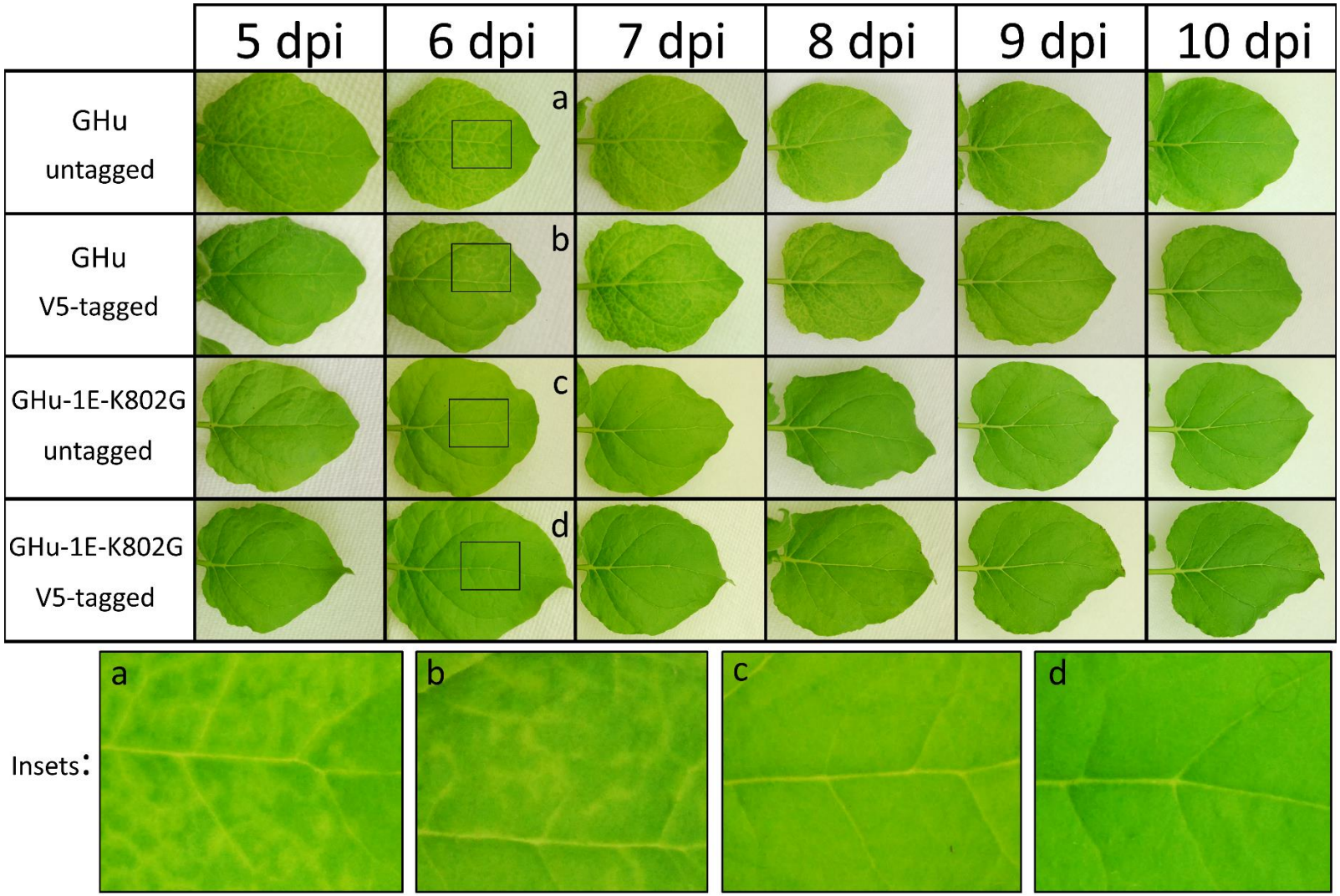


Figure 5-2. Time course western blot analysis of V5-tagged GFLV-GHu $1E^{Pol}$ extracted from systemically infected *N. benthamiana* leaves. Total soluble proteins from *N. benthamiana* mock inoculated or infected with GHu-1E:V5 were extracted in a denaturing lysis buffer (50 mM Tris-HCl, pH 6.8, 10% glycerol, 2.5% SDS, 5% β-mercaptoethanol). Two biological samples were extracted per treatment for each time point. Proteins were probed with polyclonal anti-V5 antibody (Invitrogen PA1-993) and goat anti-rabbit AP-conjugated secondary antibody (Invitrogen T2191) and developed with 1-Step NBT/BCIP solution. The predicted molecular weight of V5-tagged $1E^{Pol}$ is 93 kDA and is indicated with a red arrow head. V5-tagged $1E^{Pol}$ is detectable between 6 and 9 days post mechanical inoculation, with optimal detection at 7 to 8 days post inoculation.

Figure 5-3. Progression of vein clearing symptoms induced by grapevine fanleaf virus (GFLV) strain GHu in *Nicotiana benthamiana* following mechanical inoculation. *N. benthamiana* plants were mechanically inoculated with untagged (first and third rows) or tagged (second and fourth row) versions of wild type GFLV-GHu (first and second rows) or an asymptomatic mutant of GFLV-GHu in which lysine 802 of the $1E^{Pol}$ protein has been mutated to glycine (third and fourth rows). Apical leaves were photographed beginning at 5 days post inoculation (dpi) through 10 dpi. Insets show a close up of vein clearing symptoms (a and b) and asymptomatic leaves (c and d). The contrast for images shown in the insets was increased by 25% to better display the differences between yellow and green areas of the leaf.


5 dpi
6 dpi
7 dpi
8 dpi
9 dpi
10 dpi
GHu untagged
GHu V5-tagged
GHu-1E-K802G untagged
GHu-1E-K802G V5-tagged
a
b
c
d
Insets:
a
b
c
d

Optimization of lysis buffer for extraction of 1E^{Pol} from **N. benthamiana** *tissue*

While the lysis buffer used for the time course study allowed us to reliably detect 1E^{Pol}:V5 via western blotting, this denaturing lysis buffer is not suitable for use in AP experiments as denaturation eliminates protein-protein interactions. To identify a lysis buffer which would be compatible with AP and efficiently extract 1E^{Pol}:V5, we selected several lysis buffers from the literature which had previously been used for the study of proteins of GFLV or related viruses or for AP of plant virus proteins (Table 5-2) and conducted small-scale total protein extractions from *N. benthamiana*.

For our initial optimization, we used the GHu-1E:V5 construct. Tissue from plants infected with wild type GFLV-GHu (untagged) was used as a negative control to confirm the specificity of the polyclonal anti-V5 antibody for the V5-tagged 1E^{Pol} protein. Of the lysis buffers which allowed us to confidently detect a 1E^{Pol}:V5 band via western blotting, the commonality was the inclusion of a reducing agent such as dithiothreitol (DTT) or β-mercaptoethanol (β-ME) (Figure 5-4). Notably, the denaturing lysis buffer used for the time course study also included β-ME. To confirm the necessity of a reducing agent for reliable extraction of 1E^{Pol}:V5 from *N. benthamiana* tissue, we performed additional small-scale extractions in which components of one of the successful buffers were singly excluded. Indeed, exclusion of DTT from the extraction buffer drastically reduced our ability to detect 1E^{Pol}:V5 in total protein extracts (Figure 5-5). However, this buffer used Tris as the buffering agent, making it incompatible with downstream mass spectrometry. Thus, we tried a modified version of a lysis buffer previously used for AP of the potato leafroll virus read through protein which uses 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

(HEPES) as the buffering agent (DeBlasio et al. 2014). Amendment of this buffer with either 10, 5, or 2.5 mM DTT produced efficient $1E^{Pol}$:V5 extraction, while inclusion of 1 mM DTT was insufficient for $1E^{Pol}$:V5 extraction (Figure 5-6).

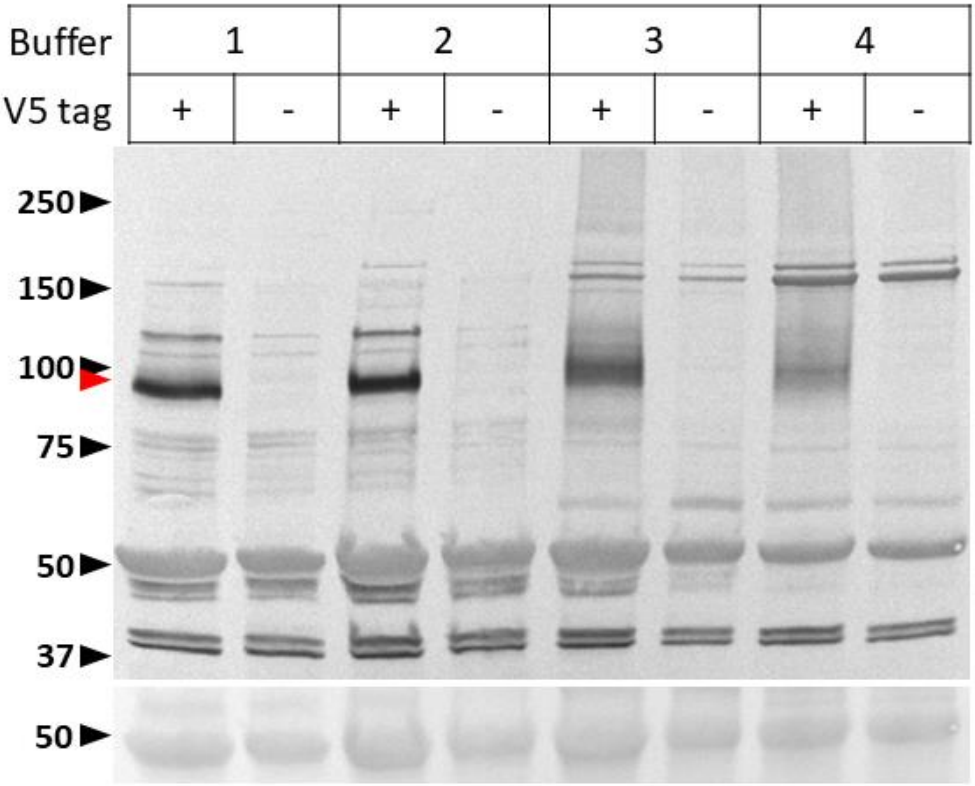


Figure 5-4. Western blot analysis of V5-tagged GFLV $1E^{Pol}$ extracted from systemically infected *N. benthamiana* leaves using different lysis buffers. Top panel shows total proteins extracted from cryogenically milled *N. benthamiana* leaves infected with either GFLV-GHu containing V5-tagged $1E^{Pol}$ (+) or wild type (untagged, -). Buffer formulations were: 1) 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10 mM DTT, 0.1% Tween-20; 2) 100 mM Tris-HCl (pH 8.1), 10 mM KCl, 5 mM $MgCl_2$, 400 mM sucrose, 10% glycerol, 1 mM β-mercaptoethanol; 3) 50 mM Tris-HCl (pH 7.6), 150 mM NaCl, 10% glycerol, 1 mM DTT, 0.1% octylphenoxy poly(ethyleneoxy)ethanol, branched (IGEPAL CA-630, Sigma-Aldrich); 4) 50 mM HEPES-KOH (pH 7.4), 110 mM KOAc, 2 mM $MgCl_2$, 0.4% TritonX-100. All buffers were also amended with 1X Halt protease inhibitor cocktail (Thermo Fisher Scientific). The predicted molecular weight of V5-tagged $1E^{Pol}$ is 93 kDa and is indicated with a red arrow head. Proteins were probed with a polyclonal anti-V5 antibody (Invitrogen PA1-993) and goat anti-rabbit AP-conjugated secondary antibody (Intvitrogen T2191) and developed with 1-Step NBT/BCIP solution. Bottom panel shows Ponceau staining of the RuBisCO large subunit. Buffer #1 was selected for additional optimization experiments.

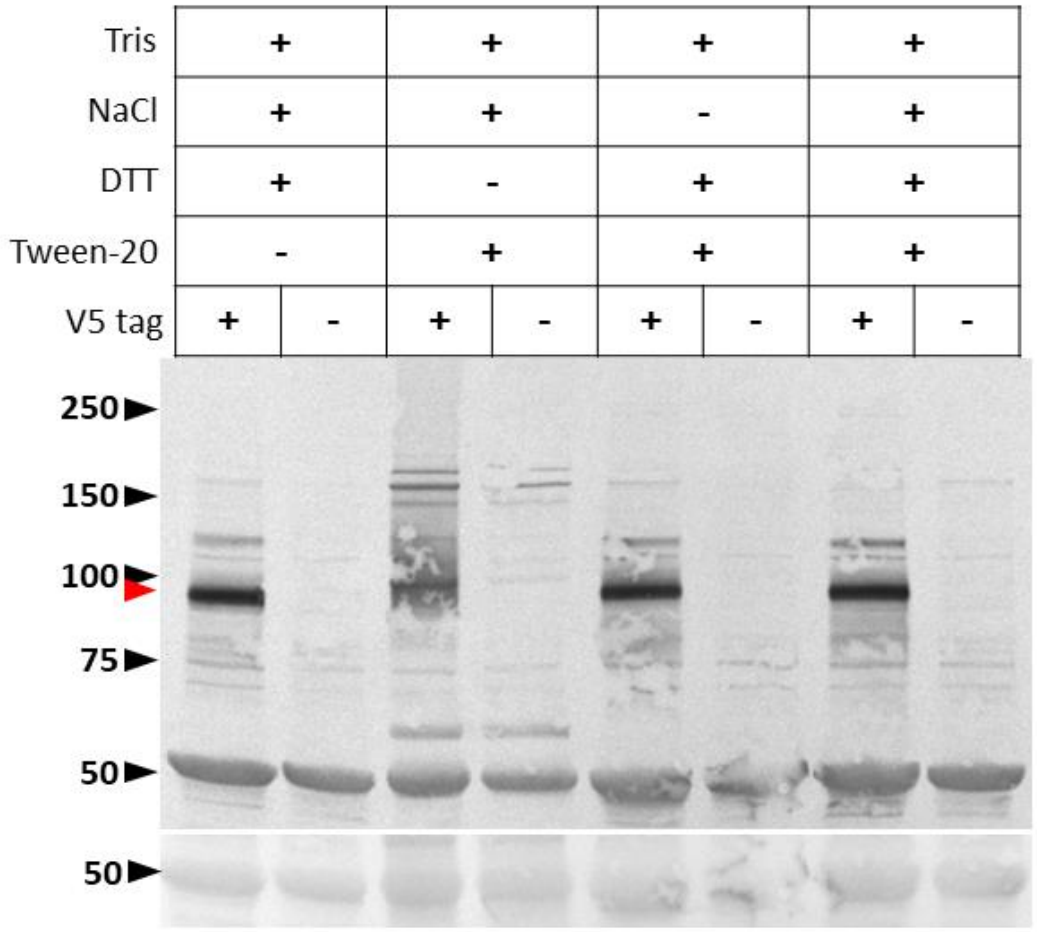


Figure 5-5. Western blot analysis of V5-tagged GFLV 1E^{Pol} extracted from systemically infected *N. benthamiana* leaves using a lysis buffer (50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10 mM DTT, 0.1% Tween-20, 1X Halt protease inhibitor cocktail) from which some components were eliminated. Top panel shows total proteins extracted from cryogenically milled *N. benthamiana* leaves infected with either GFLV-GHu containing V5-tagged 1E^{Pol} (+) or wild type (untagged, -). Lysis buffer variations were created by exclusion (-) of NaCl, DTT, or Tween-20 from the lysis buffer. The predicted molecular weight of V5-tagged 1E^{Pol} is 93 kDa and is indicated with a red arrow head. Proteins were probed with a polyclonal anti-V5 antibody (Invitrogen PA1-993) and goat anti-rabbit AP-conjugated secondary antibody (Invitrogen T2191) and developed with 1-Step NBT/BCIP solution. Bottom panel shows Ponceau staining of the RuBisCO large subunit. Single elimination of components from lysis buffer identified DTT as necessary for efficient extraction of GFLV 1E^{Pol} from *N. benthamiana* tissue. The lysis buffer supplemented with DTT was further selected for new optimization experiments.

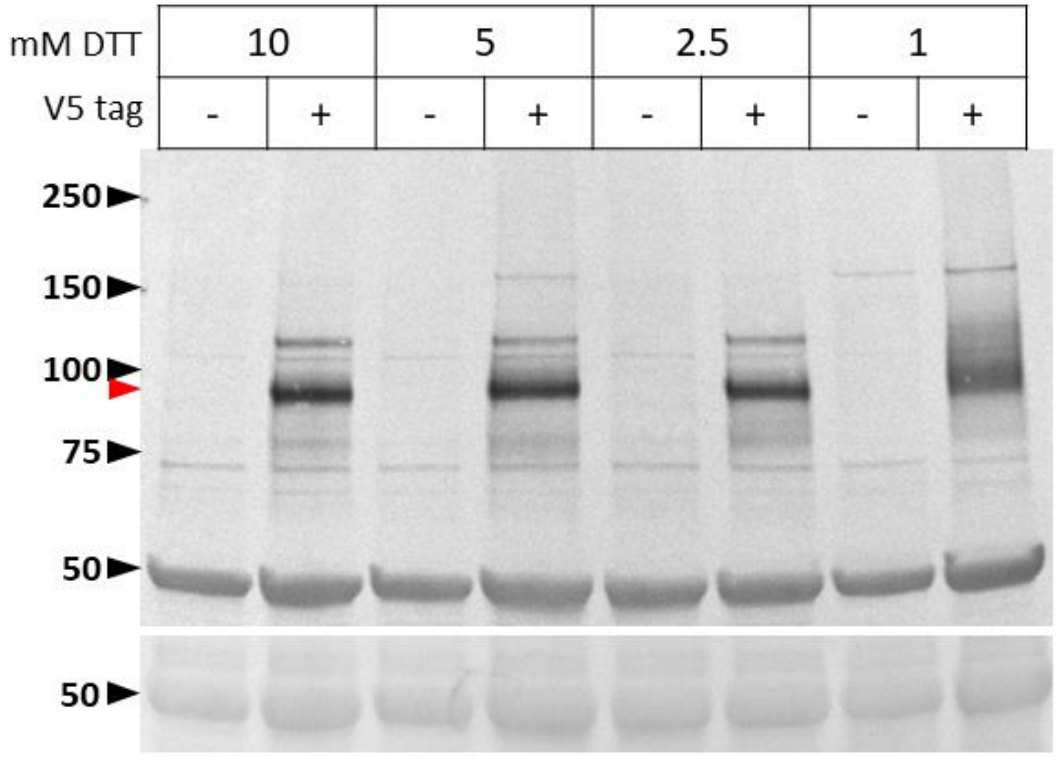


Figure 5-6. Western blot analysis of V5-tagged GFLV 1E^{Pol} extracted from systemically infected *N. benthamiana* leaves using a HEPES-based lysis buffer amended with DTT. Top panel shows total proteins extracted from cryogenically milled *N. benthamiana* leaves infected with either GFLV-GHu containing V5-tagged 1E^{Pol} (+) or wild type (untagged, -). Base lysis buffer formulation was 50 mM HEPES-KOH (pH 7.4); 110 mM potassium acetate; 2 mM $MgCl_2$; 0.4% Triton X-100, 1X Halt protease inhibitor. Base lysis buffer was amended with 10, 5, 2.5, or 1 mM DTT. The predicted molecular weight of V5-tagged 1E^{Pol} is 93 kDa and is indicated with a red arrow head. Proteins were probed with a polyclonal anti-V5 antibody (Invitrogen PA1-993) and goat anti-rabbit AP-conjugated secondary antibody (Invitrogen T2191) and developed with 1-Step NBT/BCIP solution. Bottom panel shows Ponceau staining of the RuBisCO large subunit. Optimal detection of V5-tagged GFLV 1E^{Pol} was achieved with HEPES-based lysis buffer amended with 2.5 mM DTT.

However, DTT has the potential to interfere with protein-protein interactions, including the interactions which allow antibodies to bind their antigens. Since exclusion of DTT was not possible, we sought to minimize the amount of DTT present in the final soluble fraction. We found that 1:5 dilution of supernatants with DTT-free lysis buffer had no effect on the efficiency of $1E^{Pol}$:V5 extraction (Figure 5-7). This brought the final concentration of DTT in the soluble fraction to 0.5 mM when using lysis buffer containing 2.5 mM DTT for the initial extraction.

