DISSERTATION

MANAGING POTATO TUBER NECROTIC VIRUSES: POTATO VIRUS Y AND POTATO MOP-TOP VIRUS, DIVERSITY, DETECTION AND CLONING METHODS

Submitted by

Jennifer Rushton

Department of Agricultural Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2025

Doctoral Committee:

Advisor: Vamsi Nalam

Amy Charkowski
Mark Stenglein
Robyn Roberts

ABSTRACT

MANAGING POTATO TUBER NECROTIC VIRUSES: POTATO VIRUS Y AND POTATO MOP-TOP VIRUS, DIVERSITY, DETECTION AND CLONING METHODS

Potato virus Y (PVY) and Potato mop-top virus (PMTV) are major threats to potato production globally, significantly affecting crop quality and yield. PVY, a well-known economic concern, poses a challenge due to its wide strain diversity and prevalence, particularly in the San Luis Valley (SLV), Colorado. PMTV, a single-stranded positive-sense RNA virus from the Pomovirus genus, is vectored by the soil-borne pathogen *Spongospora subterranea f. sp. subterranea (S. subterranea),* which also causes potato powdery scab. Despite their importance, there is still limited knowledge of PMTV genetics and the population structure of *S. subterranea*, as well as the combined virulence of these pathogens in the United States.

This study first investigated PVY trends and diversity in the San Luis Valley, Colorado using Colorado Seed Potato Certification Program data (2017-2022). PVY incidence fluctuated between 2.6% (2017) and 6.5% (2022) in certified seed lots, with PVY^{N-Wi} identified as the dominant strain. A field study from 2021 to 2022 further examined PVY prevalence in 600 samples collected across 30 fields, confirming PVY^{N-Wi} as the predominant strain, followed by PVY^{O} and PVY^{NTN}. The study also noted an increase in dual and triple infections, and high-throughput sequencing (HTS) of inconclusive samples

revealed polymorphisms that may hinder strain typing by traditional methods. Aside from PVY, potato virus S (PVS) was the only other potato-infecting virus detected. HTS additionally uncovered diverse non-plant infecting viruses, emphasizing the complexity of the potato virome.

In parallel, the study addressed the knowledge gaps surrounding PMTV and its vector *S. subterranea*. Sequencing of seven PMTV isolates showed low genetic variability, with one strain, ND24, containing a non-synonymous SNP in the transmembrane domain at codon 289, correlating with a mild strain instead of severe. Additionally, this work is the first to suggest diversity in *S. subterranea* populations across the United States with the use of microsatellite markers. Moreover, the study found that PMTV's low genetic diversity had no significant effect on its virulence in controlled conditions. Instead, the origin of *S. subterranea* inoculation was the greatest determinant of virulence. Primers were developed for each RNA segment of PMTV's tripartite genome, with TGB being the most abundant in both plant bioassays and sequencing read counts. Interestingly, no correlation between PMTV and *S. subterranea* counts was found in soil samples, but a clear relationship emerged in plant bioassays, suggesting that environmental factors and potato cultivars are critical in determining field infections.

This research provides valuable insights into the genetic and population structures of PVY and PMTV, highlighting the need for ongoing surveillance and improved diagnostics to manage these viral threats. The findings underscore the importance of viral diversity, vector dynamics, and environmental influences in shaping the impact of viral infections on potato production.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my graduate advisor, Dr. Vamsi Nalam. Without your support, optimism, and understanding, I would not have been able to pursue and complete this degree. Additionally, I would like to extend my sincere thanks to my other committee members: Dr. Amy Charkowski, Dr. David Holm, Dr. Robyn Roberts and Dr. Mark Stenglein. I appreciate your feedback and contributions.

Special thanks to the members of the Nalam Lab, past and present. It hasn't always been easy being a part of a new lab but the challenges have bought us closer. Another thanks to other labs in the AgBio department (Nachappa lab, Leach lab, Argueso lab) that have helped me along the way, whether it was sharing equipment or advice during troubleshooting.

I am honored to have been funded by the USDA-NIFA predoctoral fellowship since June 2023. It has allowed me to continue my research and has offered financial support.

Lastly, I want to thank my family and friends who have supported me throughout my education. A special thank you to my partner Sam Sillen, who has given me unwavering support and warm meals. Additionally, a shoutout to my biggest fan and best friend, Rachel Pappas; she encouraged me to pursue my passion as a scientist and is the reason I ended up at Colorado State University.