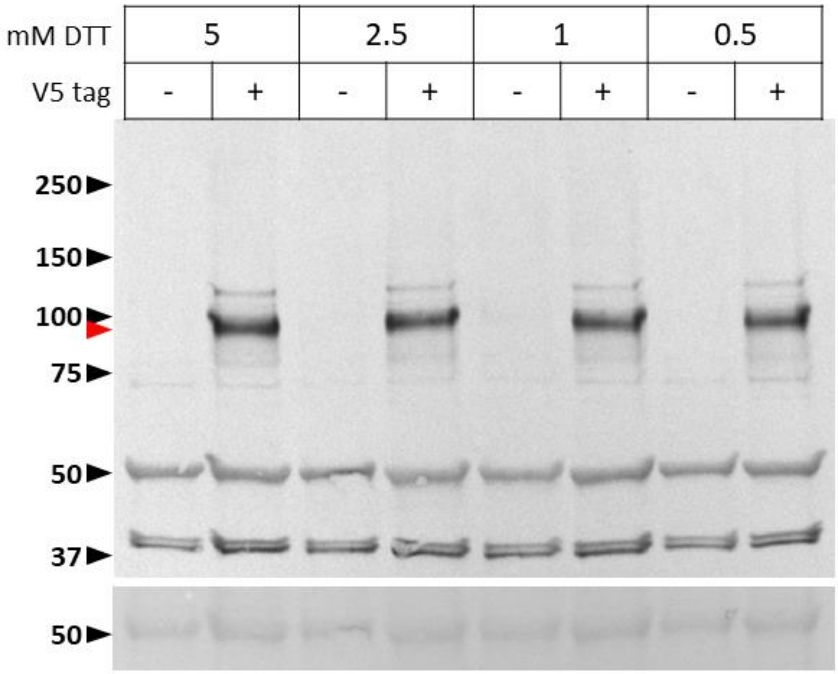


Figure 5-7. Western blot analysis of V5-tagged GFLV $1E^{Pol}$ extracted from systemically infected *N. benthamiana* leaves using a HEPES-based lysis buffer amended with DTT followed by a post-extraction dilution of DTT. Total soluble proteins from *N. benthamiana* leaves infected with GFLV-GHu containing V5-tagged $1E^{Pol}$ (+) or wild type (-) were extracted in lysis buffer containing either 5 or 2.5 mM DTT; supernatants were then diluted 1:5 with original DTT-containing lysis buffer (final DTT concentration 5 and 2.5 mM, respectively) or with DTT-free lysis buffer (final DTT concentration 1 and 0.5 mM, respectively). The predicted molecular weight of V5-tagged $1E^{Pol}$ is 93 kDa and is indicated with a red arrow head. Proteins were probed with a polyclonal anti-V5 antibody (Invitrogen PA1-993) and goat anti-rabbit AP-conjugated secondary antibody (Invitrogen T2191) and developed with 1-Step NBT/BCIP solution. Bottom panel shows Ponceau staining of the RuBisCO large subunit. A post-extraction dilution of total soluble proteins of *N. benthamiana* with DTT does not impact extraction and detection of GFLV $1E^{Pol}$ from *N. benthamiana* tissue.

Up to this point, we had utilized GHu-1E:V5 and wild type GFLV-GHu for our optimizations. To demonstrate that the lysis conditions we identified were sufficient for the extraction of other GFLV variants which were intended for AP, we performed small-scale protein extractions of F13-1E:$V5_{C-1}$ and GHu-1E-K802G:V5, and used untagged GFLV-F13 and –GHu-1E-K802G as negative controls, respectively. We also included tissue from mock-inoculated *N. benthamiana* plants. We found that the lysis conditions optimized for GHu-1E:V5 were sufficient for extraction of $1E^{Pol}$:V5 from both GFLV-F13 and GFLV-GHu-1E-K802G (Figure 5-8).

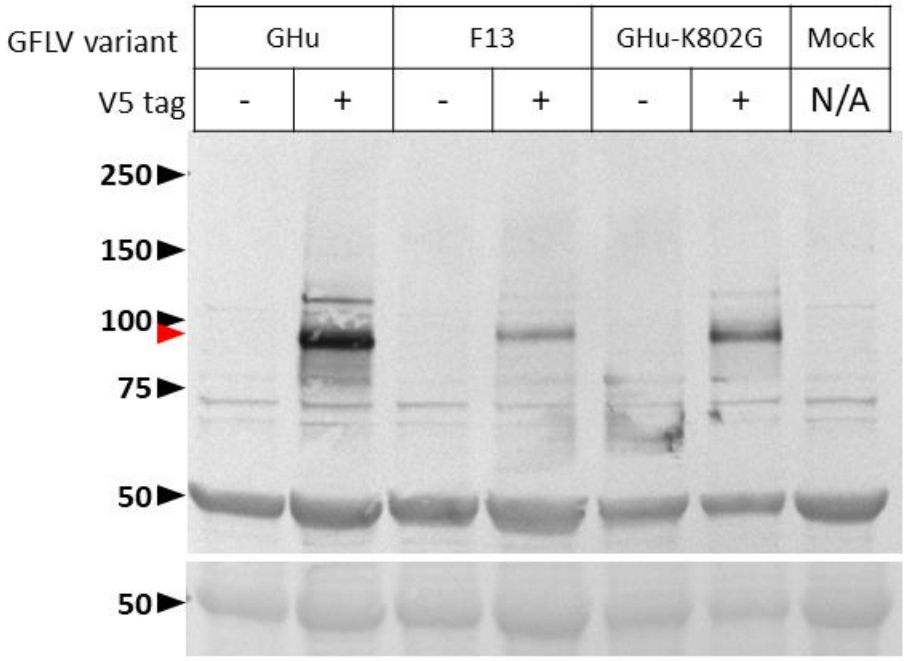


Figure 5-8. Western blot analysis of three V5-tagged GFLV $1E^{Pol}$ extracted from systemically infected *N. benthamiana* leaves. Total soluble proteins (TSP) from *N. benthamiana* infected with GFLV variants GHu, F13, or GHu-1E-K802G containing V5-tagged $1E^{Pol}$ (+) or wild type (untagged, -) were extracted and analyzed. TSP was also extracted from a healthy, mock-inoculated *N. benthamiana*. Proteins were probed with polyclonal anti-V5 antibody (Invitrogen PA1-993) and goat anti-rabbit AP-conjugated secondary antibody (Invitrogen T2191) and developed with 1-Step NBT/BCIP solution. The predicted molecular weight of V5-tagged $1E^{Pol}$ is 93 kDA and is indicated with a red arrow head. Bottom panel shows Ponceau staining of the RuBisCO large subunit. Results show that the V5 epitope tag of the three GFLV variants is successfully detected by western blot.

Proteins identified as likely interaction partners of 1E^{Pol}

For reasons that remain unknown, peptide samples prepared for mass spectrometry with OMIX C18 100 µL tips caused significant issues during mass spectrometry, especially in clogging the liquid chromatography column prior to introduction to the mass spectrometer (K. Rivera, Cold Spring Harbor Laboratory, personal communication). In an attempt to resolve these issues and remove possible contaminants, samples were treated with Millipore C18 10µL ZipTips. While these samples had fewer clogging issues, it appears this came at the expense of depth of protein identification: Mascot searches of mass spectrometry runs from OMIX-prepared samples identified an average of 740 proteins per analytical replicate; analytical replicates from ZipTip-prepared samples had an average of 460 proteins identified by Mascot. While OMIX-prepared samples generally resulted in higher numbers of proteins being identified, there was still significant variability between analytical replicates of some samples (e.g. one OMIX-prepared V5-tagged GFLV-GHu biological replicate had 1196 proteins identified in one injection run and only 53 in the second injection). Ideally, variation in protein identification levels across analytical and biological replicates should be low to enable confident label-free quantification by spectral counting (Lundgren et al. 2010). To meet this requirement to the best of our ability with the available dataset, we selected a subset of the analytical replicates from both OMIX and ZipTip prepared samples for inclusion the Scaffold analysis (Table 5-4).

Table 5-4. Analytical replicates of mass spectrometry runs of tryptic peptides from affinity purification of V5-tagged grapevine fanleaf virus (GFLV) $1E^{Pol}$ from systemically infected *Nicotiana benthamiana* tissue and their total protein identifications from Mascot database searching.

Table 5-4					
Biorep	**Tip[a]**	**Inject[b]**	**Treatment**	**# prot[c]**	**Scaffold[d]**
1	OMIX	1	GHu:V5	1387	*
1	OMIX	2	GHu:V5	874	*
1	Zip	1	GHu:V5	400	*
1	Zip	2	GHu:V5	413	*
3	OMIX	1	GHu:V5	1197	*
4	OMIX	1	GHu:V5	1196	*
4	OMIX	2	GHu:V5	53	
4	Zip	1	GHu:V5	973	*
4	Zip	2	GHu:V5	993	*
5	OMIX	1	GHu (nt)	1142	*
5	OMIX	2	GHu (nt)	1330	*
5	Zip	1	GHu (nt)	544	
5	Zip	2	GHu (nt)	560	
6	OMIX	1	GHu (nt)	1324	*
7	OMIX	1	GHu (nt)	48	
7	OMIX	2	GHu (nt)	39	
7	Zip	1	GHu (nt)	722	*
7	Zip	2	GHu (nt)	728	*
9	OMIX	1	F13:V5	980	*
9	OMIX	2	F13:V5	1178	*
9	Zip	1	F13:V5	479	*
9	Zip	2	F13:V5	497	*
10	OMIX	1	F13:V5	916	*
10	OMIX	2	F13:V5	48	
10	Zip	1	F13:V5	329	*
10	Zip	2	F13:V5	330	
11	OMIX	1	F13:V5	37	
11	OMIX	2	F13:V5	32	
11	Zip	1	F13:V5	493	*
11	Zip	2	F13:V5	507	*
13	OMIX	1	F13 (nt)	834	*
13	OMIX	2	F13 (nt)	947	*
13	Zip	1	F13 (nt)	281	*
13	Zip	2	F13 (nt)	289	*
14	Zip	1	F13 (nt)	47	
14	Zip	2	F13 (nt)	400	
14	Zip	3	F13 (nt)	465	*

Table 5-4					
Biorep	**Tip[a]**	**Inject[b]**	**Treatment**	**# prot[c]**	**Scaffold[d]**
14	Zip	4	F13 (nt)	407	
15	OMIX	1	F13 (nt)	32	
15	OMIX	2	F13 (nt)	31	
15	Zip	1	F13 (nt)	77	
15	Zip	2	F13 (nt)	45	
17	OMIX	1	GHu-K802G:V5	747	
17	OMIX	2	GHu-K802G:V5	894	*
17	Zip	1	GHu-K802G:V5	1177	*
17	Zip	2	GHu-K802G:V5	219	
19	OMIX	1	GHu-K802G:V5	1265	*
19	OMIX	2	GHu-K802G:V5	1266	*
19	Zip	1	GHu-K802G:V5	233	*
19	Zip	2	GHu-K802G:V5	236	*
20	OMIX	1	GHu-K802G:V5	1392	
20	OMIX	2	GHu-K802G:V5	68	
20	Zip	1	GHu-K802G:V5	522	*
20	Zip	2	GHu-K802G:V5	517	*

[a]Type of tip used for preparation of tryptic peptides prior to mass spectrometry analysis. OMIX = OMIX C18 100 µL tips (Agilent); Zip = C18 10µL ZipTips (Millipore).
[b]Injection number (analytical replicate)
[c]Number of proteins identified from the mass spectrometry run by Mascot database searching of a *N. benthamiana*-contaminant-nepovirus protein database
[d]* indicates that the analytical replicate was included in the Scaffold analysis

In total, 3119 proteins in 1236 clusters were identified across all the samples (2 peptides minimum per protein). There were 806 protein clusters identified across all tagged samples prior to filtering for contaminant proteins, such as those found in the untagged negative control APs (Figure 5-9A); filtering for these contaminants reduced the total number of protein clusters to 156 (Figure 5-9B). To identify proteins which were specifically enriched in certain AP conditions, we performed spectral counting (SC) in Scaffold, normalized to total spectra. ANOVA was used to identify proteins which had spectral counts which were significantly different ($p < 0.05$) between AP conditions. Interestingly, while 8 clusters were found to be common to the tagged AP categories (Figure 5-9B), only one protein cluster (apart from $1E^{Pol}$) was found to be significantly enriched in tagged GFLV samples relative to the untagged controls: the GFLV-GHu $1D^{Pro}$ cluster. This cluster includes GFLV-F13 $1D^{Pro}$. Also, spectra were identified for GFLV $1B^{Hel}$ in most tagged GFLV biological replicates, while no spectra for this protein were found in the untagged controls. However, due to the low spectral counts for this protein, it did not reach the threshold for significance by ANOVA ($p = 0.098$). Of the remaining protein clusters which met the significance threshold of $p < 0.05$, a few are specific to the asymptomatic GFLV-GHu mutant, including one chloroplast ribosomal protein, cinnamate-4-hydroxylase (a biosynthesis enzyme), and two proteins involved in protein metabolism (Table 5-5). However, the majority of specifically enriched proteins clusters are those which are specifically enriched in tagged GFLV-GHu (Table 5-5).

Among the 156 protein clusters identified (post-filtering) across all V5-tagged samples, 59 clusters are specific to GHu-1E:V5 (Figure 5-9). Among those which are

statistically significantly enriched specfically in APs from V5-tagged GFLV-GHu $1E^{Pol}$, there appear to be some common trends (Table 5-5). The following proteins functions were retrieved from UniProt (www.uniprot.org). Quite a few of the significantly enriched proteins are involved in protein metabolism, synthesis, and folding. These include key components of the ribosome (eukaryotic translation initiation factor 2 subunit β and a structural constituent of the ribosome). The 26S proteasome non-ATPase regulatory subunit 7 homolog is involved in the degradation of ubiquinated proteins, while ATP phosphoribosyltransferase HisG catalyzes the initial step for the biosynthesis of histidine. Peptidly-prolyl cis-trans isomerase accelerates protein-folding by triggering isomerization of proline peptide bonds during protein synthesis. Chaperone protein DNaJ is a heat-shock protein which aids in the re-folding of proteins. There are also proteins associated with mediating metabolic stress, such as lactoylglutathione lyase (detoxifies methyglyoxal, a normal but toxic byproduct of cellular metabolism) and catalase (decomposes the reactive oxygen species hydrogen peroxide into water and oxygen).

There are also a fair number of proteins are associated with the chloroplast. Plastid transcriptionally active 14 is crucial for thylakoid development and may play a role in gene expression within the chloroplast, while plastid transcriptionally active 12 is necessary for photomorphogenesis by regulating plastid gene expression. Similarly, pentatricopeptide repeat-containing proteins are RNA-binding proteins which are usually targeted to plastids to regulate expression of plastid RNAs. There is a pentatricopeptide-repeat protein that is enriched for GFLV-GHu alone and another which is enriched in both GFLV-GHu variants. There is also a putative chloroplast

plasma membrane-associated proteins among this group. Finally, magnesium-chelatase subunit H narrowly missed the cut-off for significance by ANOVA ($p = 0.065$), yet is another example of a chloroplast-associated protein. This protein inserts magnesium ions into chlorophyll precursors and is essential to chlorophyll synthesis. Interestingly, this protein appears in APs with both wild-type GFLV-GHu $1E^{Pol}$ and the asymptomatic GHu-1E-K802G mutant. Proteins specifically enriched in APs of V5-tagged GFLV-GHu $1E^{Pol}$ also include histone H1 (which stabilizes histone complexes), endogluconase 17 (which cleaves 1-4-β-D-glucosidic linkages in cellulose) and DNA-directed RNA polymerase subunit β.

Several of these specifically enriched proteins are helicases (Holliday junction ATP-dependent DNA helicase RuvB and ATP-dependent RNA helicase RhlB). Others are splicing factors involved in pre-mRNA processing (splicing factor 2B subunit 1 and pre-mRNA processing splicing factor 8). There are several which are associated with cellular trafficking, such as myosin (actin-associated motor protein), AP-2 complex subunit μ (an adapter protein for trafficking of clathrin-coated vesicles), and a kinesin-like protein (microtubule-associated motor protein). Interestingly, another AP-complex protein, AP-1 complex subunit μ-2, is enriched for both GHu and F13. There are also proteins for signal transduction, nutrient transport, and other processes which require transmembrane trafficking (2-oxoglutarate translocator, a SH3 domain-containing protein, plasma membrane ATPase 1, and a Membrane Occupation and Recognition Nexus [MORN] repeat-containing protein). There is both a positive (COP1-interacting protein) and negative (importin subunit β-1) regulator of abscisic acid-signaling, which controls drought stress through stomatal closure.

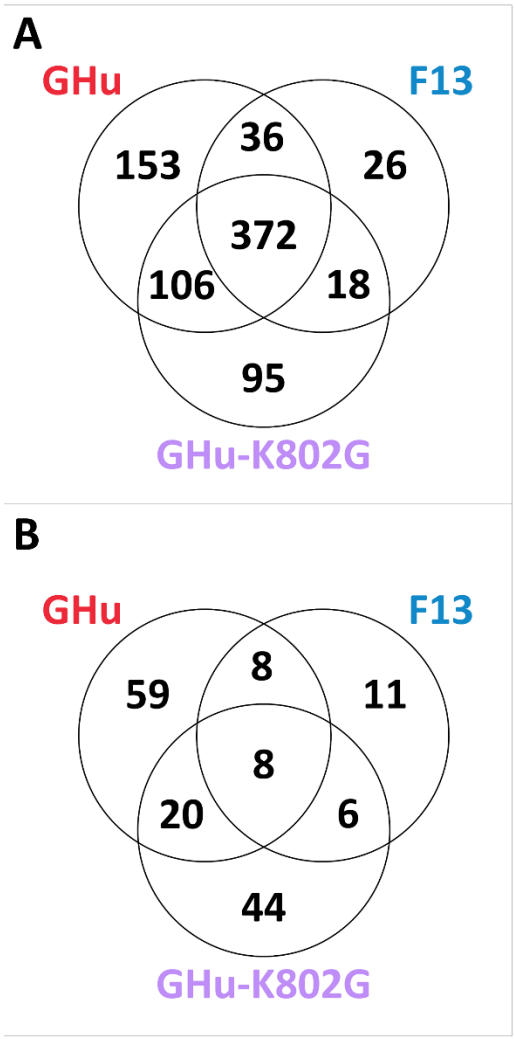


Figure 5-9. Venn diagrams showing distribution of protein clusters identified via affinity purification (AP) of V5-tagged grapevine fanleaf virus (GFLV) $1E^{Pol}$ variants from systemically infected *Nicotiana benthamiana* tissue. **A.** Distribution of all protein clusters identified in APs of V5-tagged GFLV $1E^{Pol}$. A total of 806 protein clusters are displayed. **B.** Distribution of protein clusters identified in APs of V5-tagged GFLV $1E^{Pol}$ following filtering to remove protein clusters identified in negative control APs of *N. benthamiana* tissue infected with untagged GFLV. A total of 156 protein clusters are displayed. GHu: GFLV strain GHu which produces vein clearing in *N. benthamiana.* F13: GFLV strain F13, which is asymptomatic in *N. benthamiana*. GHu-K802G: an asymptomatic mutant of GFLV-GHu in which lysine 802 of $1E^{Pol}$ has been mutated to glycine.

Table 5-5. Spectral counts and quantitative profiles for proteins of interest from affinity purifications (APs) of V5-tagged grapevine fanleaf virus (GFLV) $1E^{Pol}$ from systemically-infected *Nicotiana benthamiana* tissue.

Table 5-5

		Quantitative Profile[a]				GHu:V5[b]			F13:V5[b]			K802G:V5[b]			GHu (nt)[b]			F13 (nt)[b]	
Protein ID[c]	**p ANOVA**	**G**	**F**	**K**	**nt**	**1**	**3**	**4**	**9**	**10**	**11**	**17**	**19**	**20**	**5**	**6**	**7**	**13**	**14**
AFM91094.1_GFLV-GHu_1E	0.0038	*		*		697	135	1200	136	34	29	672	614	306	22	20	5	0	0
NP_619689.1_GFLV-F13_1E	0.028	*	*	*		226	37	377	426	116	92	220	191	106	9	8	5	0	0
AFM91094.1_GFLV-GHu_1D	0.002	*		*		34	11	33	9	0	0	20	32	10	0	0	0	0	0
NP_619689.1_GFLV-F13_1D	0.025	*	*	*		22	8	26	22	0	0	14	21	10	0	0	0	0	0
AFM91094.1_GFLV-GHu_1B	0.098					8	2	7	7	3	0	12	7	0	0	0	0	0	0
Niben101Scf05875g00012.1 Kinesin-like protein	<0.0001	*				2	2	3	0	0	0	0	0	0	0	0	0	0	0
Niben101Scf00548g00009.1 Pentatricopeptide repeat-containing protein	0.00012	*				3	2	3	0	0	0	0	2	0	0	0	0	0	0
Niben101Scf11366g02026.1 Holliday junction ATP-dependent DNA helicase RuvB	0.00019	*				5	3	9	0	0	0	0	0	0	0	0	0	0	0
Niben101Scf01063g05019.1 Peptidyl-prolyl cis-trans isomerases	0.00043	*				6	6	7	2	0	0	0	2	0	2	3	0	0	0
Niben101Scf03607g00009.1 Plastid transcriptionally active 14	0.0017	*				4	4	5	2	2	0	0	3	0	0	0	0	0	0
Niben101Scf14859g00005.1 2-oxoglutarate translocator	0.0027	*				3	3	5	0	0	0	0	0	0	0	3	0	0	0
Niben101Scf01623g01025.1 Endoglucanase	0.0033	*				6	3	5	2	0	0	3	0	0	2	0	0	0	0
Niben101Scf06078g00001.1 DNA-directed RNA polymerase subunit beta	0.0077	*				3	2	3	0	0	0	0	0	0	0	0	0	2	0
Niben101Scf14996g00009.1 Catalase	0.0079	*				10	7	12	0	0	0	0	5	0	7	7	0	5	0
Niben101Scf04156g00011.1 Importin subunit beta-1	0.014	*				3	2	5	3	0	0	2	0	0	0	0	0	0	0

Table 5-5, cont.																			
		Quantitative Profile[a]				GHu:V5[b]			F13:V5[b]			K802G:V5[b]			GHu (nt)[b]			F13 (nt)[b]	
Protein ID[c]	p ANOVA	G	F	K	nt	1	3	4	9	10	11	17	19	20	5	6	7	13	14
Niben101Scf08991g00007.1 Lactoylglutathione lyase	0.015	*				3	2	9	0	0	0	0	3	0	0	0	0	0	0
Niben101Scf04077g02006.1 26S proteasome non-ATPase reg. subunit 7 homolog A	0.015	*				2	2	2	0	0	0	0	3	0	0	0	0	0	0
Niben101Scf00262g05007.1 eIF 2 subunit beta	0.02	*				4	0	4	0	0	0	0	0	0	0	0	0	0	0
Niben101Scf02069g02008.1 COP1-interactive protein 1	0.02	*				2	0	2	0	0	0	0	0	0	0	0	0	0	0
Niben101Scf08698g02020.1 Plastid transcriptionally active 12	0.021	*				0	2	2	0	0	0	0	0	0	0	0	0	0	0
Niben101Scf08026g02006.1 AP-2 complex subunit mu	0.022	*				0	2	3	0	0	0	0	0	0	0	0	0	0	0
Niben101Scf03492g01004.1 Pre-mRNA-processing-splicing factor 8	0.022	*				4	2	5	3	0	0	0	2	0	0	2	0	0	0
Niben101Scf04172g02008.1 Myosin 2	0.024	*				7	3	9	3	0	0	7	0	0	0	0	0	0	0
Niben101Scf03070g08002.1 Structural constituent of ribosome	0.025	*				3	0	2	0	0	0	0	0	0	0	0	0	0	0
Niben101Scf06128g00004.1 Splicing factor 3B subunit 1	0.028	*				3	4	11	4	0	0	3	4	0	0	0	0	0	0
Niben101Scf10476g01004.1 Chloroplast, plasma membrane […] putative	0.035	*				2	2	7	0	0	0	2	0	0	3	0	0	0	0
Niben101Scf04959g00010.1 ATP-dependent RNA helicase RhlB	0.036	*				5	0	6	0	0	0	0	0	0	2	0	0	0	0
Niben101Scf01970g01025.1 ATP phosphoribosyltransferase	0.036	*				3	6	10	2	0	0	6	3	0	0	4	0	0	0
Niben101Scf03979g02010.1 Plasma membrane ATPase 1	0.037	*				9	5	16	5	0	0	5	7	0	6	4	3	0	0

Table 5-5, cont.																			
		Quantitative Profile[a]				GHu:V5[b]			F13:V5[b]			K802G:V5[b]			GHu (nt)[b]			F13 (nt)[b]	
Protein ID[c]	p ANOVA	G	F	K	nt	1	3	4	9	10	11	17	19	20	5	6	7	13	14
Niben101Scf03138g01010.1 Chaperone protein DnaJ	0.038	*				17	13	44	7	3	0	8	9	2	26	9	7	11	0
Niben101Scf05487g01009.1 Histone H1	0.039	*				4	0	2	0	0	0	0	0	0	0	0	0	0	0
Niben101Scf06461g06012.1 SH3 domain-containing protein	0.05	*				2	6	14	2	0	0	2	5	0	2	3	0	0	0
Niben101Scf00949g03003.1 MORN repeat-containing protein	0.05					4	0	6	2	0	0	0	0	0	0	0	0	0	0
Niben101Scf04388g00011.1 Magnesium-chelatase subunit H	0.065					4	3	10	0	0	0	5	8	0	6	0	0	0	0
Niben101Scf00650g02015.1 AP-1 complex subunit mu-2	0.045	*	*			6	9	45	0	0	0	6	0	0	0	0	0	0	0
Niben101Scf02407g03001.1 Pentatricopeptide repeat 5	0.017	*		*		3	0	2	0	0	0	3	6	0	0	0	0	0	0
Niben101Scf05848g05012.1 Isopentenyl-diphosphate Delta-isomerase II	0.025			*		0	0	3	0	0	0	3	4	0	0	0	0	0	0
Niben101Scf06211g01011.1 50S ribosomal protein L14, chloroplastic	0.041			*		3	6	14	0	0	0	9	9	0	0	0	0	0	0
Niben101Scf07400g01008.1 Cullin-associated NEDD8-dissociated protein 1	0.041			*		0	3	3	0	0	0	4	2	0	0	0	0	0	0
Niben101Scf08196g01007.1 Cinnamate-4-hydroxylase	0.02			*		0	0	0	0	0	0	3	3	0	0	0	0	0	0

[a]Affinity purification categories in which proteins were found to be significantly ($p < 0.05$) enriched are designated with a *. G = V5-tagged GFLV-GHu $1E^{Pol}$; F = V5-tagged GFLV-F13 $1E^{Pol}$; K = V5-tagged GFLV-GHu $1E^{Pol}$ mutant K802G; nt = no tag (wild-type GFLV-GHu and GFLV-F13 negative controls were pooled into a single "no tag" AP category).
[b]The columns below each AP condition designate individual biological replicates. The sample number of each biological replicate is shown.
[c]Rows are color-coded by protein function or origin. Grey rows are GFLV proteins. Yellow rows are chloroplast-associated proteins. Pink rows are helicases. Brown rows are proteins involved in protein metabolism or biosynthesis. Orange rows are proteins involved in signal transduction or transmembrane trafficking. Dark blue rows are proteins involved in mitigating metabolic stress. Purple rows are proteins involved in regulation of abscisic acid signaling. Green rows are proteins involved in cellular trafficking. Light blue rows are pre-mRNA processing proteins. Unshaded rows do not share an apparent function with other proteins in the list.

DISCUSSION

Little is known of the structure of the $1E^{Pol}$ C-terminus. Scaffolded protein modeling of this region is unreliable due to the lack of solved crystal structures of viral RdRps containing the C-terminal extension seen in nepoviruses. We had previously applied several *in silico* analyses to this region of the $1E^{Pol}$ protein to glean clues as to its structure and function. We found that this region is not strongly predicted to be inherently disordered. It also harbors structural motifs flanking the symptom determinant residue 802 which are highly conserved across sequenced GFLV isolates. Interestingly, the C-terminus of $1E^{Pol}$ is deficient in acidic residues (Table 5-6). Note that the C-terminus region in this table was chosen to corresponds to a region found to be highly sensitive to mutations which render GFLV unable to establish systemic infection in *N. benthamiana* via agroinoculation (Figures 4-1C and D, 4-3A). Protein domains involved in membrane interactions are often enriched in basic and/or hydrophobic residues. Also, attempts to add acid-rich epitope tags (e.g. FLAG variants, HA, and myc) rendered GFLV constructs non-infectious via agroinoculation. Interestingly, of the epitope tags we tried, only V5 permitted systemic infection of GFLV constructs. The V5 peptide is derived from the P and V proteins of simian virus 5 (Southern et al. 1991). The V5 tag is largely hydrophobic and contains a single acidic residue (out of a total of 14 amino acids, 7% acidic residues). In comparison, acidic residues make up 5 of 8 FLAG residues (62.5%), 11 out of 22 3XFLAG residues (50%), 4 of 10 myc residues (40%) and 2 of 9 HA residues (22%). This may be evidence that the functionality of $1E^{Pol}$ is dependent on the charge status of the C-terminus. It is also interesting that extraction of $1E^{Pol}$:V5 required the inclusion of a

reducing agent (β-ME or DTT) in the lysis buffer. It may be that a reducing environment is necessary to dissociate $1E^{Pol}$ from membranes, nucleic acids (especially RNA since it is an RdRp), or other macromolecular structures that would otherwise prevent it from being released into the soluble protein fraction during extraction.

Affinity purification of V5-tagged $1E^{Pol}$, coupled with protein identification by high-accuracy tandem mass spectrometry provided us with new insights into the putative protein interaction network of GFLV $1E^{Pol}$. We found that GFLV protein $1D^{Pro}$, the viral protease which processes viral polyproteins into individual mature peptides (Margis et al. 1991), is enriched in APs with V5-tagged $1E^{Pol}$, suggesting that $1D^{Pro}$ interacts with $1E^{Pol}$ during virus infection. It is also possible that $1D^{Pro}$ peptides are enriched in APs with $1E^{Pol}$ due to the presence of a putative 1DE precursor. Indeed, in our initial AP trials, we saw enrichment of a specific band which appears to correspond to the predicted molecular weight of a 1DE precursor (Figure 5-10). Precursors of viral proteins containing polymerase (Pol) domains have been detected in other nepovirus systmes. A polyprotein consisting of VPg, Pro, and a truncated Pol domains (VPg-Pro-Pol') were detected in cucumber infected with a subgroup C nepovirus, tomato ringspot virus (ToRSV). The VPg-Pro-Pol' was found to peripherally associate with cellular membranes and localized to virus replication complexes (Chisholm et al. 2007). A similar polyprotein consisting of VPg-Pro-Pol (that is, untruncated Pol) was also found for a subgroup B nepovirus, beet ringspot virus (Demangeat et al. 1992). The functional significance of these precursors has not been determined.

Table 5-6. Residue composition of protein $1E^{Pol}$ of grapevine fanleaf virus (GFLV) strains F13 and GHu

			Basic residues[a]		Acidic residues[b]		Hydrophobic residues[c]		Polar residues[d]		Other residues[e]	
Strain	Region	Residues	#	%	#	%	#	%	#	%	#	%
F13	Full $1E^{Pol}$	1-824	96	11.7	104	12.6	271	32.9	210	25.5	143	17.4
F13	$1E^{Pol}$ C-terminus[f]	771-824	10	18.5	2	3.7	19	35.2	18	33.3	5	9.3
GHu	Full $1E^{Pol}$	1-824	98	11.9	104	12.6	278	33.7	201	24.4	143	17.4
GHu	$1E^{Pol}$ C-terminus[f]	771-824	10	18.5	1	1.9	17	31.5	18	33.3	8	14.8

[a]Lysine and arginine
[b]Aspartic acid and glutamic acid
[c]Alanine, isoleucine, leucine, phenylalanine, tryptophan, valine
[d]Asparagine, cysteine, glutamine, serine, threonine, tyrosine
[e]Glycine, histidine, methionine, proline
[f]The C-terminus region corresponds to a region previously found to be highly sensitive to mutations which render GFLV unable to establish systemic infection in *N. benthamiana* via agroinoculation.

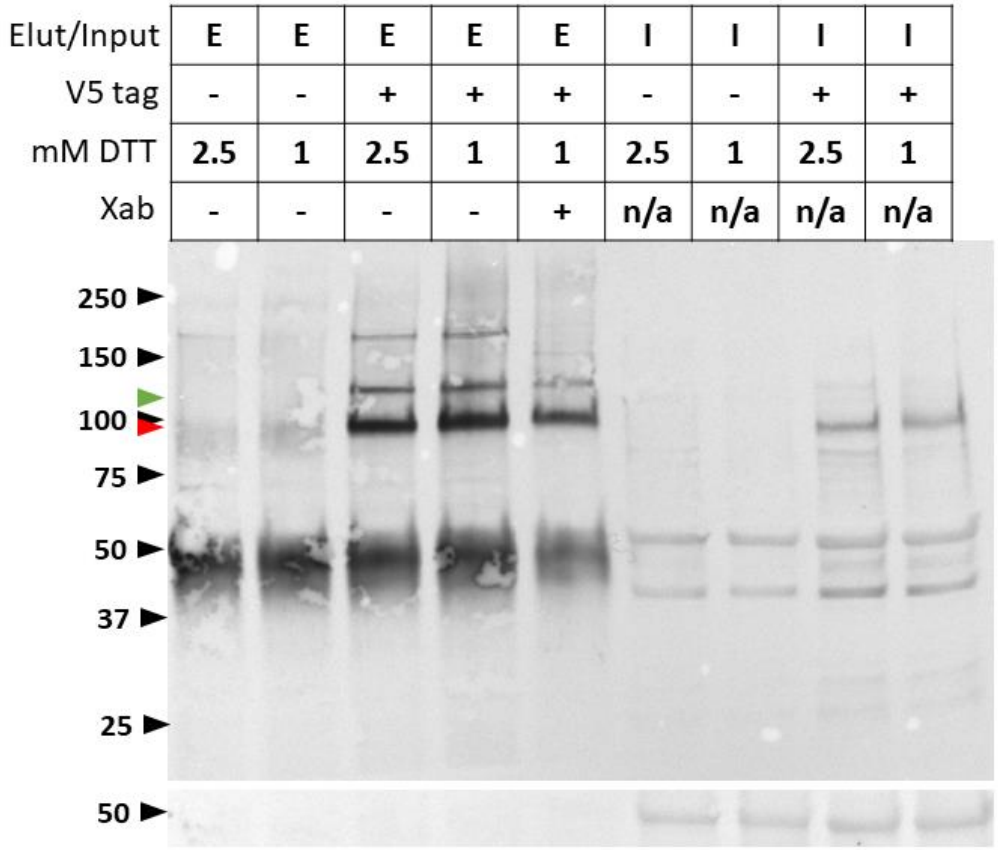


Figure 5-10. Western blot analysis of small-scale affinity purifications (AP) of V5-tagged GFLV-GHu $1E^{Pol}$ from systemically infected *N. benthamiana* leaves. Total soluble proteins from systemically-infected *N. benthamiana* were extracted in lysis buffer (50 mM HEPES-KOH (pH 7.4); 110 mM potassium acetate; 2 mM $MgCl_2$; 0.4% Triton X-100, 5 or 1 mM DTT, 1X Halt protease inhibitor), then diluted 1:5 in DTT-free lysis buffer. Inputs (I) were affinity purified with V5 polyclonal antibody (Invitrogen PA1-993) conjugated to Protein A Dynabeads (Invitrogen). Beads were washed 4 times with DTT-free lysis buffer and twice with phosphate buffered saline. AP products were eluted from beads at 70°C for 10 minutes in 50 µL of 2X Laemmli Buffer supplemented with 5% β-mercaptoethanol (β-ME) and 40 µL were loaded onto the gel. Input fractions were mixed 1:1 with 2X Laemmli Buffer supplemented with 5% β-ME and denatured at 70°C for 10 minutes and 40 µL of prepared samples were loaded onto the gel. The top row of lane labels indicates whether the lane contains elution products (E) or input fractions (I) of the AP. The second row indicates whether the tissue used for the AP contained V5-tagged (+) or untagged (-) GFLV-GHu. The third row indicates the concentration of DTT included in the initial lysis buffer. The fourth row indicates whether the V5 polyclonal antibody used for AP had been cross-absorbed against total soluble proteins from healthy (uninfected) *N. benthamiana* tissue (+) or had been untreated (-). Proteins were probed with a polyclonal anti-V5 antibody (Invitrogen PA1-993, non-cross-absorbed) and goat anti-rabbit AP-conjugated secondary antibody (Invitrogen T2191) and developed with 1-Step NBT/BCIP solution. The predicted molecular weight of V5-tagged $1E^{Pol}$ is 93 kDa and is indicated with a red arrow head. The predicted molecular weight of a putative V5-tagged 1DE precursor is 117 kDa and is indicated with a green arrow head. Bottom panel shows Ponceau staining of the RuBisCO large subunit.

We also found that GFLV protein $1B^{Hel}$ appears to be enriched in GFLV $1E^{Pol}$ APs. Protein $1B^{Hel}$ is a putative helicase and is predicted to be an RNA-binding protein. That $1B^{Hel}$ is co-purified with a viral RNA-dependent RNA polymerase is perhaps unsurprising. It would be interesting to determine if $1B^{Hel}$ enrichment is the result of direct interactions between $1E^{Pol}$ and $1B^{Hel}$, or if they co-enrich due to mutual interaction with viral RNA or other protein interactors. There is evidence that $1E^{Pol}$ does not localize to membranes when expressed alone in *N. benthamiana* leaves and thus must require other viral factors for localization to the replication complex (C. Schmitt-Keichinger and F. Berthold, unpublished results presented in Schmitt-Keichinger et al. 2017). It has been proposed that $1B^{Hel}$ may be this factor since it is the only GFLV protein that contains predicted transmembrane domains (Schmitt-Keichinger et al. 2017).

Beyond these two viral proteins, host proteins which were found by spectral counting to be specifically enriched in particular AP conditions are mostly those which were specifically found in APs of GFLV-GHu. Amongst these proteins, there appear to be some trends in regards to protein function. Proteins playing a role in protein metabolism and folding, cellular trafficking, and RNA processing are well represented, such as components of the 26S proteasome, chaperone proteins, structural components of the ribosome, pre-mRNA splicing factors, and motor proteins such as kinesin and myosin. This follows a trend seen in the protein interactome of PLRV virions (DeBlasio et al. 2014). Given its central role in viral replication and localization within viral replication (Schmitt-Keichinger et al. 2017), one might expect GFLV $1E^{Pol}$ to act within these pathways regardless of strain. However, the fact that a

majority of specficially enriched proteins are found in GFLV-GHu APs could be because GFLV-GHu is the only symptomatic isolate of GFLV which we studied. Studies investigating plant systemic responses to virus infection via transcriptomics often find more intense systemic responses to infections with severe viral strains compared to mild strains in terms of both number of differentially expressed genes (DEGs) and fold changes of DEGs (Osterbaan and Fuchs, 2019). We have evidence of this fold-change differential between symptomatic and asymptomatic viral infections since many of the proteins found specifically enriched in APs of V5-tagged GFLV-GHu were not entirely absent from APs of the other GFLV strains (Table 5-5), but merely had low spectral counts. Normalization of spectral counts or protein abundance against total $1E^{Pol}$ abundance within individual biological replicates could help delineate whether differences in protein abundance across AP conditions is due to actual differences in protein abundance during viral infection or merely an artifact of the variation in the efficiency of $1E^{Pol}$ affinity purification.

A similar analysis could determine the specificity of interactions between GFLV-GHu $1E^{Pol}$ and chloroplast-associated proteins. GFLV-GHu APs were found to enrich plastid transcriptionally active 14 and 12, which regulate plastid gene expression along with pentatricopeptide repeat-containing RNA-binding proteins. Given that GFLV-GHu is the only chlorosis-inducing strain included in this analysis, it would make sense that these proteins are indeed specific to GFLV-GHu.

This is not the first study to explore host factors involved in GFLV infection. The interactions which enable cell-to-cell movement of GFLV, mediated by the $2B^{MP}$ movement protein, have been studied extensively. Class XI myosins are required for

cell-to-cell movement of GFLV virions (Amari et al. 2011) and GFLV replication is sensitive to brefeldin A (a disruptor of the COP-associated secretory pathway), suggesting that COP-dependent cellular trafficking plays a role in GFLV replication (Ritzenthaler et al. 2002). It should be noted that a COP1-interactive protein was among the proteins shown to be specifically enriched in GFLV-GHu APs in this study. And though a specific interaction between $2B^{MP}$ and $2C^{CP}$ has not been directly demonstrated, it has been shown that there is a required compatibility between $2B^{MP}$ and $2C^{CP}$ for cell-to-cell movement to occur (Belin et al. 1999). While APs have been performed on $2B^{MP}$ to identify specific interactors (Laporte et al. 2003), it may be worth revisiting such studies through high-resolution mass spectrometry to increase the depth of identified interactors. Apart from $2B^{MP}$, and now $1E^{Pol}$, investigations into the host interaction networks of GFLV proteins remain largely unknown.

My proteomics work is very encouraging. Some putative viral and host protein interaction partners of GFLV-GHu $1E^{Pol}$ were identified by LC-tandem MS although much remains to be done in unpacking the results of this preliminary study. Additional label-free quantification methods could reveal trends that are masked by the reliance of spectral counting on low variance across analytical and biological replicates. For example, relative protein abundance could be calculated by integration of precursor ion peaks. This would enable us to normalize the relative abundance of proteins of interest against the relative abundance of $1E^{Pol}$ in a given biological replicate, or determine if there is a linear correlation between $1E^{Pol}$ abundance and the abundance of proteins of interest. And while ANOVA of spectral counts provided us with insights into some specifically-enriched proteins, statistical analyses tailored to the analysis of

AP-MS data, such as Significance Analysis of INTeractome (SAINT, Choi et al. 2010), may provide more robust statistical support of these early findings. Additionally, beyond the current dataset, an additional replication of the AP-MS experiments will be needed for validation of these results. Lessons learned in preparation of the AP peptides for MS will hopefully aid in the development of a MS protocol better tailored to these types of samples, along with improvements to the AP protocol to reduce non-specific contaminants.

REFERENCES

Amari, K., Lerich, A., Schmitt-Keichinger, C., Dolja, V. V, and Ritzenthaler, C. 2011. Tubule-guided cell-to-cell movement of a plant virus requires class XI myosin wotors. PLoS Pathog. 7:1–11

Andret-Link, P., Schmitt-Keichinger, C., Demangeat, G., Komar, V., and Fuchs, M. 2004. The specific transmission of *Grapevine fanleaf virus* by its nematode vector *Xiphinema index* is solely determined by the viral coat protein. Virology 320:12–22

Belin, C., Schmitt, C., Demangeat, G., Komar, V., Pinck, L., and Fuchs, M. 2001. Involvement of RNA2-encoded proteins in the specific transmission of *Grapevine fanleaf virus* by its nematode vector *Xiphinema index*. Virology 291:161–171

Belin, C., Schmitt, C., Gaire, F., Walter, B., Demangeat, G., and Pinck, L. 1999. The nine C-terminal residues of the grapevine fanleaf nepovirus movement protein are critical for systemic virus spread. J. Gen. Virol. 80:1347–1356

Candresse, T., Morch, M. D., and Dunez, J. 1990. Multiple alignment and hierachical clustering of conserved amino acid sequences in the replication-associated proteins of plant RNA viruses. Res. Virol. 141:315–329

Chisholm, J., Zhang, G., Wang, A., and Sanfaçon, H. 2007. Peripheral association of a polyprotein precursor form of the RNA-dependent RNA polymerase of *Tomato ringspot virus* with the membrane-bound viral replication complex. Virology 368:133–144

Choi, H., Larsen, B., Lin, Z.-Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z. S., Tyers, M., Gingras, A.-C., and Nesvizhskii, A. I. 2010. SAINT: probabilistic scoring of affinity purification–mass spectrometry data. Nat. Methods. 8:70

DeBlasio, S. L., Johnson, R. S., Mahoney, J., Karasev, A. V, Gray, S. M., MacCoss, M. J., and Cilia, M. 2014. Insights into the polerovirus–plant interactome revealed by coimmunoprecipitation and mass spectrometry. Mol. Plant-Microbe Interact. 28:467–481

DeBlasio, S. L., Johnson, R., Sweeney, M. M., Karasev, A., Gray, S. M., MacCoss, M. J., and Cilia, M. 2015. Potato leafroll virus structural proteins manipulate overlapping, yet distinct protein interaction networks during infection. Proteomics. 15:2098–2112

Demangeat, G., Hemmer, O., Reinbolt, J., Mayo, M.A., and Fritsch, C. 1992. Virus-specific proteins in cells infected with tomato black ring nepovirus: evidence for protleolytic processing *in vivo.* J Gen Virol. 73: 1609-1614.