TABLE OF CONTENTS

CHAPTER 1- INTRODUCTION

Potato viruses pose a significant threat to global potato production (Chikh-Ali & Karasev, 2023; Karasev & Gray, 2013). Producers face reduced yield and quality due to the challenges of managing complex viruses, exacerbated by the absence of resistant cultivars (Karasev et al., 2010; Whitworth et al., 2009). The vegetative propagation of potatoes makes them particularly susceptible to viral accumulation over time, complicating efforts to maintain sustainable and profitable production. Additionally, the indirect financial impacts of these viruses include increased production costs and expenses related to disease management (Frost et al., 2013). To prevent viral accumulation and limit their spread, detection is a necessary tool for management. Effective virus detection and resistance are crucial for growers and the future of the potato industry (Chikh-Ali et al., 2020). Seed lots, growers, and breeding programs struggle with vector-borne tuber-necrotic viruses, such as the aphid-transmitted potato virus Y (PVY) and the plasmodiophorid-transmitted potato mop-top virus (PMTV).

The detection of viruses such as PVY and PMTV is crucial for limiting their spread, particularly within seed certification programs, which seek to minimize the viral load in seed tubers. Seed certification programs previously relied on visual assessment of foliar symptoms to determine PVY incidence, however this method was problematic given the rise of PVY strains that lack foliar symptoms (Frost et al., 2013). Specifically, the tuber necrotic strains $PVY^{N\text{-}WI}$ and PVY^{NTN} have become the dominant strain over foliar symptoms inducing PVY^{O} in the western United States and Canada (Funke et al., 2017; MacKenzie et

al., 2018; Tran et al., 2022). PVY^{NTN} is a recombinant strain first reported in the United States in the 2000s and has been on the rise since (Gray et al., 2010; Green et al., 2017; Karasev et al., 2010). The rise of recombinant strains is a result of the viral genomes' ability to mutate and recombine at rapid rates. Characterizing the strains of PVY present in fields is important to discovering new recombinants, investigating strain prevalence, and assessing mixed infections that could give rise to recombination events. PVY's capacity to adapt and the shift to recombinant strains has rendered visual assessments less effective, necessitating more precise molecular detection methods. The most common tools for PVY detection include enzyme-linked immunosorbent assay (ELISA), immunocapture, and polymerase chain reaction (PCR). Among these, PCR offers the highest resolution for differentiating between virus strains, particularly in identifying recombinant forms of PVY (Chikh-Ali & Karasev, 2015). Understanding the diversity and prevalence of PVY strains is crucial for managing outbreaks, as strain-specific management strategies, such as *Ny* resistance genes, can be more effective. Given the importance PVY detection, we aimed to explore strain composition in the San Luis Valley, Colorado in a two-year field study. We hypothesized that we would find a recombination event considering PVY's adaptability. We investigated PVY prevalence and characterized strains in the San Luis Valley, Colorado. strains using ELISA, PCR and NGS technologies.

PMTV, another significant threat to potato production, is vectored by a soil-borne pathogenic plasmodiophorid, *Spongospora subterranea f. sp. Subterranea (S. subterranea),* the causal agent of potato powdery scab disease (Jones & Harrison, 1969). *S. subterranea* produces resilient resting spores that can survive in the soil for more 40 years

without the potato host, and PMTV can remain infectious in the long-lived resting spores of its vector (Kirk, 2008; Santala et al., 2010; Torrance, 2008). Current management strategies for PMTV are limited, as chemical controls are ineffective against both the virus and its vector, and genetic resistance to PMTV and *S. subterranea* remains scarce. Notably, recent research in the San Luis Valley suggests that the use of recommended chemicals and cultivars may even increase soil inoculum levels (Zeng et al., 2020). This study, along with several others, raise important questions about the factors influencing the soil inoculum required for PMTV tuber infections (Davey et al., 2014; Merz, n.d.; Shah et al., 2012.; Van De Graaf et al., 2005). Until recently, detection of PMTV in field soils was labor-intensive and required a bioassay that is both time-consuming and impractical for large-scale field studies. However, advancements in digital droplet PCR (ddPCR) have revolutionized the ability to detect PMTV at low levels in soil samples, providing a powerful tool for understanding the virus-vector relationship (Pandey et al., 2020). In our study, we used ddPCR to evaluate detection methods and the relationship of the pathogen in field soils in addition a plant bioassay. Additionally, we studied their relationship in regard to population and genome diversity. We hypothesized that the pathogens would have a correlation in soil and inoculated plants. Furthermore, we expected PMTV and *S. subterranea* to be genetically different based on location.