Digiaro, M., Elbeaino, T., and Martelli, G. P. 2017. *Grapevine fanleaf virus* and other Old World nepoviruses. Pages 47–82 in: Grapevine Viruses: Molecular Biology, Diagnostics and Management, B. Meng, G.P. Martelli, D.A. Golino, and M. Fuchs, eds. Springer International Publishing, Cham.

Dufresne, P. J., Thivierge, K., Cotton, S., Beauchemin, C., Ide, C., Ubalijoro, E., Laliberté, J. F., and Fortin, M. G. 2008. Heat shock 70 protein interaction with Turnip mosaic virus RNA-dependent RNA polymerase within virus-induced membrane vesicles. Virology 374:217–227

Elbeaino, T., Kiyi, H., Boutarfa, R., Minafra, A., Martelli, G. P., and Digiaro, M. 2014. Phylogenetic and recombination analysis of the homing protein domain of grapevine fanleaf virus (GFLV) isolates associated with “yellow mosaic” and “infectious malformation” syndromes in grapevine. Arch. Virol. 159:2757–2764

Gaire, F., Schmitt, C., Stussi-Garaud, C., Pinck, L., and Ritzenthaler, C. 1999. Protein 2A of grapevine fanleaf nepovirus is implicated in RNA2 replication and colocalizes to the replication site. Virology 264:25–36

Jovel, J., Walker, M., and Sanfaçon, H. 2007. Recovery of *Nicotiana benthamiana* plants from a necrotic response induced by a nepovirus is associated with RNA silencing but not with reduced virus titer. J Virol. 81:122285–122297

Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. 2008. ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics 24:2534–2536

Laporte, C., Vetter, G., Loudes, A.-M., Robinson, D. G., Hillmer, S., Stussi-Garaud, C., and Ritzenthaler, C. 2003. Involvement of the secretory pathway and the cytoskeleton in intracellular targeting and tubule assembly of *Grapevine fanleaf virus* movement protein in tobacco BY-2 cells. Plant C 15:2058–2075

Lundgren, D. H., Hwang, S.-I., Wu, L., and Han, D. K. 2010. Role of spectral counting in quantitative proteomics. Expert Rev. Proteomics 7:39–53

MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26:966–968

Margis, R., Viry, M., Pinck, M., and Pinck, L. 1991. Cloning and *in vitro* characterization of the grapevine fanleaf virus proteinase cistron. Virology 185:779–787

Martin, I. R., Vigne, E., Berthold, F., Komar, V., Lemaire, O., Fuchs, M., and Schmitt-Keichinger, C. 2018. The 50 distal amino acids of the $2A^{HP}$ homing protein of *Grapevine fanleaf virus* elicit a hypersensitive reaction on *Nicotiana occidentalis*. Mol. Plant Pathol. 19:731–743

Nölke, G., Cobanov, P., Uhde-Holzem, K., Reustle, G., Fischer, R., and Schillberg, S. 2009. *Grapevine fanleaf virus* (GFLV)-specific antibodies confer GFLV and *Arabis mosaic virus* (ArMV) resistance in *Nicotiana benthamiana*. Mol. Plant Pathol. 10:41–49

Osterbaan, L. J., Choi, J., Kenney, J., Flasco, M., Vigne, E., Schmitt-Keichinger, C., Rebelo, A. R., Heck, M., and Fuchs, M. 2019a. The identity of a single residue of the RNA-dependent RNA polymerase of grapevine fanleaf virus modulates vein clearing in *Nicotiana benthamiana*. Mol. Plant-Microbe Interact. 32:790–801

Osterbaan, L.J., and Fuchs, M. 2019b. Dynamic interactions between plant viruses and their hosts for symptom development. J. Plant Pathol. https://doi.org/10.1007/s42161-019-00323-5

Osterbaan, L. J., Schmitt-Keichinger, C., Vigne, E., and Fuchs, M. 2018. Optimal systemic grapevine fanleaf virus infection in *Nicotiana benthamiana* following agroinoculation. J. Virol. Methods. 257:16–21

Perkins, D. N., Pappin, D. J. C., Creasy, D. M., and Cottrell, J. S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567

Pinck, M., Reinbolt, J., Loudes, A. M., Le Ret, M., and Pinck, L. 1991. Primary structure and location of the genome-linked protein (VPg) of grapevine fanleaf nepovirus. FEBS Lett. 284:117–119

Ritzenthaler, C., Laporte, C., Gaire, F., Dunoyer, P., Schmitt, C., Duval, S., Piéquet, A., Loudes, A.-M., Rohfritsch, O., Stussi-Garaud, C., and Pfeiffer, P. 2002. Grapevine fanleaf virus replication occurs on endoplasmic reticulum-derived membranes. J. Virol. 76:8808–8819

Schellenberger, P., Andret-Link, P., Schmitt-Keichinger, C., Bergdoll, M., Marmonier, A., Vigne, E., Lemaire, O., Fuchs, M., Demangeat, G., and Ritzenthaler, C. 2010. A stretch of 11 amino acids in the βB- βC loop of the coat protein of *Grapevine fanleaf virus* is essential for transmission by the nematode *Xiphinema index*. J. Virol. 84:7924–7933

Schmitt-Keichinger, C., Hemmer, C., Berthold, F., and Ritzenthaler, C. 2017. Molecular, Cellular, and Structural Biology of *Grapevine fanleaf virus*. Pages 83–107 in: Grapevine Viruses: Molecular Biology, Diagnostics and Management, B. Meng, G.P. Martelli, D.A. Golino, and M. Fuchs, eds. Springer International Publishing, Cham.

Southern, J.A., Young, D.F., Heaney, F., Baumgartner, W.K., and Randall, R.E. 1991. Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with differential biological characteristics. J Gen Virol. 72:1551-1557

Steil, B. P., and Barton, D. J. 2009. Cis-active RNA elements (CREs) and picornavirus RNA replication. Virus Res. 139:240–252

Vigne, E., Gottula, J. W., Schmitt-Keichinger, C., Komar, V., Ackerer, L., Belval, L., Rakotomalala, L., Lemaire, O., Ritzenthaler, C., and Fuchs, M. 2013. A strain-specific segment of the RNA-dependent RNA polymerase of grapevine fanleaf virus determines symptoms in *Nicotiana* species. J. Gen. Virol. 94:2803–2813

Viry, M., Serghini, M. A., Hans, F., Ritzenthaler, C., Pinck, M., and Pinck, L. 1993. Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. J. Gen. Virol. 74:169–174

CHAPTER 6

PERSPECTIVES AND FUTURE DIRECTIONS

PERSPECTIVES ON GRAPEVINE FANLEAF VIRUS AGROINOCULATION IN NICOTIANA BENTHAMIANA

A major focus of my work was to generate and optimize molecular tools for the study of virus-host interactions of grapevine fanleaf virus (GFLV). Through the work presented in Chapter 3, I demonstrated that *Agrobacterium tumefaciens*-mediated inoculation (agroinoculation) is an effective and reliable means of establishing systemic GFLV infection in the herbaceous host *Nicotiana benthamiana.* In Chapter 3, I describe my work to clone cDNAs of both genomic RNAs (RNA1 and RNA2) of GFLV strains F13 and GHu into binary plasmids and optimize parameters of the agroinoculation protocol to maximize systemic infection frequency.

Some factors of the agroinoculation protocol such as the genomic makeup of the inoculum, the temperature at which plants are maintained, and the identity of a co-infiltrated silencing suppressor, have a substantial effect of the frequency of GFLV systemic infection in *N. benthamiana*. Since each cDNA of both RNA1 (1) and RNA2 (2) of both GFLV strains F13 (F) and GHu (G) was cloned individually, I was able to mix-and-match the genomic makeup of the inocula. I found that G1G2 and F1G2 outperformed F1F2, with G1F2 having a middling performance, suggesting that overall GFLV-F13 RNA2 is not well suited for GFLV agroinoculation in *N. benthamiana*. I found that rates of systemic GFLV infection were higher when plants were maintained at 25°C rather than 28°C following agroinoculation. This is consistent with other reports showing that *A. tumefaciens* T-DNA transfer is more

efficient at lower temperatures (Annamalai and Rao 2005; Vaghchhipawala et al. 2011). It is possible that lower temperatures may also reduce plant RNA silencing responses, tilting the scales of host-pathogen interactions in favor of the virus. Finally, co-infiltration with *A. tumefaciens* harboring a construct of the p24 silencing suppressor protein from grapevine leafroll associated virus 2 (GLRaV-2) boosted the GFLV systemic infection rate beyond that of a no silencing suppressor control and those of other common silencing suppressors such a tomato bushy stunt virus P19, cucumber mosaic virus 2b, and potato virus Y helper component-protease (HC-Pro). It is not clear why the silencing suppressor of a phloem-limited virus should be the most beneficial to GFLV, which has no-known tissue tropism. The profound effect of p24 is puzzling because its mode of action does not appear to differ drastically from that of P19, 2b, and HC-Pro (Li et al. 2017). More research would be needed to understand how p24 facilitates GFLV infection in *N. benthamiana*.

Some factors of the agroinoculation protocol do not have an effect on the frequency of GFLV infection. For example, the inclusion of acetosyringone in bacterial culture media did not impact infection frequency. This is at odds with recommended protocols for other agroinoculation systems which found that acetosyringone was necessary for successful infection (Sheikholeslam and Weeks 1987). Acetosyringone is produced by plant cells upon wounding. In a natural infection, this wounding signal stimulates *A. tumefaciens* to enact T-DNA transfer (Stachel et al. 1985). In agroinoculation, this T-DNA transfer is how viral cDNA is delivered to plant cells, enabling the establishment of the viral infection. The infiltration step of agroinoculation inflicts some incidental wounding to *N.*

benthamiana leaves, so it is possible that *N. benthamiana* produces enough endogenous acetosyringone to stimulate effective T-DNA transfer, thus eliminating the need to amend *A. tumefaciens* culture media with exogenous acetosyringone (Godwin et al. 1991). Additionally, the optical density (OD_{600}) of the infiltration culture does not affect GFLV agroinoculation outcomes. However, it should be noted that higher OD_{600} values (e.g. 1.0 and 2.0, as compared to 0.1 and 0.5) produced more consistent rates of systemic infection across independent inoculations. Thus I recommend adjusting infiltration cultures to an OD_{600} of 1.0 to 2.0 for GFLV agroinoculation.

Finally, I showed that the binary plasmids harboring GFLV cDNAs are amenable to direct plasmid mutagenesis. In direct plasmid mutagenesis, the desired alterations (substitutions, insertions, or deletions) are coded into a single pair of mutagenic primers which are then used in PCR with the parental plasmid as the template. Each cycle of amplification produces new linear copies of the plasmid, except that the copies contain the desired alterations. The final product mixture is an excess of altered, linear plasmid copies with a small amount of unaltered template plasmid. The template is removed via *Dpn*I digestion and the altered copies are circularized and transformed into *E. coli* for scale-up and confirmation. Thus, introduction of alterations via direct plasmid mutagenesis requires a single PCR. This drastically reduces the time-frame and at-the-bench-time for the creation of GFLV mutants compared to the legacy mutagenesis techniques of PCR cloning (i.e. restriction enzyme digestion and ligation of compatible fragments) and overlap-extension PCR. This enabled me to perform detailed reverse genetic studies to identify

and characterize a symptom determinant of GFLV-GHu in *N. benthamiana* (Chapter 4, described below).

FUTURE DIRECTIONS FOR GRAPEVINE FANLEAF VIRUS AGROINOCULATION

Further reverse genetics of GFLV in Nicotiana *spp.*

The identification and characterization of the vein clearing symptom determinant of GFLV-GHu is just the tip of the iceberg. The manipulations possible through direct plasmid mutagenesis are impressive; in most cases, the only limitation is one's imagination (funding and time notwithstanding). In my own work, I have successfully mutated up to 90 bp, inserted up to 66 bp, and deleted up to 8000 bp of sequence in single independent mutagenesis reactions. Several outstanding questions about GFLV molecular biology come to mind as excellent candidates for reverse genetic approaches enabled by this GFLV agroinoculation system. Preliminary work in our lab has identified a GFLV fusion protein as a potential silencing suppressor (unpublished results). It would be informative to delete or mutate sites of functional significance within the GFLV cDNAs, then characterize the effect these manipulations have on GFLV infection and/or the ability of the fusion protein to act as a silencing suppressor.

Protein $2A^{HP}$ of GFLV-F13 has been identified as an avirulence factor in *Nicotiana occidentalis* (Martin et al. 2018). Specifically, the induction of a hypersensitive response (HR) was mapped to the C-terminal 50 residues of $2A^{HP}$. These authors observed that of these residues, only 17 are not conserved between strains F13 and GHu and thus would be excellent candidates for targeted mutagenesis

to determine if any are individually responsible for HR induction. There would be some technical difficulties to overcome in this line of work. Firstly, the GFLV agroinoculation protocol has been optimized for establishment of infection in *N. benthamiana* and, to my knowledge, this protocol has not yet been adapted to *N. occidentalis.* However, successful agroinoculation of *N. occidentalis* has been achieved with other plant viruses, such as apple chlorotic leaf spot virus (ACLSV) and potato virus X (Zhang and Jelkmann 2017; Fujita et al. 2018), so it is probable that this system could be adapted for inoculation of *N. occidentalis* with GFLV. Secondly, my work showed that F13 RNA2 does not perform well in agroinoculation of *N. benthamiana* (Chapter 3). Thus establishment of systemic infection with F13 RNA2 mutants could be difficult, though it is currently unknown if F13 RNA2's poor performance in agroinoculation of *N. benthamiana* would also translate to difficulties in *N. occidentalis.* This potential roadblock could be addressed by utilizing a gain-of-function rather than loss-of-function approach. Rather than mutating F13 $2A^{HP}$ and monitoring for loss of HR induction, cumulative mutations could be introduced to GHu $2A^{HP}$ to determine what sequence is sufficient for HR induction.

Agroinoculation of GFLV in grapevine

While the recent advancements in the identification and characterization of GFLV symptom determinants in herbaceous hosts is exciting, the brass ring of GFLV research remains the study of GFLV in its natural host, grapevine. Much effort has been expended on developing rapid, reliable, and cost-effective methods for inoculating grapevine with GFLV and results have been mixed (Valat et al. 2003). Inoculation via heterografting has been successful, but remains a laborious and

technically challenging approach. Inoculation via viruliferous *Xiphinema index* nematodes, the specific vector of GFLV, has also met with some success, but also suffers from a long time frame (9 months) (Valat et al. 2003). As an added drawback, this technique is not feasible for researchers located in regions where *X. index* is not native (such as upstate New York) and thus regulators may be wary of allowing even laboratory colonies of *X. index.*

Previous attempts to adapt an early version of the GFLV agroinoculation system to grapevine were unsuccessful: while GFLV could be detected in inoculated grape leaves at nine days post inoculation, no systemic infection was ever confirmed (Gottula, pg. 103, 2014). Yet grapevines are not impervious to agroinoculation. Other grapevine viruses, such as grapevine Pinot gris virus (GPGV) and grapevine red blotch virus (GRBV) have been successfully established in grapevine via agroinoculation variations, such as agrodrench (i.e. soaking of grape roots in *A. tumefaciens* solution), vacuum infiltration (i.e. whole plantlets are forcibly infiltrated by application and release of epi-atmospheric pressures), and pinprick inoculation (Tarquini et al. 2019; Yepes et al. 2018). Though both GFLV strains F13 and GHu were originally isolated from naturally-infected grapevine, both strains were passage through the herbaceous host *Chenopodium quinoa* prior to cDNA construction and cloning. Like most viruses, nepoviruses such as GFLV exist as quasispecies: populations of closely related individual genomes which exist in a "genomic space" around a theoretical consensus sequence (Naraghi-Arani et al. 2001). Primary passage of F13 and GHu through *C. quinoa* may have inadvertently selected for individuals within the quasispeces which are adapted to replication in herbaceous hosts, possibly at the expense of viral fitness

in grapevine. Thus for adaptation of the GFLV agroinoculation system to grapevine it would be advisable to begin with the direct cloning of GFLV cDNA from a naturally-infected grapevine, without primary passage of the isolate to a herbaceous host. Cloning technology has advanced substantially since GFLV-F13 (Viry et al. 1993) and GFLV-GHu (Vigne et al. 2013) were cloned such that production of full-length cDNAs of GFLV from total RNA extracted from grapevine is highly feasible. In fact, a similar direct-cloning approach was successful for the cloning of papaya ringspot virus and papaya leaf distortion mosaic virus from papaya, ACLSV from apple, and GLRaV-2 from grapevine (Tuo et al. 2017; Zhang and Jelkmann 2017; Kurth et al. 2012).

The agroinoculation protocol itself will likely require substantial optimization for adaptation to grapevine. A recent paper describes the successful agroinoculation of apple seedlings via vacuum infiltration for establishment of ACLSV infection (Zhang and Jelkmann 2017). A similar approach could be utilized for grape seedlings or tissue culture plantlets. The establishment of a population of grapevines infected with engineered GFLV would be useful for direct study, or for use as inoculum for graft-inoculation of specific cultivars of interest. For inoculation of mature grapevines, the agrodrench approach may be worth consideration as this method has been successful for establishment of GPGV in grapevine (Tarquini et al. 2019).

The ability to directly and reliably inoculate grapevine with engineered GFLV mutants would be a major boon to the molecular characterization of GFLV in its natural host. Some specific potential projects utilizing this technique are described below and are presented within the context of the other major findings of this thesis.

PERSPETIVES ON THE VEIN CLEARING SYMPTOM DETERMINANT OF GFLV STRAIN GHU

Through the work presented in Chapter 4, I demonstrated that vein clearing caused by GFLV strain GHu in *N. benthamiana* is modulated by residue 802 of protein $1E^{Pol}$. I found that the native lysine (K) produces wild type vein clearing within a GFLV-GHu RNA1 background and that these symptoms are phenocopied by substitution of lysine with arginine, another basic residue (Appendix I, unpublished results). Symptoms are abolished when lysine is replaced by glycine, alanine, or glutamine. Weak symptoms (faint vein clearing which fades rapidly compared to a wild type GFLV-GHu control) are produced when lysine 802 is replaced by proline or asparagine (Chapter 4), or by serine or threonine (Appendix I, unpublished results). Mutation of lysine 802 to glutamic acid generally produced asymptomatic infections: over the course of five independent inoculations, ten plants became infected with GHu K802E, nine of which were asymptomatic. However, in one inoculation, a single *N. benthamiana* plant infected with GHu K802E displayed vein clearing consistent with wild-type GFLV-GHu and the K802E mutation was found to be retained in the symptomatic plant. This led me to label GHu K802E as producing "inconsistent" symptoms rather than fully asymptomatic (Chapter 4). Follow-up work on the symptomatic GHu K802E-infected plant revealed that symptoms were likely due to a reversion of the K802E mutation to the wild type lysine codon which went undetected in the initial symptomatic GHu K802E-infected plant (Appendix I, unpublished results). It should be noted that lysine 802 is not sufficient for vein clearing since a

GFLV-F13 mutant in which the native glycine was mutated to lysine (F13 G802K) failed to produce vein clearing in *N. benthamiana.*

I also found that vein clearing development was not correlated with overall virus titer, as determined by RT-qPCR of GFLV RNA1. The comparison of silent and nonsynonymous mutations to codon 802 allowed me to determine that symptoms were dependent on the protein sequence, rather than the nucleotide sequence, of $1E^{Pol}$. These results were further bolstered by the finding that both GFLV and tobacco rattle virus (TRV) virus-induced gene silencing (VIGS) vectors harboring portions of the $1E^{Pol}$ coding sequence containing codon 802 failed to induce vein clearing in *N. benthamiana.*

I followed up these *in planta* studies with *in silico* characterization of the $1E^{Pol}$ C-terminus. Prediction of the tertiary structure of the $1E^{Pol}$ C-terminus proved unreliable due to the lack of solved crystal structures for proteins with sequence homology to this region. Note that the RNA-dependent RNA polymerase (RdRp) "core" of $1E^{Pol}$ could be predicted with good confidence since the crystal structures of several viral RdRps, such as the RdRp of poliovirus, have been solved (Hansen et al. 1997). Prediction of the secondary structure of $1E^{Pol}$ yielded some interesting results. I found that residue 802 is flanked by conserved stable secondary structures. Downstream of residue 802 is a strongly predicted alpha-helix while upstream is a string of residues which is predicted to be highly ordered across a broad range of sequenced GFLV isolates. Such features suggest this region may be amenable to crystallization, a prerequisite for structure determination via X-ray crystallography.