Although detection is vital to management and valuable to studying the epidemiology and pathogenicity of viruses, host resistance remains the most promising long-term strategy for managing PVY and PMTV. With a lack of resistance in commercial cultivars, breeding programs are actively trying to create resistant cultivars (Chikh-Ali et al., 2020;

Funke et al., 2017; Whitworth et al., 2009). There are three known *Ry* genes in potato that cause extreme resistance genes to PVY, however they are challenging to incorporate due to unwanted traits and the tetraploidy of the potato's genome (Jansky et al., 2016). To expand known resistance to both viruses, we proposed the use of full-length infectious clones tagged with a marker, *A. majus* Rosea1 MYB-transcription factor, to visually evaluate resistance. The Rosea1 transcription factor induces the expression of genes involved in the biosynthesis of anthocyanin (Cordero et al., 2017). In plants infected with a Rosea1-tagged viral clone, upon viral replication, the biosynthesis of anthocyanin is induced resulting in the accumulation of anthocyanins visible as red/dark pigmentation allowing for the visual tracking of virus movement. We hypothesized that full-length recombinant infectious viral clones could be used to track PVY and PMTV in the potato host for rapid resistance detection.

Overall, our study aimed to address virus management and vector dynamics from a molecular perspective, focusing on detection tools and cloning techniques. Over the course of two years, we conducted a field study in the San Luis Valley to investigate PVY strain diversity and recombination events, using PCR and high-throughput sequencing technologies. We also developed ddPCR-based tools to detect PMTV in field soils and employed molecular techniques to assess the genetic diversity of both PMTV and its vector. We aimed to gain a deeper understanding of PMTV's relationship with its host and vector, *S. subterranea,* by exploring genetic factors that could influence the virus-vector relationship. We assessed the populations individually through whole transcriptome sequencing (PMTV) and microsatellite markers (*S. subterranea)*. Although unsuccessful,

we explored various cloning techniques to address host resistance to PVY and PMTV in efforts to advance virus management. The main goal of our research was providing detection tools and to analyze genetic data that could aid in the future of potato tuber necrotic virus management. The specific objectives and hypothesis are as follows:

1. Characterize PVY strains and analyze mixed infections for recombination and virome in the San Luis Valley, Colorado
    - *Hypothesis: $PVY^{N\text{-}Wi}$ will be the most predominant strain with new recombination detected and other economically important potato viruses based on preliminary findings in post-harvest testing data showing mixed infections and inconclusive results.*
2. Sequence and analyze PMTV genomes and diversity
    - *Hypothesis: There would be some genetic diversity in new PMTV isolates due to the lack of PMTV isolates sequenced in the United States and greater diversity worldwide.*
3. Develop primers for PMTV detection and analysis of tripartite genome and relationship to vector
    - *Hypothesis: New primers used to detect all genome segments would be a useful tool for examining PMTV in soils and would allow us to see correlation between the virus and vector because of the sensitivity and accuracy of the detection tools.*
4. Construct recombinant infectious clones of PVY and PMTV and use clones for rapid resistance detection

- *Hypothesis: Rosea-1 tagged infectious clones would allow for rapid resistance detection because of the anthocyanin production of the MYB-transcription factor.*

REFERENCES

Chikh-Ali, M., & Karasev, A. V. (2015). Immunocapture-multiplex RT-PCR for the simultaneous detection and identification of plant viruses and their strains: Study case, potato virus Y (PVY). Methods in Molecular Biology, 1302, 177–186. https://doi.org/10.1007/978-1-4939-2620-6_14

Chikh-Ali, M., & Karasev, A. V. (2023). Virus diseases of potato and their control. In Potato Production Worldwide (pp. 199–212). Elsevier. https://doi.org/10.1016/b978-0-12-822925-5.00008-6

Chikh-Ali, M., Tran, L. T., Price, W. J., & Karasev, A. V. (2020). Effects of the age-related resistance to potato virus y in potato on the systemic spread of the virus, incidence of the potato tuber necrotic ringspot disease, tuber yield, and translocation rates into progeny tubers. Plant Disease, 104(1), 269–275. https://doi.org/10.1094/PDIS-06-19-1201-RE/ASSET/IMAGES/LARGE/PDIS-06-19-1201-RE_F7.JPEG

Cordero, T., Mohamed, M. A., López-Moya, J. J., & Daròs, J. A. (2017). A recombinant Potato virus Y infectious clone tagged with the Rosea1 visual marker (PVY-Ros1) facilitates the analysis of viral infectivity and allows the production of large amounts of anthocyanins in plants. Frontiers in Microbiology, 8(APR), 611. https://doi.org/10.3389/FMICB.2017.00611/FULL

Davey, T., Carnegie, S. F., Saddler, G. S., & Mitchell, W. J. (2014). The importance of the infected seed tuber and soil inoculum in transmitting Potato mop-top virus to potato plants. Plant Pathology, 63(1), 88–97. https://doi.org/10.1111/PPA.12074

Frost, K. E., Groves, R. L., & Charkowski, A. O. (2013). Integrated control of potato pathogens through seed potato certification and provision of clean seed potatoes. Plant Disease, 97(10), 1268–1280. https://doi.org/10.1094/PDIS-05-13-0477-FE/SUPPL_FILE/PDIS-05-13-0477-FEE.PDF

Funke, C. N., Nikolaeva, O. V., Green, K. J., Tran, L. T., Chikh-Ali, M., Quintero-Ferrer, A., Cating, R. A., Frost, K. E., Hamm, P. B., Olsen, N., Pavek, M. J., Gray, S. M., Crosslin, J. M., & Karasev, A. V. (2017). Strain-specific resistance to potato virus Y (PVY) in potato and its effect on the relative abundance of pvy strains in commercial potato fields. Plant Disease, 101(1), 20–28. https://doi.org/10.1094/PDIS-06-16-0901-RE/ASSET/IMAGES/LARGE/PDIS-06-16-0901-RE_F5.JPEG

Gray, S., De Boer, S., Lorenzen, J., Karasev, A., Whitworth, J., Nolte, P., Singh, R., Boucher, A., & Xu, H. (2010). Potato virus Y: An evolving concern for potato crops in the United States and Canada. Plant Disease, 94(12), 1384–1397. https://doi.org/10.1094/PDIS-02-10-0124

Green, K. J., Brown, C. J., Gray, S. M., & Karasev, A. V. (2017). Phylogenetic study of recombinant strains of Potato virus Y. Virology, 507, 40–52. https://doi.org/10.1016/j.virol.2017.03.018

Jansky, S. H., Charkowski, A. O., Douches, D. S., Gusmini, G., Richael, C., Bethke, P. C., Spooner, D. M., Novy, R. G., De Jong, H., De Jong, W. S., Bamberg, J. B., Thompson, A. L., Bizimungu, B., Holm, D. G., Brown, C. R., Haynes, K. G., Sathuvalli, V. R., Veilleux, R. E., Creighton Miller, J., … Jiang, J. (2016). Reinventing Potato as a Diploid Inbred Line–Based Crop. Crop Science, 56(4), 1412–1422. https://doi.org/10.2135/CROPSCI2015.12.0740

Jones, R. A. C., & Harrison, B. D. (1969). The behaviour of potato mop-top virus in soil, and evidence for its transmission by Spongospora subterranea (Wallr.) Lagerh. Annals of Applied Biology, 63(1), 1–17. https://doi.org/10.1111/j.1744-7348.1969.tb05461.x

Karasev, A. V., & Gray, S. M. (2013). Continuous and emerging challenges of potato virus y in potato. Annual Review of Phytopathology, 51, 571–586. https://doi.org/10.1146/annurev-phyto-082712-102332

Karasev, A. V., Nikolaeva, O. V., Hu, X., Sielaff, Z., Whitworth, J., Lorenzen, J. H., & Gray, S. M. (2010). Serological properties of ordinary and necrotic isolates of Potato virus Y: A case study of PVYN misidentification. American Journal of Potato Research, 87(1), 1–9. https://doi.org/10.1007/S12230-009-9110-2/TABLES/2

Kirk, H. G. (2008). Mop-top virus, relationship to its vector. American Journal of Potato Research, 85(4), 261–265. https://doi.org/10.1007/s12230-008-9021-7

MacKenzie, T. D. B., Lavoie, J., Nie, X., & Singh, M. (2018). Differential Spread of Potato virus Y (PVY) Strains O, N:O and NTN in the Field: Implications for the Rise of Recombinant PVY Strains in New Brunswick, Canada. American Journal of Potato Research, 95(3), 301–310. https://doi.org/10.1007/S12230-018-9632-6/FIGURES/2

Merz, U. (n.d.). Infectivity, inoculum density and germination of Spongospora subterranea resting spores: a soiution-culture test system'.