FUTURE DIRECTIONS FOR THE CHARACTERIZATION OF THE GFLV VEIN CLEARING SYMPTOM DETERMINANT

In the quest to delineate the mechanism of vein clearing symptom production by GFLV-GHu in *N. benthamiana,* the first logical step was the identification of protein interaction partners of $1E^{Pol}$ by affinity purification and mass spectrometry (AP-MS). The preliminary results of that approach are reported in Chapter 5 and future directions evolving from that study are detailed below. However, there remain some unanswered questions about the GFLV-GHu symptom determinant which AP-MS cannot address directly.

Firstly, while lysine 802 has been found (along with arginine) to be necessary for vein clearing, the minimal sufficient symptom determinant remains unknown. Since lysine 802 is not sufficient on its own to produce vein clearing within an F13 RNA1 background, there must be some other component within the C-terminus of GFLV-GHu $1E^{Pol}$ which is necessary for symptom development. In their original characterization of the GFLV-GHu symptom determinant, Vigne et al. (2013) found that residues 689 to 824 (nucleotides 2065-2472) of $1E^{Pol}$ were sufficient for vein clearing in *N. benthamiana* (and for chlorotic spots in *Nicotiana clevelandii),* suggesting that this "other component" likely lies within the ultimate 136 residues of the $1E^{Pol}$ protein. While it may be feasible to identify the minimal sufficient symptom determinant by making cumulative mutations to GFLV-F13 RNA1 until a symptomatic variant is found, it would be helpful to have additional insights into the nature of the $1E^{Pol}$ C-terminus to narrow the scope of such a fishing expedition.

For example, determination of the tertiary structure of the $1E^{Pol}$ C-terminus could be enormously helpful in identifying regions of $1E^{Pol}$ which may interact with residue 802. This could be achieved by X-ray crystallography of the full $1E^{Pol}$ protein, or, if the full protein proves recalcitrant to crystallization, of the C-terminus alone. A similar truncation approach to crystallization of a plant viral protein was recently utilized to solve the structure of the N-terminus of the potato leafroll virus read through protein (Dr. Michelle Heck, personal communication). In addition to providing insights on the minimal sufficient symptom determinant, knowing the tertiary structure of the $1E^{Pol}$ C-terminus would enhance interpretation of the protein interactome data generated via AP-MS. If solved structures are available for any of the putative interactors, *in silico* docking simulations could be run to determine the three dimensional geometry of the interactions. This could go a long way in elucidating the mechanism by which $1E^{Pol}$ interacts with host components to trigger vein clearing in *N. benthamiana.*

It could also be worthwhile to characterize some of the GFLV mutants produced in Chapter 4 and Appendix I in *N. clevelandii.* Archived tissue (i.e. whole leaves stored at -80°C) is available for the vast majority of these mutants such that they could be mechanically inoculated to *N. clevelandii* with relative ease. If it is found that lysine 802 acts as a symptom determinant for chlorotic spots in *N. clevelandii*, then this host could prove a valuable point of comparison for results obtained with *N. benthamiana*. One could envision running assays such as AP-MS, transcriptomics, and GFLV quantification in parallel in both *N. benthamiana* and *N.*

clevelandii and comparing the data to determine which results are consistent across different hosts.

PERSPECTIVES ON THE PRELIMARY IDENTIFICATION OF 1E^{POL} INTERACTION PARTNERS

Through the work presented in Chapter 5, I was able to lay the groundwork for mapping the protein interactome of GFLV 1E^{Pol} via affinity purification and mass spectrometry (AP-MS) and completed a preliminary survey of the 1E^{Pol} protein interactome. A major step forward for this line of inquiry was the design and creation of infectious GFLV constructs containing epitope-tagged 1E^{Pol}. Protein 1E^{Pol} was initially recalcitrant to epitope-tagging, as constructs harboring FLAG, myc, or HA tags were unable to establish systemic infection in *N. benthamiana* via agroinoculation. GFLV-GHu variants tagged with epitope V5, derived from the P and V proteins of simian virus 5, were systemically infectious in *N. benthamiana* and V5-tagged 1E^{Pol} could be detected in total soluble proteins from *N. benthamiana* via western blotting using a commercially available anti-V5 antibody (Invitrogen PA1-993). Tagging GFLV-F13 was trickier and I found that a one amino acid truncation of the 1E^{Pol} C-terminus was necessary to rescue systemic infectivity of the V5-tagged GFLV-F13 construct.

Even with systemically infectious V5-tagged GFLV constructs established, additional optimization was required prior to AP. I found that a reducing agent, such as dithiothreitol or β-mercaptoethanol, was required in the lysis buffer used for protein extraction in order for 1E^{Pol} to remain in the soluble fraction. I also found that 1E^{Pol} titer appears to correlate with GFLV RNA1 titer over the early days of infection (three

to thirteen days post mechanical inoculation) and that initial detection of $1E^{Pol}$ and GFLV RNA1 by western blotting and RT-qPCR, respectively, corresponds to the onset of vein clearing for the symptomatic construct GHu-1E:V5. I found that $1E^{Pol}$ detection by western blotting was strongest between seven and eight days post mechanical inoculation.

While I was able to reliably and confidently detect V5-tagged $1E^{Pol}$ within *N. benthamiana* total soluble proteins using a polyclonal anti-V5 antibody, there were consistent non-specific bands from both virus-infected and healthy *N. benthamiana* tissue. Non-specific binding of host or viral proteins to the antibody during AP can mask or eliminate low-abundance, specific interactors from the AP products, decreasing the depth to which the protein interactome is detected. To reduce the non-specific background binding of anti-V5 antibody species to host and viral proteins, the polyclonal anti-V5 antibody used for AP was cross-absorbed against a blend of *N. benthamiana* tissue infected with untagged versions of the GFLV variants used for AP.

AP using the cross-absorbed anti-V5 antibody conjugated to Protein A Dynabeads was carried out on three biological replicates (pools of systemically infected *N. benthamiana* tissue) for each of the following GFLV variants: GHu-1E:V5 (symptomatic strain), F13-1E:$V5_{C\text{-}1}$ (asymptomatic strain), GHu-1E-K802G:V5 (an asymptomatic mutant of GHu), wild type (untagged) GFLV-GHu, and wild type (untagged) GFLV-F13. Tryptic peptides from each AP were analyzed by tandem mass spectrometry.

I used label-free quantification by spectral counting to identify proteins which were specifically enriched in particular AP conditions. GFLV protein $1D^{Pro}$, the viral protease, is enriched in $1E^{Pol}$ APs, as is protein $1B^{Hel}$, the viral helicase. GFLV proteins $2B^{MP}$ and $2C^{CP}$ were non-specifically purified in all APs (including negative controls), suggesting that these viral proteins non-specifically bind to the antibody and/or beads. A number of host proteins were shown to be specifically enriched in V5-tagged GFLV-GHu $1E^{Pol}$ APs, including several involved in chloroplast function. The functional significance of these putative interactions will require both validation of the identified interactors by replicate APs and characterization of the effect which perturbation of these interactors has on $1E^{Pol}$ function and/or GFLV infection.

FUTURE DIRECTIONS FOR GFLV PROTEOMICS

Further characterization of the GFLV $1E^{Pol}$ interactome in N. benthamiana

The data presented in Chapter 5 is preliminary and will require at least one round of replication before it will be ready for full peer-review and publication. In the initial AP experiment, I included GFLV variants which produced wild-type vein clearing (GFLV-GHu) or were asymptomatic (GFLV-F13 and GFLV-GHu-1E-K802G). The amount of anti-V5 antibody available at the time of the experiment was a limiting factor, forcing me to leave out GHu-1E-K802P as one of treatments. GHu-1E-K802P produces faint vein clearing which fades more rapidly than vein clearing produced by wild type GFLV-GHu. If materials and funding allow, I would include GHu-1E-K802P in the next round of AP experiments in order to have data from all three “classes” of vein clearing produced by GFLV variants.

There were significant issues in optimizing preparation of tryptic peptides from these APs for mass spectrometry analysis. Aliquots of 20% and 5% of the prepared peptides clogged the liquid chromatography column used for resolution of the peptides prior to mass spectrometry; mass spectrometry was eventually performed on 2% aliquots of the samples (K. Rivera, personal communication). It remains unknown what caused this clogging issue, though the use of a pre-trap to eliminate most non-peptide molecules from the mass spectrometry analysis may resolve this issue.

Non-specific binding of proteins to antibodies and/or beads can make it difficult to confidently detect low-abundance protein interactions in AP. Cross-absorption against representative protein extractions lacking the bait protein can reduce the background binding of antibody species to non-baits, as was implemented for the APs presented in this work. However, this does not appear to have completely eliminated non-specific binding of GFLV proteins $2B^{MP}$ and $2C^{CP}$. It is possible that these proteins non-specifically bind the Protein A Dynabeads themselves. Pre-clearing of tissue lysates with non-IgG coated beads prior to AP could eliminate some of this non-specific binding.

Of course, even with sufficient replication and additional optimization, AP-MS is merely a starting point in identifying protein-protein interactions. Putative interactors of $1E^{Pol}$ will need validation. This could be achieved through yeast two-hybrid assays and/or VIGS of putative interactors to determine what effect(s) silencing these host genes have on GFLV infection or symptom production in *N. benthamiana.* Putative interactions could also be probed by protein interaction reporter (PIR) technology. In PIR, protein mixtures (such as proteins extracted from GFLV-infected

tissue) are treated with a chemical cross-linker which contains an MS-labile bridge connected to a mass-encoded tag. Reactive residues (such as lysine) within a defined molecular distance of one another become cross-linked and these linked residues (and the proteins they come from) are then identified via mass spectrometry. This approach could be particularly informative in the context of the GFLV-GHu symptom determinant since residue 802 is a lysine and thus would be one of the mapped contact points within the data set.

Proteomics of GFLV 1E^{Pol} in grapevine

As mentioned above, while using herbaceous hosts for the molecular characterization of GFLV has allowed us to make great advancements that would be otherwise impossible given the technical challenges of working with grapevines, the ultimate goal should be to apply and replicate these findings in the natural host. Given that a preliminary dataset of GFLV 1E^{Pol} protein interactors in *N. benthamiana* is already available (Chapter 5), it would be logical to pursue a similar dataset from grapevine. Comparisons of these datasets could allow for the identification of conserved interactions (likely essential to virus replication) and those interactions which are host-specific.

The ability to rapidly and reliably inoculate grapevine (such as via agroinoculation) with the GFLV constructs utilized in Chapter 5 would certainly speed this process. I discuss above some strategies that could help finally "crack the code" of GFLV agroinoculation in grapevine. But even as the GFLV research community eagerly awaits progress on that front, steps could be made now to move forward on proteomics of GFLV in grapevine. While never able to achieve systemic infection,

Gottula's attempts to design a GFLV agroinoculation system did confirm GFLV infection at up to 9 days post inoculation in agroinfiltrated leaves of grapevine (Gottula, pg. 103, 2014). Even locally-infected tissue can yield useful data on protein-protein interactions, as was done for virions of potato leafroll virus in *N. benthamiana* (DeBlasio et al. 2014). At the very least, locally-infected tissue could be used to optimize lysis buffer, protein extraction, and AP conditions for AP of $1E^{Pol}$ from grapevine in preparation for AP from systemically infected grapevine. And as a contingency plan against GFLV agroinoculation continuing to prove intractable, work could begin now on transferring, at minimum, GHu-1E:V5 and F13-1E:$V5_{C\text{-}1}$ (and their untagged parental constructs) to grapevine via heterologous grafting for the production of grapevine systemically infected with V5-tagged GFLV. Actually, efforts to transmit GHu, GHu-K802K, GHu K802P, and GHu-K802G to grapevine via heterologous grafting is underway in collaboration with Emmanuelle Vigne at INRA in Colmar, France.

An additional contingency could be the use of antibody serum raised against a synthetic peptide of $1E^{Pol}$, as described in Vigne et al. (2013). While my own attempts to utilize such an anti-1E serum for detection of $1E^{Pol}$ in *N. benthamiana* were unsuccessful, thus necessitating the creation and use of the V5-tagged $1E^{Pol}$ constructs described in Chapter 5, it may be worth revisiting this approach. My own experience with the anti-1E serum was that it suffered from high background (non-specific binding), making it difficult to confidently detect $1E^{Pol}$ (unpublished results). The cross-absorption protocol described in Chapter 5 to reduce non-specific binding of the polyclonal anti-V5 antibody could easily be applied to an anti-1E serum and could

render the serum more useful for the detection (and subsequent AP) of untagged $1E^{Pol}$. In such a case, most any GFLV-infected grapevine currently available could be used for AP of $1E^{Pol}$ without the need to infect grapevines with engineered GFLV constructs.

NEW TOOLS FOR GFLV MANAGEMENT IN GRAPEVINE

Given its worldwide distribution and the severity of its impacts on grapevine vigor and productivity, GFLV is considered one of the most adverse grapevine viruses. Thus there has been great interest, and effort, in developing strategies to control the spread and effects of GFLV.

Early pathogen-derived resistance efforts focused on identifying attenuated strains of GFLV for use in cross-protection (Huss et al. 1989), whereby vines are purposefully inoculated with a weakened or asymptomatic strain to prevent subsequent infection by virulent strains. Work is underway to transfer some of the asymptomatic variants of GFLV described in chapter 4 to grapevine via heterografting (Emmanuelle Vigne, personal communication). If any of these variants are found to be attenuated or asymptomatic in grapevine, it would be interesting to test their ability to provide cross-protection in grapevine against more severe isolates of GFLV.

To date, no natural resistance to GFLV has been found in grapevine and thus most efforts have focused on the use of transgenic approaches to develop novel sources of resistance (Oliver and Fuchs 2011). Since GFLV is nematode-vectored, field resistance should be achievable by the development of a GFLV-resistant rootstock without the need to alter the nigh sacred scion. While numerous projects have produced transformed grapevines harboring portions of the GFLV genome

(Fuchs and Lemaire 2017), only a few have been field-tested. And while field tests showed promising results (Vigne et al. 2004), the experiments were dogged by public-relations woes, culminating in vandalism of the test plots and abandonment of the project in France (Fuchs and Lemaire 2017). This has not appeared to diminish the interest of grape growers in genetically-modified GFLV-resistant grapevines, particularly in the US (Marc Fuchs, personal communication).

Genome-editing technologies (most notably CRISPR-Cas9) which allow for "clean" edits (i.e. no introduction of foreign genetic material) of genomes, are gaining considerable attention. However, the deployment of such technologies for the production of GFLV-resistance grapevines will require additional insights into the virus-host interactions that occur during infection in order to identify promising host candidate genes for editing. While remarkable progress has been made in detailing the molecular events of GFLV infection (for review, see Schmitt-Keichinger et al. 2017), systems biology techniques such as transcriptomics (Hily et al. 2018) and proteomics (Chapter 5) should provide further insights. Grapevine homologs of $1E^{Pol}$ interactors identified in Chapter 5 could be excellent candidates for genome-editing and expanding GFLV proteomics to grapevine will only further focus such efforts.

The future of GFLV molecular biology research is bright. The tools and techniques reported here provide new and improved tools for reverse genetics and systems biology approaches to fill in the mechanistic gaps in our understanding of GFLV symptomatology. The next great challenge will be to efficiently and vigorously apply these approaches to grapevine to map new avenues towards durable and effective GFLV management in the field.

REFERENCES

Annamalai, P., and Rao, A. L. N. 2005. Delivery and expression of functional viral RNA genomes *in planta* by agroinfiltration. in: Current Protocols in Microbiology, John Wiley & Sons, Inc.

DeBlasio, S. L., Johnson, R. S., Mahoney, J., Karasev, A. V, Gray, S. M., MacCoss, M. J., and Cilia, M. 2014. Insights into the polerovirus–plant interactome revealed by coimmunoprecipitation and mass spectrometry. Mol. Plant-Microbe Interact. 28:467–481

Fuchs, M., and Lemaire, O. 2017. Novel approaches for viral disease management. Pages 599–621 in: Grapevine Viruses: Molecular Biology, Diagnostics and Management, B. Meng, G.P. Martelli, D.A. Golino, and M. Fuchs, eds. Springer International Publishing, Cham.

Fujita, N., Komatsu, K., Ayukawa, Y., Matsuo, Y., Hashimoto, M., Netsu, O., Teraoka, T., Yamaji, Y., Namba, S., and Arie, T. 2018. N-terminal region of cysteine-rich protein (CRP) in carlaviruses is involved in the determination of symptom types. Mol. Plant Pathol. 19:180–190

Godwin, I., Todd, G., Ford-Lloyd, B., and Newbury, H. J. 1991. The effects of acetosyringone and pH on *Agrobacterium*-mediated transformation vary according to plant species. Plant Cell Rep. 9:671–675

Gottula, J. W. 2014. Grapevine fanleaf virus: biology, biotechnology and resistance. Ph.D. Dissertation. Cornell University

Hansen, J. L., Long, A. M., and Schultz, S. C. 1997. Structure of the RNA-dependent RNA polymerase of poliovirus. Structure 5:1109–1122

Hily, J.-M., Demanèche, S., Poulicard, N., Tannières, M., Djennane, S., Beuve, M., Vigne, E., Demangeat, G., Komar, V., Gertz, C., Marmonier, A., Hemmer, C., Vigneron, S., Marais, A., Candresse, T., Simonet, P., and Lemaire, O. 2018. Metagenomic-based impact study of transgenic grapevine rootstock on its associated virome and soil bacteriome. Plant Biotechnol. J. 16:208–220

Huss, B., Walter, B., and Fuchs, M. 1989. Cross-protection between arabis mosaic virus and grapevine fan leaf virus isolates in *Chenopodium quinoa*. Ann. Appl. Biol. 114:45–60

Kurth, E. G., Peremyslov, V. V, Prokhnevsky, A. I., Kasschau, K. D., Miller, M., Carrington, J. C., and Dolja, V. V. 2012. Virus-derived gene expression and RNA interference vector for grapevine. J. Virol. 86:6002–6009

Li, M., Zhang, J., Feng, M., Wang, X., Luo, C., Wang, Q., and Cheng, Y. 2017. Characterization of silencing suppressor p24 of *Grapevine leafroll-associated virus 2*. Mol. Plant Pathol. 19:355-368

Martin, I. R., Vigne, E., Berthold, F., Komar, V., Lemaire, O., Fuchs, M., and Schmitt-Keichinger, C. 2018. The 50 distal amino acids of the $2A^{HP}$ homing protein of *Grapevine fanleaf virus* elicit a hypersensitive reaction on *Nicotiana occidentalis*. Mol. Plant Pathol. 19:731–743

Naraghi-Arani, P., Daubert, S., and Rowhani, A. 2001. Quasispecies nature of the genome of *Grapevine fanleaf virus*. J. Gen. Virol. 82:1791–1795

Oliver, J. E., and Fuchs, M. 2011. Tolerance and resistance to viruses and their vectors in Vitis sp.: a virologist's perspective of the literature. Am. J. Enol. Vitic. 62:438–451

Schmitt-Keichinger, C., Hemmer, C., Berthold, F., and Ritzenthaler, C. 2017. Molecular, Cellular, and Structural Biology of *Grapevine fanleaf virus*. Pages 83–107 in: Grapevine Viruses: Molecular Biology, Diagnostics and Management, B. Meng, G.P. Martelli, D.A. Golino, and M. Fuchs, eds. Springer International Publishing, Cham.

Sheikholeslam, S. N., and Weeks, D. P. 1987. Acetosyringone promotes high efficiency transformation of *Arabidopsis thaliana* explants by *Agrobacterium tumefaciens*. Plant Mol. Biol. 8:291–298

Stachel, S. E., Messens, E., Van Montagu, M., and Zambryski, P. 1985. Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. Nature 318:624–629

Tarquini, G., Zaina, G., Ermacora, P., De Amicis, F., Franco-Orozco, B., Loi, N., Martini, M., Bianchi, G. L., Pagliari, L., Firrao, G., de Paoli, E., and Musetti, R. 2019. Agroinoculation of *Grapevine Pinot gris virus* in tobacco and grapevine provides insights on viral pathogenesis. PLoS One. 14:e0214010

Tuo, D., Fu, L., Shen, W., Li, X., Zhou, P., and Yan, P. 2017. Generation of stable infectious clones of plant viruses by using *Rhizobium radiobacter* for both cloning and inoculation. Virology 510:99–103

Vaghchhipawala, Z., Rojas, C. M., Senthil-Kumar, M., and Mysore, K. S. 2011. Agroinoculation and agroinfiltration: simple tools for complex gene function analyses. Pages 65–76 in: Plant Reverse Genetics: Methods and Protocols, A. Pereira, ed. Humana Press, Totowa, NJ.

Valat, L., Burrus, M., Fuchs, M., and Mauro, M. C. 2003. Review of techniques to inoculate grapevines with Grapevine fanleaf virus: lessons and perspectives. Am. J. Enol. Vitic. 54:279–285

Vigne, E., Gottula, J. W., Schmitt-Keichinger, C., Komar, V., Ackerer, L., Belval, L., Rakotomalala, L., Lemaire, O., Ritzenthaler, C., and Fuchs, M. 2013. A strain-specific segment of the RNA-dependent RNA polymerase of grapevine fanleaf virus determines symptoms in *Nicotiana* species. J. Gen. Virol. 94:2803–2813

Vigne, E., Komar, V., and Fuchs, M. 2004. Field safety assessment of recombination in transgenic grapevines expressing the coat protein gene of *Grapevine fanleaf virus*. Transgenic Res. 13:165–179

Viry, M., Serghini, M. A., Hans, F., Ritzenthaler, C., Pinck, M., and Pinck, L. 1993. Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. J. Gen. Virol. 74:169–174

Yepes, L. M., Cieniewicz, E., Krenz, B., McLane, H., Thompson, J. R., Perry, K. L., and Fuchs, M. 2018. Causative role of grapevine red blotch virus in red blotch disease. Phytopathology 108:902–909

Zhang, L., and Jelkmann, W. 2017. Construction of full-length infectious cDNA clones of *Apple chlorotic leaf spot virus* and their agroinoculation to woody plants by a novel method of vacuum infiltration. Plant Dis. 101:2110–2115

APPENDIX

ADDITIONAL GRAPEVINE FANLEAF VIRUS MUTANTS AND CONSTRUCTS OF INTEREST

ABSTRACT

While the "file drawer problem" is usually framed within the context of a publication bias that favors statistically significant results over those that fail to reject the null hypothesis, one could argue that the problem extends even to studies which do not make extensive use of statistics. For every radical idea and exciting new finding that makes it to press, there are dozens of projects that got abandoned early-on due to technical difficulties, discouraging preliminary results, unexpected outcomes that are challenging to interpret, or simply a lack of time and resources to bring the project to completion. While such projects are unlikely to see publication in the peer-reviewed literature, it seems a disservice to the scientific community to leave these results confined to the file drawer (or, more realistically these days, to a twelfth tier subfolder on one's hard drive labeled "Other Stuff"). Most of my research is presented in the preceding chapters. What follows is the research that did not yield anticipated results or which did not fit into the central scope of my thesis work.

ADDITIONAL GRAPEVINE FANLEAF VIRUS 1E^{POL} MUTANTS

A note about the term "local expression"

When I first started engineering and characterizing GFLV 1E^{Pol} mutants, my workflow consisted of agroinoculating *N. benthamiana* with my designed constructs, then waiting 2 weeks until testing the plants for systemic infection by DAS-ELISA. In much of my early work (before the GFLV agroinoculation protocol was fully optimized), it was not unusual for entire agroinoculation experiments to fail, with even the wild type controls failing to establish systemic infection. In an effort to track down the cause of these failures, I decided to see if I could actually detect the virus in the inoculated leaves at 3 to 4 days post inoculation. My reasoning was that a positive virus signal in the inoculated leaves should indicate that the viral constructs were actually being expressed. A negative signal could indicate that something had gone wrong with the construct or the *A. tumefaciens* culture itself to prevent virus expression.

I found that most constructs were positive for GFLV by DAS-ELISA in inoculated leaves at 3 to 4 days post inoculation, though this was not a guarantee that it would produce a systemic infection. In contrast, some constructs never produced a positive GFLV signal in inoculated leaves; such constructs always failed to establish a systemic infection.

I began to call constructs which produced a positive GFLV signal in inoculated leaves at 3 to 4 days post inoculation as "locally infectious" because I interpreted the positive GFLV signal to indicate that the virus was actively replicating in the inoculated leaves. However, it was later pointed out to me that expression of GFLV in

the inoculated leaves could be driven directly by the cauliflower mosaic virus (CaMV) 35S expression cassette to a high enough level to produce a positive GFLV signal in double-antibody sandwich (DAS) enzyme-linked immunosorbent assay (ELISA), without any viral replication. Though suggestions were made for ways to distinguish between 35S-driven expression and viral replication, I did not have the time to pursue them. Thus, to more accurately encompass both hypotheses, constructs as described above have been termed here as capable of "local expression."

Additional mutants targeting GFLV-GHu 1E^{Pol} residue 802

Chapter 4 describes my work to characterize the GFLV-GHu symptom determinant for vein clearing in *N. benthamiana*, which was identified as residue 802 of protein 1E^{Pol}. At the time of publication of that chapter as a peer-reviewed journal article, I had only completed characterization of a handful of mutants, specifically mutations of GFLV-GHu 1E^{Pol} lysine 802 to glycine, alanine, and glutamine (asymptomatic); proline and asparagine (weakly symptomatic); glutamic acid (inconsistent symptoms); and arginine and tyrosine (non-infectious via agroinoculation). Of the twenty standard amino acids, this left eleven amino acids remaining to be characterized for their effects on vein clearing in *N. benthamiana* (Table A1). This work was intended to investigate whether amino acids with similar physicochemical properties would generate mutants that were phenotypically identical.

Table A1. Grapevine fanleaf virus (GFLV) strain GHu mutants targeting residue 802 of the RNA-dependent RNA polymerase ($1E^{Pol}$).

Construct name	NT	AA	Local expression[a]	Systemic infection	Symptoms [b]
GHu (wt)	AAG	K	+	+	+
GHu K802R [AGA]	AGA	R	N/T	+	+
GHu K802S	AGT	S	N/T	+	Weak
GHu K802T	ACC	T	N/T	+	Weak
GHu K802R [CGT][c]	CGT	R	-	-	N/A
GHu K802Y [TAC][c]	TAC	Y	+	-	N/A
GHu K802C	TGT	C	N/T	-	N/A
GHu K802D	GAT	D	N/T	-	N/A
GHu K802F	TTT	F	N/T	-	N/A
GHu K802H [CAC]	CAC	H	N/T	-	N/A
GHu K802H [CAT]	CAT	H	N/T	-	N/A
GHu K802I	ATT	I	N/T	-	N/A
GHu K802M	ATG	M	N/T	-	N/A
GHu K802Y [TAT]	TAT	Y	N/T	-	N/A
GHu K802E (revertant)[d]	GAG (AAG)	E (K)	N/T	+	- (+)
GHu K802W	TGG	W	Cloning failed[e]	N/A	N/A

[a]"+" indicates that agroinoculated leaves collected 3 to 4 days post inoculation tested positive for GFLV by double antibody sandwich enzyme-linked immunsorbant assay (DAS-ELISA). "-" indicates that agroinoculated leaves collected 3 to 4 days post inoculation tested negative for GFLV by DAS-ELISA. "N/T" indicates that the construct was not tested for local infection.

[b]"+" indicates plants infected with the construct displayed distinct vein clearing on upper, non-inoculated leaves. "-" indicates plants infected with the construct remained asymptomatic. "Weak" indicates that plants displayed faint vein clearing which faded more rapidly compared to a wild type control. "N/A" indicates non-applicable.

[c]Both of these constructs are included in Table 4-6 and are also reported in Chapter 4. They are included here for comparison to GHu K802R [AGA] and GHu K802Y [TAT].

[d]This construct produced a single revertant, with the results of the revertant form shown in parenthesis.

[e]Despite repeated attempts, I was unable to produce a successfully mutated plasmid containing the TGG codon at position 802.

Of these additional mutants, the most exciting was GHu K802R [AGA], in which the native lysine was substituted by arginine. Both are basic residues. This mutant produced distinct vein clearing symptoms which were indistinguishable from those of a wild type GFLV-GHu control (Figure A1; Table A1). It is interesting to note that I had previously characterized an arginine mutant, GHu K802R [CGT], which was unable to establish systemic infection via agroinoculation in *N. benthamiana* (Chapter 4). However, it should be noted that GHu K802R [CGT] was not locally expressed in inoculated leaves (in two independent agroinoculations). This suggests that this version of the virus was unable to replicate at all or that there was some error in the plasmid or issue with the *A. tumefaciens* culture to prevent expression of the construct.

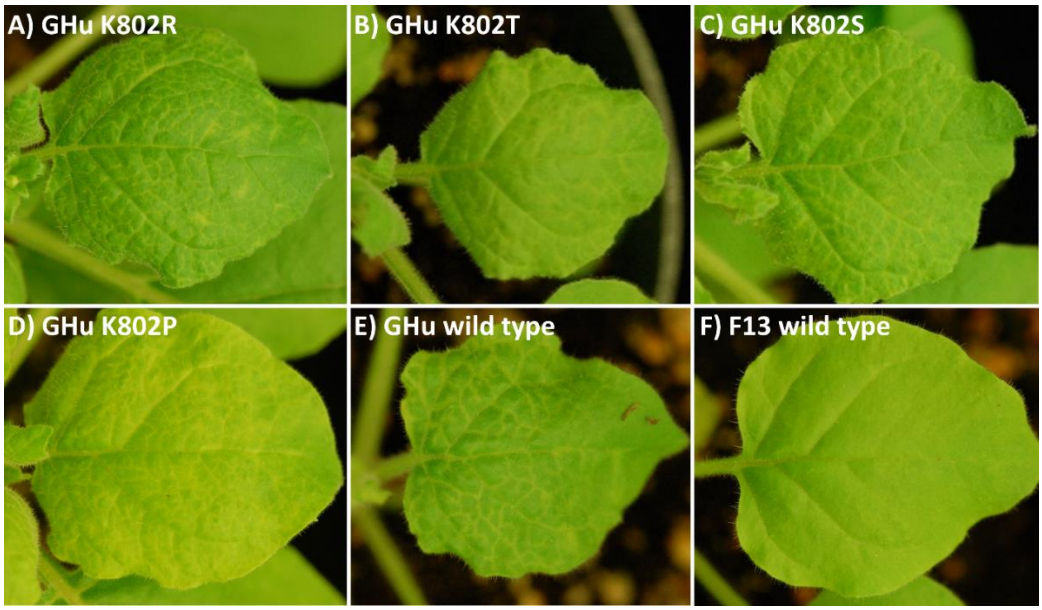


Figure A1. Grapevine fanleaf virus (GFLV) strains and mutants and their behavior in *Nicotiana benthamiana*

Mutants GHu K802S and GHu K802T are also noteworthy in that they produced weak vein clearing symptoms which faded more rapidly than a wild type

GFLV-GHu control (Figure A1; Table A1). This is identical to the symptom phenotypes of GHu K802P and GHu K802N (Chapter 4). Serine, threonine, and asparagine are all polar residues, while proline is a non-polar amino acid with a unique structure that constrains the peptide backbone.

Mutants featuring most of the remaining standard amino acids were all unable to establish systemic infection in *N. benthamiana* via agroinoculation: cysteine, aspartic acid, phenylalanine, histidine (codons CAC and CAT), isoleucine, methionine, and tyrosine (Table A1). Mutation to tyrosine (GHu K802Y [TAC]) had been previously attempted (Chapter 4) and this construct was found to be locally expressed but unable to establish systemic infection via agroinoculation. An additional tyrosine mutant utilizing a different codon (GHu K802Y [TAT]) was unable to establish systemic infection, though its local expression was not determined (Table A1). Despite multiple attempts, I was unable to produce a successfully mutated plasmid containing "TGG" (which codes for tryptophan) at codon 802 and thus was unable to determine the effects of substituting lysine 802 for tryptophan within GFLV-GHu $1E^{Pol}$. Mutants featuring either valine or leucine (both small, non-polar amino acids) were not included in the experiment design.

I also have data to follow up the story of GHu K802E. In Chapter 4, GHu K802E is listed as "inconsistent" (Table 4-6). This is because in the majority of inoculations, GHu K802E produced asymptomatic infections. However, in one mechanical inoculation, a single plant produced typical vein clearing, while the four remaining systemically infected plants were asymptomatic. I sequenced viral progeny from all four plants and found the K802E mutation to be conserved in each plant. I did

not have time to further characterize this construct before the submission of Chapter 4 to peer review. However, I later mechanically passaged archived tissue (whole leaves frozen at -80°C) from both symptomatic and asymptomatic plants of GHu K802E. All of the plants inoculated with the asymptomatic isolate of GHu K802E became systemically infectious and remained asymptomatic. In contrast, all of the plants inoculated with the symptomatic isolate of GHu K802E produced vein clearing symptoms indistinguishable from a GFLV-GHu wild type control. Sequencing of the viral progeny revealed that the K802E mutation was retained in the asymptomatic infections. However, the symptomatic plants contained viral progeny in which the $1E^{Pol}$ 802 codon had reverted from GAG (glutamic acid, E) to the wild type AAG (lysine, K). It is interesting that this is the only case of a revertant that I encountered during the characterization of the GFLV-GHu symptom determinant, especially given that it occurred when an acidic residue had been substituted. The GFLV $1E^{Pol}$ C-terminus is deficient in acidic residues and fusion of highly-acidic epitope tags to the C-terminus render the virus non-infectious (Chapter 5). This reversion could indicate that acidic residues are selected against within the GFLV $1E^{Pol}$ C-terminus.

A working hypothesis on the structure of the $1E^{Pol}$ C-terminus

The results from these additional mutants have allowed me to form a working hypothesis about the structure of the $1E^{Pol}$ C-terminus. Residue 802 is flanked by conserved ordered motifs (Chapter 4). I hypothesize that residue 802 serves as a structural hinge between these two motifs and determines their geometry relative to one another. This geometry in turn modulates a protein-protein interaction with a yet-to-be-identified partner which is necessary for vein clearing.

Lysine and arginine (both basic residues) produce wild type vein clearing symptoms. I hypothesize that basic residues allow the conserved motifs to adopt a geometry that enables the putative protein-protein interaction(s) necessary for vein clearing. Asparagine, serine, and threonine (all polar residues) produce weak vein clearing symptoms, as does proline (a non-polar, but sterically hindered, residue). I hypothesize that polar residues allow the conserved motifs to adopt a geometry similar to, but not identical to, the native conformation, thus partially disrupting or weakening (though not abolishing) the putative protein-protein interaction and thus trigger some level of symptom expression, though not enough to reproduce wild type vein clearing. Proline, with its rigid constraint of the peptide backbone, may "lock" the conserved motifs into a similar suboptimal conformation, similarly disrupting the putative protein-protein interaction. Glycine and alanine (non-polar residues with minimal steric hindrance) and glutamic acid (acidic residue) produce asymptomatic infections. I hypothesize these small, non-polar residues and acidic residues prevent the conserved motifs from achieving their optimal geometry, thereby abolishing the putative protein-protein interaction. The same may be true for glutamine, even though it is a polar residue and thus would be expected to produce weak symptoms, according to this working hypothesis.

This hypothesis does not address those residues which render GFLV-GHu non-infectious via agroinoculation: tyrosine, cysteine, aspartic acid, phenylalanine, histidine, isoleucine, and methionine. Without further investigation, I cannot determine if the failure of these constructs to establish systemic infection is due to an actual change to GFLV biology or rather merely due to some failure of the

agroinoculation itself (e.g. incomplete T-DNA transfer due to undetected plasmid mutations). Therefore, I cannot confidently assign a biological significance to the impact of these mutations on the function of $1E^{Pol}$ or its role in symptom development

Serendipitous mutants featuring truncations and deletions of GFLV $1E^{Pol}$

While the direct plasmid mutagenesis process is generally straightforward and reliable, it (like all PCR-based applications) is prone to some level of error. For any given construct, I routinely screened at least eight clones for the engineered mutation. In the vast majority of cases, I obtained at least one (though typically more) successfully mutated clone within the first round of screening. Most of the "failed" clones fell into one of two categories: those which contained no alterations (i.e. identical to the parental plasmid) or which were only partially mutated and thus contained one or two base pair deletions within the targeted codon. In this second category of "failed" constructs, the deletions triggered a frameshift that often produced premature stop codons which should, theoretically, cause a C-terminal truncation of the $1E^{Pol}$ protein during translation of the RNA1 open reading frame. I had the chance to test three such serendipitous $1E^{Pol}$ C-terminal truncation mutants in agroinoculation (Figure A2).

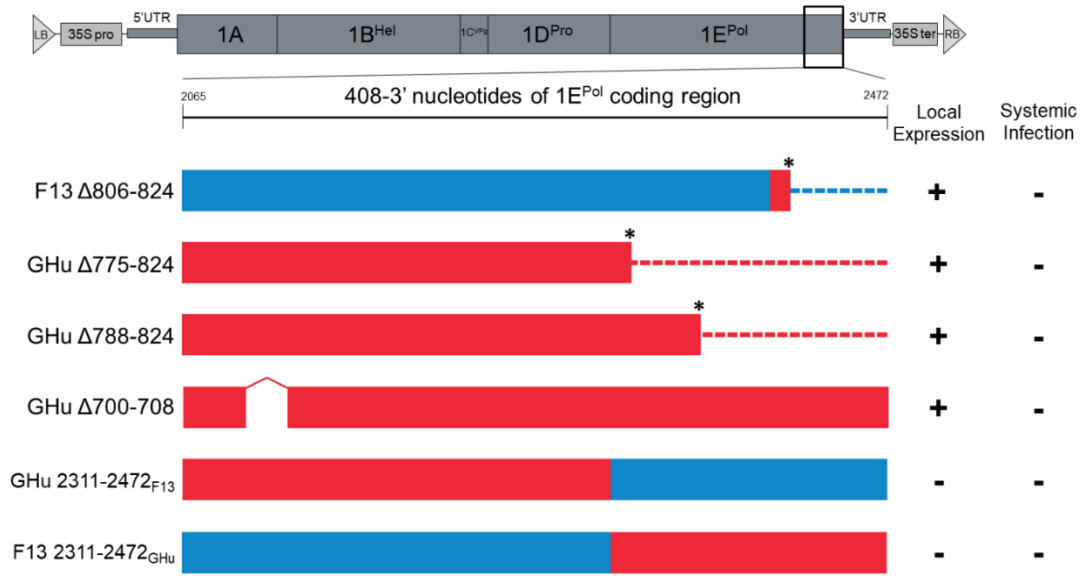


Figure A2. Grapevine fanleaf virus (GFLV) $1E^{Pol}$ chimera, deletion, and truncations mutants and their behavior in *A. tumefaciens*-mediated inoculation (agroinoculation) of *Nicotiana benthamiana*. ***Top*** Schematic of GFLV RNA1 cDNA within the T-DNA region of pCLEAN-G181, a binary vector. The unshaded box at the 3' end of the $1E^{Pol}$ coding region indicates the 408 nucleotide stretch that was targeted for mutagenesis in the creation of the mutants and is equivalent to the full-length boxes shown in the bottom of the figure. ***Bottom*** Schematics of the ultimate 408 nucleotides of the $1E^{Pol}$ coding region of several GFLV RNA1 mutants. The first column is the names of the mutants, with the first word of each mutant names indicating which GFLV strain (F13 or GHu) was used as the template for mutant creation. The second word of each mutant name indicates what modifications were made to the parental sequence. The second column shows a schematic of the mutated region. A "*" indicates the site of a premature stop codon. Dashed lines indicate portions of the $1E^{Pol}$ coding sequence which are untranslated due to a premature stop codon. Red boxes and dashed lines indicate sequences of GHu origin; F13 sequences are shown in blue. A "^" indicates a region of deleted sequence. The third column indicates whether the mutant was capable of local expression, as determined by GFLV double antibody sandwich enzyme-linked immuosorbent assay (DAS-ELISA) of agroinoculated leaves sampled at three to four days post inoculation. The fourth column indicates whether the mutant was capable of systemic infection, as determined by GFLV DAS-ELISA of apical, uninoculated leaves at 2 to 3 weeks post inoculation.

The first, F13 Δ806-824, was an unintentional result of the process to create F13 2404-2424_{GHu} (Figure 4-1D). Thus, F13 Δ806-824 is a partial chimera, with nucleotides 2404-2415 (residues 802-805) of F13 $1E^{Pol}$ replaced with the corresponding sequence from GFLV-GHu. A premature stop codon at codon 806 prevents residues 806 to 824 of $1E^{Pol}$ from being translated and the remainder of the $1E^{Pol}$ coding sequence is of F13 origin. The second and third constructs, GHu Δ775-

824 and GHu Δ788-824 are unintentional products of cloning GHu V779A and GHu V791V (Table 4-5), respectively. Both contain premature stop codons which prevent translation of portions of the $1E^{Pol}$ C-terminus. All of these constructs were capable of local expression in *N. benthamiana*, but were unable to establish systemic infection (Figure A2).

A more unusual serendipitous mutant was GHu Δ700-708, an unintentional product of cloning GHu 2065-2130_{F13} (Figure 4-1). The mutant contained a deletion of 27 nucleotides, resulting in the deletion of nine residues without the introduction of a frameshift (and thus no premature stop codon). In theory, this construct should have a nearly full-length $1E^{Pol}$ protein, with the exception of the nine missing residues. As with the C-terminal truncation mutants described above, this construct was capable of local expression in agroinoculation but failed to establish systemic infection.

As I reported in Chapter 4, attempts to mutate regions of the $1E^{Pol}$ C-terminus often rendered the virus unable to establish systemic infection via agroinoculation (Figure 4-1C and D). What went unreported in Chapter 4 is whether the constructs were capable of local expression. All of the constructs in Figure 4-1C were found to be capable of local expression via agroinoculation. This includes those mutants targeting the distal portions of the $1E^{Pol}$ C-terminus: GHu 2311-2400_{F13}, GHu 2401-2472_{F13}, F13 2311-2400_{GHu}, and F13 2401-2472_{GHu}. Due to the lack of systemic infectivity of these constructs, I created chimeras targeting larger regions in the hopes that the larger chimeric regions would reunite unidentified functional motifs that may have been split-up by the earlier chimeras. Thus I created GHu 2311-2472_{F13} and F13

2311-2472$_{F13.}$ Notably, both of these constructs were incapable of even local expression via agroinoculation (Figure A2).

Thus it appears that truncation or deletions within the 1E^{Pol} C-terminus do not interfere with the ability of mutants to be locally expressed via agroinoculation. It could be informative to further characterize these mutants to determine if "local expression" actually translates to local, active replication of the virus. This could imply that these alterations of 1E^{Pol} do not affect its role in viral replication but may instead interfere with an as-yet-unidentified role of 1E^{Pol} in establishing systemic infection.

Another serendipitous mutation within the GFLV 1E^{Pol} C-terminus

Even the most experienced among us still make silly errors from time to time. We might make a batch of bacterial culture plates with agarose instead of agar (guilty!). We might be baffled as to why all our DNA gels keep coming up blank, only to later realize that we made our gels with water instead of TAE (also guilty!). And we just might, over the course of 6 years and dozens of cloning projects, manage to mistype a primer sequence on an order form and thus order an incorrect mutagenesis primer. And we may even go so far as to not realize this error until we've sequenced our putative clones and find ourselves confused as to why, for every clone, the valine codon we were trying to mutate to alanine is coming back as a leucine (guilty!). This is how mutant GHu V779L came about in the attempt to create GHu V779A (Table 4-5). I tested GHu V779L in agroinoculation twice. In the first inoculation, one out of the eight inoculated plants became systemically infected. The infection was

asymptomatic. In the second agroinoculation, none of the six inoculated plants became infected.

A TRANSLATION ENHANCER ELEMENT APPEARS TO HAVE LITTLE IMPACT ON GFLV AGROINOCULATION

To create the first-generation GFLV agroinoculation vectors, John Gottula, a former graduate student in the Fuchs lab, cloned GFLV genomic cDNAs into the expression cassette of pEPT8, then subcloned the GFLV sequences plus expression cassette into the binary plasmid pGA482G (Gottula, pg. 77-119, 2014). The pEPT8 expression cassette contains a CaMV 35S promoter, a 5' unstructured leader sequence from alfalfa mosaic virus (AlMV) RNA4, and a CaMV 35S terminator. Each GFLV cDNA was cloned between the AlMV sequence and CaMV 35S terminator. AlMV RNA4, a subgenomic mRNA of RNA3, codes for the AlMV coat protein. The 5' untranslated region of RNA4 has been found to increase the translation efficiency of chimeric mRNAs (Jobling and Gehrke 1987) and is thus included in the pEPT8 vector to enhance translation of sequences cloned into the expression cassette.

Gottula theorized that inclusion of this genetic element could interfere with the activity of the native GFLV 5' UTRs and thus might explain some of the poor performance of the early GFLV agroinoculation system. In fact, Gottula cites preliminary, unpublished data suggesting that removal of the AlMV sequence might enhance the performance of the first-generation GFLV agroinoculation vectors (Gottula, pg. 106, 268, 2014).

Since I faced similar struggles in the early days of redesigning and developing a GFLV agroinoculation system, I decided to investigate this theory. I removed the

AlMV-derived sequence from the binary plasmid containing the GFLV-F13 RNA1 cDNA via direct plasmid mutagenesis. This allowed me to directly compare GFLV-F13 RNA1 cDNA constructs that differed only in the presence or absence of a 5' leader sequence of AlMV RNA4 (Figure A3-A). At the time, I was also investigating the infectivity of several GFLV-GHu RNA2 derivatives which were part of my attempt to redesign a GFLV-based virus-induced gene silencing (VIGS) vector (see below). These included a wild-type GFLV-GHu RNA2 (Figure A3-B and C); a GHu RNA2 containing a multiple cloning site (MCS) and $1D^{Pro}$ cleavage recognition sequence between $2A^{HP}$ and $2B^{MP}$ (Figure A3-B); and a GHu RNA2 to which a "self-cleaving" peptide derived from the foot and mouth disease virus (FMDV) 2A protein had been introduced as a C-terminal fusion of $2C^{CP}$ (Figure A3-C). I performed three independent agroinoculations in which I mixed-and-matched the two F13 RNA1 constructs with the three GHu RNA2 constructs. The results are summarized in Table A2. Overall the combination of wild type F13 RNA1 (containing the AlMV sequence) with wild type GHu RNA2 had the highest average infection rate at 93.3% over the three agroinoculations. The worst performing combination was F13 ΔAlMV with GHu:F2A at 21.7%.

Figure A3. Schematics of the T-DNA regions of binary plasmids containing grapevine fanleaf virus (GFLV) cDNAs. **A.** Schematics of GFLV-F13 RNA1 cDNAs within the T-DNA region of pCLEAN-G181. 1A, protein of unknown function; $1B^{Hel}$, putative helicase; $1C^{VPg}$, genome-linked viral protein; $1D^{Pro}$, protease; $1E^{Pol}$, RNA-dependent RNA polymerase. **B.** Schematics of GFLV-GHu RNA2 cDNAs within the T-DNA region of pCLEAN-G181. $2A^{HP}$, putative homing protein; MCS, *BsiW*I and *Mlu*I multiple cloning site; $2B^{MP}$, movement protein; $2C^{CP}$, coat protein. The dotted box at the end of the cDNA in **B** is equivalent to the areas expanded in **C.** In **A** and **B,** grey triangles represent the left (L) and right (R) border sequences of the T-DNA region. Grey arrows represent the promoter sequence of the cauliflower mosaic virus (CaMV) 35S RNA. Grey trapezoids represent the 5' unstructured leader sequence from alfalfa mosaic virus (AlMV) RNA4. Narrow filled boxes represent untranslated regions (UTRs). Large filled boxes represent the open reading frame (ORF), with cleavage sites represented by vertical lines dividing the ORF. Grey octagons represent the CaMV 35S terminator sequence. **C.** Schematics of the region spanning the $2C^{CP}$ C-terminus and 3'UTR of several stages of a redesigned GFLV VIGS/protein expression vector. In each pair of schematics, the top schematic is of the nucleotide sequence of the construct. Open triangles indicate a restriction enzyme recognition site. LIC, ligation independent cloning site; FMDV 2A, "auto-catalytic" peptide 2A from foot and mouth disease virus. The bottom schematic of each pair is of the theoretical polypeptide sequence of the construct. Rounded boxes represent individual polypeptides. Filled triangles indicate a ribosome skipping site that prevents the formation of a glycine-proline peptide bond and thus produces separate polypeptides. Arrows extending from boxes indicate that the sequence continues beyond the illustrated region. In **A** through **C**, sequences of GFLV-F13 origin are shown in blue (ORFs) and light blue (UTRs); GFLV-GHu sequences are shown in red (ORFs) and pink (UTRs); LIC sequences are shown in green; FMDV 2A sequences are shown in yellow. The names of the constructs are shown along the left side.

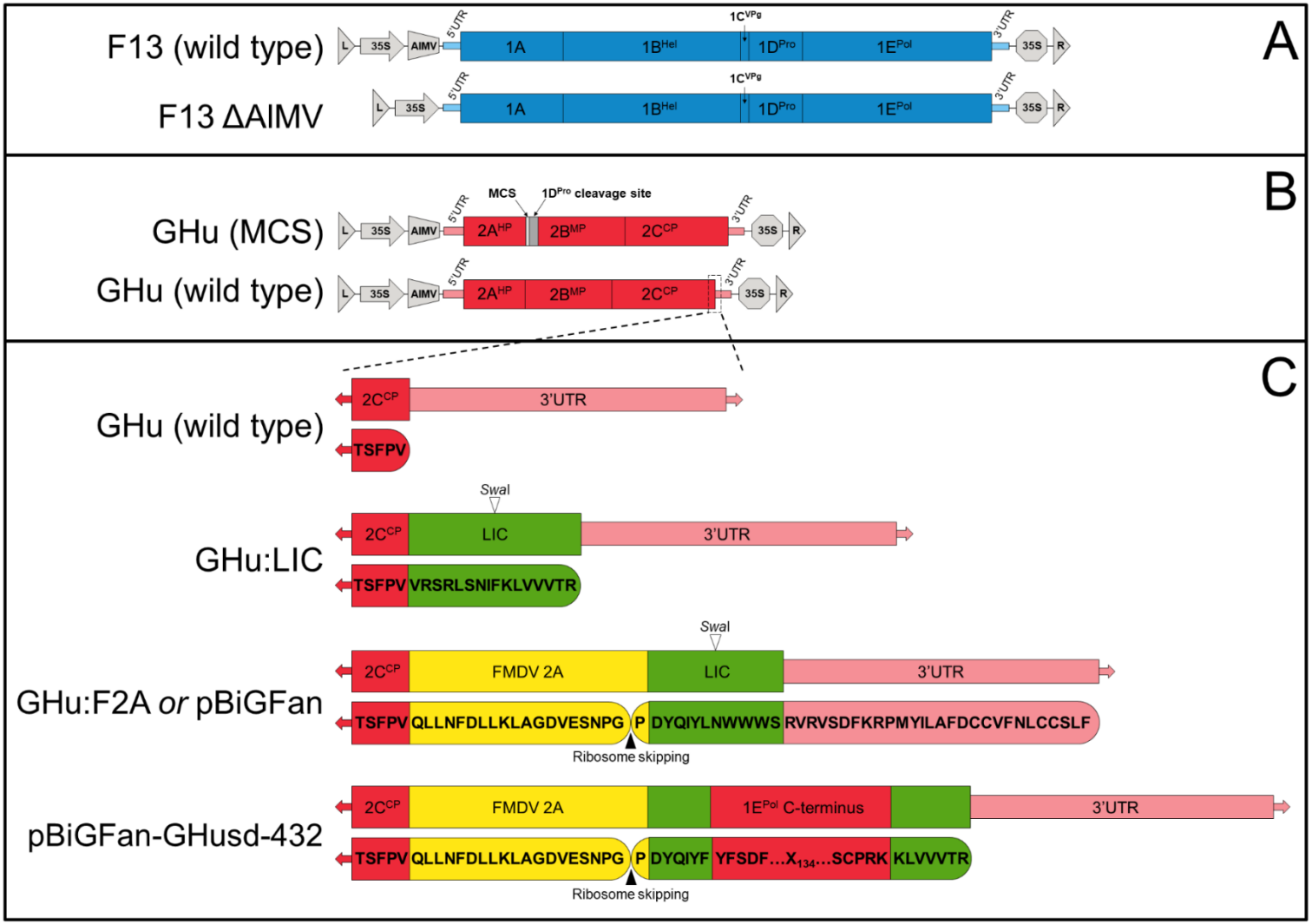

A
F13 (wild type)
F13 ΔAIMV
L
35S
AIMV
5'UTR
1A
1B^Hel
1C^VPg
1D^Pro
1E^Pol
3'UTR
R
B
GHu (MCS)
GHu (wild type)
MCS
1D^Pro cleavage site
2A^HP
2B^MP
2C^CP
C
GHu (wild type)
3'UTR
TSFPV
SwaI
GHu:LIC
LIC
VRSRLSNIFKLVVVTR
GHu:F2A or pBiGFan
FMDV 2A
QLLNFDLLKLAGDVESNPG
P
DYQIYLNWWWS
RVRVSDFKRPMYILAFDCCVFNLCCSLF
Ribosome skipping
pBiGFan-GHusd-432
1E^Pol C-terminus
DYQIYF
YFSDF…X134…SCPRK
KLVVVTR

Table A2. Agroinoculation experiments to compare infection rates of various grapevine fanleaf virus (GFLV) RNA1 and RNA2 constructs in *Nicotiana benthamiana*

		Trial 1			Trial 2			Trial 3			
RNA1	**RNA2**	**Inoc.[a]**	**Infect.[b]**	**% infected**	**Inoc.**	**Infect.**	**% infected**	**Inoc.**	**Infect.**	**% infected**	**Ave.[c]**
F13 ΔAlMV	GHu (wt)	7	7	100.0%	6	3	50.0%	5	1	20.0%	56.7%
F13 (wt)	GHu (wt)	7	7	100.0%	6	6	100.0%	5	4	80.0%	93.3%
F13 ΔAlMV	GHu:F2A	7	2	28.6%	6	1	16.7%	5	1	20.0%	21.7%
F13 (wt)	GHu:F2A	7	4	57.1%	6	2	33.3%	5	0	0.0%	30.2%
F13 ΔAlMV	GHu (MCS)	7	4	57.1%	6	3	50.0%	5	0	0.0%	35.7%
F13 (wt)	GHu (MCS)	7	2	28.6%	6	0	0.0%	5	2	40.0%	22.9%

[a]Inoc. = number of plants inoculated with the indicated RNA1/RNA2 combination in the given agroinoculation trial

[b]Infect. = number of plants which became systemically infected with GFLV following agroinoculation, as determined by double antibody sandwich enzyme-linked immunosorbent assay at three to four weeks post inoculation.

[c]Ave. = average rate of infection (percent infection) for the indicated RNA1/RNA2 combination across the three agroinoculation experiments.

I performed a series of ANOVAs to determine if any of the differences in infection rates were statistically significant. First, I ran a two-way ANOVA, which allowed me to treat the RNA1 and RNA2 components as separate, independent factors and also determine if any interactions between RNA1 and RNA2 affected infection rates. Both RNA1 and the interaction of RNA1 and RNA2 had no significant effect on infection rates ($p = 0.396$ and 0.289, respectively). The only significant variable was RNA2 ($p = 0.011$). This suggests that at least one pair of RNA2 constructs performed significantly differently from each other in agroinoculation. A Tukey's honest significant difference (HSD) test revealed that GHu:F2A and GHu (MCS) were significantly different from wild-type GHu RNA2 ($p = 0.016$ and 0.025, respectively). This suggests that the identity of the RNA2 construct had an effect on infection rates, while the presence or absence of the AlMV sequence within the F13 RNA1 binary plasmid had no effect.

Both GHu:F2A and GHu (MCS) were prototype constructs and not intended for standard use in GFLV agroinoculation (see below). However, the wild type GHu RNA2 has been used extensively (in fact, it was the RNA2 source for the vast majority of the GFLV agroinoculations reported in Chapters 3, 4, and 5) and I wanted to know if the AlMV sequence had an effect on the infection rate of F13 RNA1 in combination with wild type GHu RNA2. So I ran a one-way ANOVA on a subset of the data containing results only from wild type GHu RNA2. I found no significant difference between infection rates of F13 wild type and F13 ΔAlMV ($p = 0.205$). Thus, while there is a numerical difference in the average infection rates of F13 wild type (93.3%)

and F13 ΔAlMV (56.7%) when co-inoculated with wild type GHu RNA2, this discrepancy is not statistically different.

These results are at odds with the preliminary results reported by Gottula, who claimed that constructs lacking the AlMV sequence had 100% infectivity in *N. benthamiana*, while constructs with the AlMV sequence had infection rates of 0 to 1.6% (Gottula, pg. 268, 2014). Several factors could explain this discrepancy. First, it is not clear from Gottula's work if the comparison was made between GFLV constructs differing only in the presence or absence of the AlMV sequence or if the compared constructs contained other modifications, such as fluorescent protein tags that were cloned within the RNA2 cDNA constructs; it is not clear if the constructs are even of the same GFLV strain. Therefore, there may be other confounding factors in Gottula's preliminary data to explain the difference in infection frequency between constructs with and without the 5' leader sequence from AlMV RNA4. Second, the majority of Gottula's GFLV constructs were based on the binary vector pGA482G. This vector was originally designed for *A. tumefaciens*-mediated stable plant transformation and thus contains other modules within the T-DNA region (such as plant reporter gene for kanamycin resistance) which are not necessary for agroinoculation. In contrast, the constructs I utilized for my investigation are based on the binary vector pCLEAN-G181, the minimal T-DNA region of which contains only the GFLV cDNAs and pEPT8-derived 35S expression cassette. It is possible that the effects of the AlMV leader sequence on GFLV infectivity via agroinoculation are mediated in part by the binary vector from which these genetic elements are launched *in planta*. The lack of effect of the AlMV sequence I found may be specific to

pCLEAN-G181. Finally, my study was rather small and consisted of a total of 18 plants per treatment over three independent inoculations. Perhaps a larger study would have sufficient power to detect a significant difference between the infection rates of the constructs under investigation.

ATTEMPTS TO REDESIGN A GFLV-BASED VIGS AND PROTEIN EXPRESSION VECTOR

One of the successes of Gottula's early GFLV agroinoculation system was the engineering of GFLV-GHu RNA2 into a virus-induced gene silencing (VIGS) vector for gene silencing in *N. benthamiana.* While originally intended for VIGS, the structure of the construct also allows it to be used for protein expression (see below). This construct went by several aliases over the course of its use. In Gottula's doctoral thesis, it is variably referred to as "RNA2" "pROK2" "pROX" and "pG_R". These are, to the best of my knowledge, equivalent to the plasmid pG_R which I used in Chapter 4 for VIGS in *N. benthamiana.* In pG_R, the GFLV cDNA has been modified to include a *BsiW*I/*Mlu*I cloning site between the $2A^{HP}$ and $2B^{MP}$ coding sequences. Also, a 75 bp fragment coding for the $1D^{Pro}$-$1E^{Pol}$ junction has been included between the cloning site and $2B^{MP}$ coding region, which allows for fragments inserted into the cloning site to be cleaved from $2B^{MP}$ by the GFLV $1D^{Pro}$ protease.

While this construct was a major step forward for GFLV molecular biology research, it suffered from several drawbacks. First, since GFLV genome expression proceeds by monocistronic translation and the cloning site is within the RNA2 ORF, any fragments inserted into the cloning site must be fully translatable (i.e. contain no premature stop codons) and the nucleotide length must be a multiple of three to

prevent the introduction of a frameshift. Second, inserts are cloned as a C-terminal fusion of $2A^{HP}$, which localizes to the site of GFLV genome replication (Gaire et al. 1999). This could prevent the translated sequences from being localized to other cellular compartments. While this is not an issue for VIGS (in which only the nucleotide sequence is necessary for silencing), this could be problematic if one is using the vector for protein expression, especially if characterization of the expressed protein is dependent on correct cellular localization. Finally, the binary vector which pG_R is based on, pG2b, contains an additional expression module within the T-DNA region. This additional module codes for the 2b silencing suppressor protein of cucumber mosaic virus (CMV). The purpose of this additional module was to reduce or eliminate the need to co-infiltrate a silencing suppressor during agroinoculation and thereby simplify the agroinoculation protocol. However, this means CMV 2b is always present in any agroinoculation utilizing pG_R. While to date I am not aware of any instances of CMV 2b interfering with GFLV agroinoculation, it seems preferable to introduce additional elements to one's experiments by *choice* rather than by default.

I set out to redesign the GFLV VIGS/protein expression vector with the intent of overcoming the limitations outlined above. I used a two-pronged approach. In the first redesign, I merely moved the GFLV cDNA of pG_R to a different binary plasmid. In the second redesign, I made new modifications to the cDNA of GFLV-GHu RNA2 to rebuild the VIGS/protein expression vector from the ground up.

Redesign I: GHu (MCS)

To overcome the drawback of the CMV 2b expression module of pG_R, I subcloned the modified cDNA of GFLV-GHu RNA2 within the 35S expression

cassette from pG_R into the binary plasmid pCLEAN-G181 by amplification with primers ROXGFLVtoCLEANG181-F and ROXGFLVtoCLEANG181-R, followed by digestion with *Not*I and ligation into *Not*I-digested pCLEAN-G181 (Table A3). The resulting construct is equivalent to construct "GHu (MCS)" in Table A2 and Figure A3-B. GHu (MCS) was systemically infectious via agroinoculation in *N. benthamiana* when combined with either F13 RNA1 wild type or F13 ΔAlMV (Table A2). However, the pCLEAN-G181 MCS contains a *BsiW*I cut site, which functionally prevents the use of the *BsiW*I cut site within the GFLV cDNA for the introduction of exogenous sequences. To overcome this obstacle, I deleted the *BsiW*I cut site from the pCLEAN-G181 MCS of GHu (MCS) via mutagenic PCR using primers CLEAN-ROX-deleteBsiWI-F and CLEAN-ROX-deleteBsiWI-R (Table A4). This produced construct GHu (MCS) v2 (not shown). The infectivity of GHu (MCS) v2 in agroinoculation of *N. benthamiana* remains untested, nor has the VIGS capacity of GHu (MCS) or GHu MCS v2 been assessed. However, in theory, GHu (MCS) v2 should perform similary to pG_R in VIGS assays.

Table A3. Primers for PCR cloning.

Construct	Fragment template	Destination vector	Forward primer	Forward primer sequence (5'-3')[a]	Forward primer RE[b]
			Reverse primer	Reverse primer sequence (5'-3')[a]	Reverse primer RE[b]
GHu (MCS)	pG_R	pCLEAN-G181	ROXGFLVtoCLEANG181-F	TAAGCGCGGCCGCTCTGATCAAGATCTCCGG	*Not*I
			ROXGFLVtoCLEANG181-R	TAAGCGCGGCCGCAAGATCTTAGTACTGATT	*Not*I
pBiGFan-GHusd-432	pCLEAN-GHu-1	pBiGFan	GHusdtopBiGFan-F	*gactatcaaatatatttt*CTATTTTTCTGATTT	*Swa*I (LIC)
			GHusdtopBiGFan-R	*gaccaccaccaatttt*CTTCCTCGGGCATGA	*Swa*I (LIC)
pCLEAN-GHu-2	pGA282G-GHu-2	pCLEAN-G181	GHu2toCLEANG181-F2	TAAGCGCGGCCGCGTTGATGAAGCTTCTAGA	*Not*I
			GHu2toCLEANG181-R3	TAAGCGCGGCCGCTACCGTTAACGAGCTCTAG	*Not*I

[a]Restriction enzyme cut sites underlined. Bases making up the overhanging sequences of the ligation independent cloning (LIC) site are shown in lower case italics.

[b]RE: restriction enzymes used to cut vectors and PCR-amplified inserts. "(LIC)" indicates that the destination vector was linearized with the designated restriction enzyme.

Table A4. Primers for mutagenesis of grapevine fanleaf virus (GFLV) cDNAs

Construct Name	Template	Forward Primer	Forward Primer Sequence (5'-3')[b]	Reverse Primer	Reverse Primer Sequence (5'-3')[b]	Designed[a]
F13-1E-L800H+G802K	pCLEAN-F13-1E-G802K	F13-1E-L800H+G802K_For	TGATACGCAAcacTTTAAGAACAATCTTT	F13-1E-L800H+G802K_Rev	GTCACAACCCTGGTTGCT	+
GHu K802D	pCLEAN-GHu-1	GHu-1E-K802D-GAT_For	TGCGCACTTTgatAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802H [CAT]	pCLEAN-GHu-1	GHu-1E-K802H-CAT_For	TGCGCACTTTcatAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802H [CAC]	pCLEAN-GHu-1	GHu-1E-K802H-CAC_For	TGCGCACTTTcacAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802S	pCLEAN-GHu-1	GHu-1E-K802S_For	TGCGCACTTTagtAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802T	pCLEAN-GHu-1	GHu-1E-K802T_For	TGCGCACTTTaccAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802C	pCLEAN-GHu-1	GHu-1E-K802C-For	TGCGCACTTTtgtAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802R [AGA]	pCLEAN-GHu-1	GHu-1E-K802R-AGA_For	TGCGCACTTTagaAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802Y [TAT]	pCLEAN-GHu-1	GHu-1E-K802Y-TAT_For	TGCGCACTTTtatAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802I [ATT]	pCLEAN-GHu-1	GHu-1E-K802I-ATT_For	TGCGCACTTTattAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802M	pCLEAN-GHu-1	GHu-1E-K802M_For	TGCGCACTTTagtAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802W	pCLEAN-GHu-1	GHu-1E-K802W_For	TGCGCACTTTtggAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+
GHu K802F	pCLEAN-GHu-1	GHu-1E-K802F_For	TGCGCACTTTtttAATAGCCTTTTG	GHu-1E-K802G_Rev	GTACTACTCACAGTTTTGG	+

Table A4, cont.

Construct Name	Template	Forward Primer	Forward Primer Sequence (5'-3')[b]	Reverse Primer	Reverse Primer Sequence (5'-3')[b]	Designed[a]
GHu V779L	pCLEAN-GHu-1	GHu-1E-V779A-F[c]	CGTGAGTGAGcttTGTTTGAAATGTTGC	GHu-1E-V779A-V-R	TAAACCCCTGGAGTTTGTTC	+
F13 Δ806-824	pCLEAN-F13-1	F13-1E-2404-24-F	tttgaagacgCATCTTAAGGCTTTGAGG	F13-1E-2404-24-R	aggctattcttAAAAAGTTGCGTATCAGTC	-
GHu Δ788-824	pCLEAN-GHu-1	GHu-1E-V791V-F	TTTAGGTGTTgttGCCAAAACTG	GHu-1E-V791V-A-R	CATCTGTGGCAACATTTC	-
GHu Δ700-708	pCLEAN-GHu-1	GHu1-F13sd-A-F2	tcttatgaggggggtgaagctttaaaggaaattTTCTCTTTTTGTGAGACTGC	GHu1-F13sd-A-R2	gtttgttttggcatggactttcttgatatgaatGGGACTTTCAAAATCAGAAAAATAG	-
GHu Δ775-824	pCLEAN-GHu-1	GHu-1E-V779A-F	CGTGAGTGAGcttTGTTTGAAATGTTGC	GHu-1E-V779A-V-R	TAAACCCCTGGAGTTTGTTC	-
GHu 2311-2472$_{F13}$	pCLEAN-GHu1-2311-2400	GHu1-F13sd-E-F	ttgaggaaaattcagaatcatacatgccttaggaaaTAATTCTTCCAACCCTTGG	GHu1-F13sd-ABCDE-R	agccttaagatgagtctttaaaagattgttgccaaaAAGTTGCGTATCAGTCAC	+
F13 2311-2472$_{GHu}$	pCLEAN-F131-2311-2400	F131-Ghusd-E-F	ctgaggcgtgttcaggctcactcatgccgaggaagTAAGCCTTCCAATTCTTGG	F131-Ghusd-ABCDE-R	agtcttaagatgcgtcttcaaaaggctattcttaaaGTGCGCAGTACTACTCAC	+
GHu K802R [CGT]	pCLEAN-GHu-1	GHu-1E-K802R-F	TGCGCACTTTcgtAATAGCCTTTT	GHu-1E-K802G-R	GTACTACTCACAGTTTTGG	+
GHu K802Y [TAC]	pCLEAN-GHu-1	GHu-1E-K802Y-F	TGCGCACTTTtacAATAGCCTTTTG	GHu-1E-K802G-R	GTACTACTCACAGTTTTGG	+
F13 ΔAlMV	pCLEAN-F13-1	DeleteAlMVCLEANF131-F	GTCGACATGAAAATTTCCC	DeleteAlMVRNA4-R	ACGTGTCCTCTCCAAATG	+
GHu (MCS) v2	pCLEAN-GHu-ROXv1	CLEAN-ROX-deleteBsiWI-F	TACGGGCCCTCGAGTCGA	CLEAN-ROX-deleteBsiWI-R	CCTTAATTAAGCGGCCGCAAG	+
GHu:LIC	pCLEAN-GHu-2	AddLICtopCLEAN-GHu2-35S-F	tttaaattggtggtggtcacgcgtTAGGGTATCTGACTTAAAAG	AddLICtopCLEAN-GHu2-35S-R	tatatttgatagtctcgagcgtacGACTGGAAAACTGGTTCTTC	+

Table A4, cont.

Construct Name	Template	Forward Primer	Forward Primer Sequence (5'-3')[b]	Reverse Primer	Reverse Primer Sequence (5'-3')[b]	Designed[a]
GHu:FMDV 2a *or* pBiGFan	pCLEAN-GHu2-LIC	AddFMDV2a-F	gcgggagacgtcgagtccaaccccg ggcccGACTATCAAATAT ATTTAAATTGGTG	AddFMDV2a-R	aagcttaagaaggtcaaaattcaa cagctgGACTGGAAAAC TGGTTCTTC	+
pCLEAN-mGHu2	pBiGFan	RestoreSTOPGH u2FMDV2a-F	TTAGGGTATCTGACTTT AAAAGACCC	RestoreSTOPG Hu2FMDV2a-R	ACGCGTGACCACCAC CAA	+

[a]"+" indicates that the construct was produced as designed. "-" indicates that the construct is an accidental result of PCR error during mutagenesis using the listed template and primers.

[b]Mutagenic bases are shown in lower cases. Bases which bind to the template plasmid are shown in upper case.

[c]This primer was originally intended to mutate GFLV-GHu 1E^{Pol} valine 779 to alanine. The primer was ordered incorrectly and contains mutagenic bases which code for leucine.

Redesign II: pBiGFan, rationale for design elements

I based the redesigned GFLV VIGS/protein expression vector on GFLV-GHu RNA2 since it had been shown to outperform GFLV-F13 RNA2 in GFLV agroinoculation (Chapter 3).

Next, I considered the nature of the cloning site within the GFLV RNA2 ORF. In pG_R, the cloning site consists of *BsiW*I and *Mlu*I cut sites and requires a ligation step to clone the desired fragment into the prepared vector. PCR cloning is already a laborious process, consisting of digestion and clean-up of the destination vector, PCR amplification and clean-up of the insert fragment, restriction enzyme digestion and clean-up of the insert fragment, and finally ligation of the insert fragment and vector prior to transformation into *E. coli*. Each clean-up step represents a potential loss of material. Furthermore, PCR cloning relies on restriction enzyme digestion of the insert and vector to produce compatible ends. If one or both of the enzymes cut sites is present in the insert, it can be difficult or impossible to clone the insert into the vector via PCR cloning. Thus, I opted to employ a ligation independent cloning (LIC) site instead of a traditional restriction enzyme-based cloning site.

In LIC, a single restriction enzyme is used to linearize the plasmid. The 12 to 15 nucleotides flanking this cut site must lack one of the four bases (e.g. contain no "G"). Then, the ends of the plasmid are "chewed back" by the 3' to 5' exonuclease activity of T4 polymerase. The T4 polymerase reaction is supplemented with the single dNTP species which is lacking in the regions flanking the cut site (e.g. dGTP). When the T4 polymerase encounters the first instance of the lacking residue, (e.g. "G"), a dNTP (dGTP) is incorporated by the polymerase, preventing further "chew

back". The result is long overhanging sequences. The insert sequence is amplified using primers which are complementary to the overhang sites and the amplified insert sequence undergoes similar T4 "chew back" to produce overhang sequences complementary to those of the vector. The extended length of the complementary overhang sequences (12 to 15 bp, compared to 4 to 6 bp overhangs produced by most restriction enzyme digests) provides sufficient stability via base-pairing alone for the vector-insert complex to be transformed directly into *E. coli* without ligation. Thus, LIC relies on a single enzyme cut site for linearization of the plasmid and does not require restriction enzyme digestion of the insert sequence, eliminating the possibility of inserts being incompatible with the vector due to the presence of enzyme cut sites within the insert sequence.

I also reconsidered the location and nature of the linker sequence between the GFLV polyprotein and the inserted sequence. In pG_R, inserted sequences are translated within the P2 polyprotein as a C-terminal fusion of the $2A^{HP}$ protein. A linker sequence, downstream of the insertion site, coding for the $1D^{Pro}$-$1E^{Pol}$ junction allows $2B^{MP}$ to be cleaved from the inserted sequence by $1D^{Pro}$. I wanted the inserted sequences of the redesigned vector to be fully cleaved from any GFLV-encoded proteins. This should prevent the inserted sequences from interfering with the function of the GFLV-encoded proteins and also prevent strict localization of the inserted sequences to the site of the fusion partner. Thus, I chose to move the insertion site of the new vector to the C-terminus of $2C^{CP}$. By placing the insertion site downstream of the GFLV-encoded polyprotein, this removed the requirements that the inserted sequences be both fully-translatable and have nucleotide lengths of multiples of three.

Instead of using the $1D^{Pro}$-$1E^{Pol}$ junction sequence for cleavage of the inserted sequences from the P2 polyprotein, I chose to utilize a protease-independent cleavage site. This would allow me to use any GFLV RNA1 cDNA construct in combination with the redesigned RNA2 vector without worrying about differences in $1D^{Pro}$ activity introducing variability in the cleavage of the inserted sequences from the polyprotein. For protease-independent cleavage, I chose the foot and mouth disease virus (FMDV) 2A catalytic peptide (F2A), which induces sequence-dependent "auto-cleavage" of a characterized glycine-proline peptide bond through ribosome skipping during translation (Gopinath et al. 2000; Ryan and Drew 1994).

Redesign II: pBiGFan, cloning process

With these design elements in mind, I began the cloning process by moving the cDNA of GFLV-GHu RNA2 within the CaMV 35S expression cassette from pGA482G-GHu-2 into binary plasmid pCLEAN-G181 to produce pCLEAN-GHu-2 (Chapter 3). This construct has been used extensively as the RNA2 source for GFLV agroinoculation (Chapters 3, 4, and 5).

The next step was to add the LIC site between the $2C^{CP}$ C-terminus and 3' UTR. This was accomplished via mutagenic PCR of pCLEAN-GHu-2 using primers AddLICtopCLEAN-GHu2-35S-F and –R (Table A4).

The resulting construct, GHu:LIC (Figure A3-C), was tested for infectivity in agroinoculation once. While this construct failed to establish systemic infection in *N. benthamiana* via agroinoculation, the positive controls for that particular inoculation also failed (data not shown), so it is still unclear if GHu:LIC is truly non-infectious via

agroinoculation or if some other issue in the inoculation prevented it from establishing infection.

Next, I inserted an F2A sequence between the $2C^{CP}$ C-terminus and the LIC site by mutagenic PCR of GHu:LIC using primers AddFMDV2a-F and –R (Table A4). This produced construct GHu:F2A (Figure A3-C), which was shown to be systemically infectious in combination with either F13 RNA1 wild type or F13 ΔAlMV (Table A2). Since this construct contained all the design elements needed for the redesigned GFLV VIGS/protein expression vector, I renamed it pBiGFan (binary plasmid of grapevine fanleaf virus) to distinguish it from the GFLV-GHu RNA2 constructs from which it was derived. It should be noted that, though the F2A sequence was cloned directly downstream and in-frame with the $2C^{CP}$ coding sequence, the LIC site is 32 nucleotides long (i.e. not a multiple of three) which introduced a frameshift downstream of the F2A sequence. As a result, translation of P2 should theoretically progress well into the 3'UTR sequence, producing a novel polypeptide during GFLV genome expression (Figure A3-C). In an attempt to prevent expression of this novel peptide, which could have unknown effects on bioassays utilizing pBiGFan, I tried to re-establish a stop codon at the C-terminal end of the LIC site within pBiGFan by mutagenic PCR of pBiGFan with primers RestoreSTOPGHu2FMDV2a-F and –R. The resulting construct was unable to establish systemic infection via agroinoculation in *N. benthamiana* in combination with either F13 RNA1 wild type or F13 ΔAlMV (data not shown). Thus, going forward, I utilized pBiGFan, despite the presence of the novel theoretical peptide.

Construct pBiGFan appears to be unstable for expression of exogenous sequences

The completion of the pBiGFan construct coincided with my attempts to utilize the original GFLV VIGS vector, pG_R, to determine whether the GFLV-GHu symptom determinant was capable of inducing vein clearing via VIGS (Chapter 4). To complement that work, I cloned the "full length" GFLV-GHu symptom determinant (i.e. the ultimate 432 nt of the $1E^{Pol}$ coding region) into the LIC site of pBiGFan to create construct pBiGFan-GHusd-432. The fragment was cloned in-frame so as to be fully translatable; this also restored the stop codon at the 3' end of the LIC site (Figure A3-C).

I agroinoculated sixteen *N. benthamiana* plants with pBiGFan-GHusd-LIC combined with pGA482G-F13-1 and the GLRaV-2 p24 construct. Twelve plants became systemically infected, as determined by DAS-ELISA of apical leaves at 3 weeks post inoculation, and all were asymptomatic. I performed RT-PCR on total RNA extracted from apical leaves using primers pBiGFan-MCS-F (TTG GAC CAA GCA AAG AGT ATT A) and pBiGFan-MCS-R (AAA CAA GTT AAA ACA ACT ACA ACA), which flank the LIC insertion site. Gel electrophoresis of the RT-PCR products revealed that most of the virus progeny had not retained the original insertion (Figure A4). Three plants failed to produce an RT-PCR product, despite producing a positive signal in GFLV DAS-ELISA (Figure A4, samples 9, 12, and 13). I sequenced all of the RT-PCR products using pBiGFan-MCS-F and –R to determine the actual sequences of the viral progeny. Three plants were infected with virus progeny that perfectly matched the original construct (Figure A4, samples 6, 7, and, 16). Two plants produced small bands (<200 bp, less than the theoretical size, 277 bp, of the

expected band for pBiGFan lacking an insert); and both the forward and reverse sequencing runs for both samples had significant background noise, preventing meaningful analysis of the sequencing results.

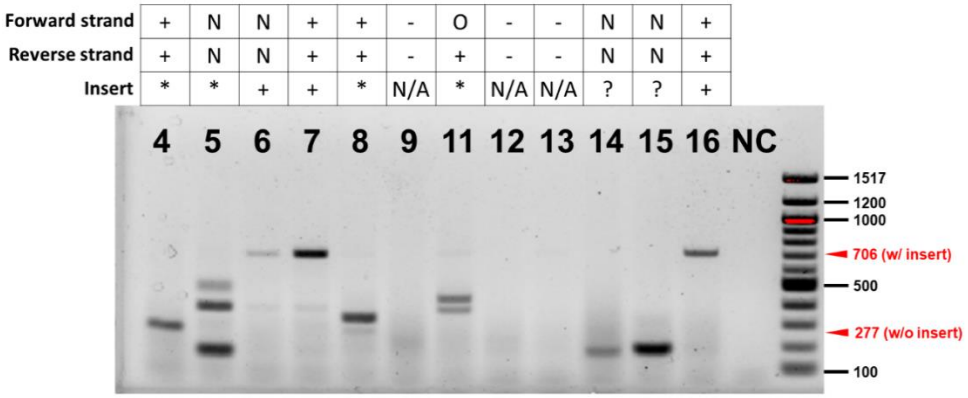


Figure A4. RT-PCR amplification of insertion site of progeny viruses from plants agroinoculated with pBiGFan-GHusd-432 that tested positive for GFLV by DAS-ELISA at 3 weeks post inoculation. The first and second rows indicate the quality of the sequencing of the RT-PCR product using the forward and reverse primers, respectively. "+" indicates that the sequencing ran well. "N" indicates the sequencing run was noisy or deteriorated. "-" indicates the sequencing run failed. "O" indicates the sequencing run contained significant overlaps (double peaks). The third row indicates the status of the inserted sequence (a C-terminal portion of the GFLV-GHu $1E^{Pol}$ coding sequence) in the viral progeny. "+" indicates that fidelity of the insertion was maintained and that the RT-PCR product matched the expected sequence. "*" indicates that the inserted sequence was retained in part. "?" indicates that there was insufficient sequencing data to determine the status of the inserted sequence. "N/A" indicates not applicable (due to failure of both sequencing runs). Size standards are shown along the right. Red arrows indicate the expected band size for progeny which retained the inserted sequence ("w/ insert") or which lack the inserted sequence ("w/o insert").

At first glance, it appears that the insert is missing from sample 4 as the band size closely matches the expected band size of pBiGFan without an insert (Figure A4, sample 4). Close inspection of the sequencing results instead reveals that while the F2A and 5' LIC site sequences are present, only a small fragment of the inserted GHu symptom determinant sequence was retained: TAT TTT TCT GAC TTT, coding for YFSDF. This is immediately followed by a truncated 3' UTR sequence (missing the first four nucleotides), without the intervening 3' LIC site. Interestingly, the 3' end of this retained sequence overlaps with nucleotides 5 to 13 (TCT GAC TTT) of the

GFLV-GHu RNA2 3'UTR. It is possible that the inserted fragment got spliced out during viral genome replication under selective pressure to partially reconstitute the wild-type 3' UTR.

Sample 5 produced multiple bands (Figure A4, sample 5). The sequencing results appear to be largely for the middle band. This GFLV "isolate" retained the first 6 nucleotides of the F2A sequence, followed by approximately 163 nt of the 3' terminal sequence of the GFLV-GHu symptom determinant (such an insert would be expected to produce a 370 bp PCR fragment). Both the forward and reverse sequencing reactions had high background noise, likely due to the mixed nature of the sequencing templates, which prevented more conclusive analysis.

In sample 8, the FMDV 2A sequence is retained, along with the LIC sites flanking the insertion site. However, the insertion has been reduced to the first 30 bp and final 13 bp of the GFLV-GHu symptom determinant (nucleotides 2040 to 2069 and 2460 to 2472 of the $1E^{Pol}$ coding region). This "splice" creates a premature stop codon, such that only "YFSDFESPV" of the symptom determinant region is predicted to be translated.

In sample 11, the forward sequencing reaction had significant overlap signals and was not included in the analysis. The reverse reaction clearly sequenced the ultimate 130 bp of the GFLV-GHu symptom determinant within the LIC site, however the sequencing gets messy (overlap) upstream of this sequence, preventing clear interpretation. However, upstream of the overlap portion, the sequencing deconvolutes and clearly matches to a 3' portion of the $2C^{CP}$ coding sequencing. Given that sample

11 clearly has two bands, it could be that there are two types of rearrangements which differ in what portion of the insert was retained (Figure A4, sample 11).

In summary, pBiGFan appears to be highly unstable for the introduction of exogenous genetic material to GFLV RNA2. While I did successfully clone portions of the *N. benthamiana* phytoene desaturase (pds) gene into pBiGFan as a proof-of-concept VIGS construct, I was never able to get this construct to establish systemic infection via agroinoculation (data not shown). Fortunately, around this time, I achieved reliable success with agroinoculation of the re-engineered GFLV agroinoculation vectors, including success in direct plasmid mutagenesis to produce GFLV chimeras (Chapter 3). This negated the need for a GFLV VIGS/protein expression vector for this project and I abandoned pBiGFan to its obsolescence.

REFERENCES

Gaire, F., Schmitt, C., Stussi-Garaud, C., Pinck, L., and Ritzenthaler, C. 1999. Protein 2A of grapevine fanleaf nepovirus is implicated in RNA2 replication and colocalizes to the replication site. Virology 264:25–36

Gopinath, K., Wellink, J., Porta, C., Taylor, K. M., Lomonossoff, G. P., and van Kammen, A. 2000. Engineering cowpea mosaic virus RNA-2 into a vector to express heterologous proteins in plants. Virology 267:159–173

Gottula, J. W. 2014. Grapevine fanleaf virus: biology, biotechnology and resistance. Ph.D. Dissertation. Cornell University

Jobling, S. A., and Gehrke, L. 1987. Enhanced translation of chimaeric messenger RNAs containing a plant viral untranslated leader sequence. Nature 325:622–625

Ryan, M. D., and Drew, J. 1994. Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. EMBO J. 13:928–933