Pandey, B., Mallik, I., & Gudmestad, N. C. (2020). Development and Application of a Real-Time Reverse-Transcription PCR and Droplet Digital PCR Assays for the Direct Detection of Potato mop top virus in Soil. Phytopathology, 110(1), 58–67. https://doi.org/10.1094/PHYTO-05-19-0185-FI/ASSET/IMAGES/LARGE/PHYTO-05-19-0185-FI_T3.JPEG

Santala, J., Samuilova, O., Hannukkala, A., Latvala, S., Kortemaa, H., Beuch, U., Kvarnheden, A., Persson, P., Topp, K., Ørstad, K., Spetz, C., Nielsen, S. L., Kirk, H. G., Budziszewska, M., Wieczorek, P., Obrepalska-Steplowska, A., Pospieszny, H.,

Kryszczuk, A., Sztangret-Wiśniewska, J., ... Valkonen, J. P. T. (2010). Detection, distribution and control of Potato mop-top virus, a soil-borne virus, in northern Europe. Annals of Applied Biology, 157(2), 163–178. https://doi.org/10.1111/j.1744-7348.2010.00423.x

Shah, F. A., Falloon, R. E., Butler, R. C., & Lister, R. A. (2012). Low amounts of Spongospora subterranea sporosorus inoculum cause severe powdery scab, root galling and reduced water use in potato (Solanum tuberosum). https://doi.org/10.1007/s13313-011-0110-6

Torrance, L. (2008). Pomovirus. In Encyclopedia of Virology: Volume 1-5 (Vols. 1–5, pp. V4-282-V4-287). Elsevier. https://doi.org/10.1016/B978-012374410-4.00543-4

Tran, L. T., Green, K. J., Rodriguez-Rodriguez, M., Orellana, G. E., Funke, C. N., Nikolaeva, O. V., Quintero-Ferrer, A., Chikh-Ali, M., Woodell, L., Olsen, N., & Karasev, A. V. (2022). Prevalence of Recombinant Strains of Potato Virus Y in Seed Potato Planted in Idaho and Washington States Between 2011 and 2021. Plant Disease, 106(3), 810–817. https://doi.org/10.1094/PDIS-08-21-1852-SR/ASSET/IMAGES/LARGE/PDIS-08-21-1852-SRF3.JPEG

Van De Graaf, P., Lees, A. K., Wale, S. J., & Duncan, J. M. (2005). Effect of soil inoculum level and environmental factors on potato powdery scab caused by Spongospora subterranea. Plant Pathology, 54(1), 22–28. https://doi.org/10.1111/J.1365-3059.2005.01111.X

Whitworth, J. L., & Crosslin, J. M. (2012). Detection of Potato mop top virus (Furovirus) on potato in southeast Idaho. Https://Doi.Org/10.1094/PDIS-08-12-0707-PDN, 97(1), 149. https://doi.org/10.1094/PDIS-08-12-0707-PDN

Whitworth, J. L., Novy, R. G., Hall, D. G., Crosslin, J. M., & Brown, C. R. (2009). Characterization of broad spectrum potato virus Y resistance in a Solanum tuberosum ssp. andigena-Derived population and select breeding clones using molecular markers, grafting, and field inoculations. American Journal of Potato Research, 86(4), 286–296. https://doi.org/10.1007/S12230-009-9082-2/TABLES/6

Zeng, Y., Fulladolsa, A. C., Cordova, A. M., O'Neill, P., Gray, S. M., & Charkowski, A. O. (2020). Evaluation of effects of chemical soil treatments and potato cultivars on spongospora subterranea soil inoculum and incidence of powdery scab and potato mop-top virus in potato. Plant Disease, 104(11), 2807–2816. https://doi.org/10.1094/PDIS-10-19-2202-RE/ASSET/IMAGES/LARGE/PDIS-10-19-2202-REF5.JPEG

CHAPTER 2 – DETECTION AND CHARACTERIZATION OF POTATO VIRUS Y STRAINS AND MIXED INFECTIONS IN SAN LUIS VALLEY, COLORADO

INTRODUCTION

Potato virus Y (PVY, *Potyvirus*) is the most important virus affecting potatoes worldwide. Described for the first time in the 1930s (Smith 1931), the virus is a significant threat to seed and commercial potato production (Kreuze et al. 2020). The virus also has a relatively wide host range and is considered a serious pathogen on other agriculturally important crops in the Solanaceae family (Kerlan and Moury 2008). Potatoes are the world's third most important crop (FAOSTAT, 2022), contributing to global food security due to their adaptability and ability to produce high yields per unit area (Devaux et al. 2020). Managing PVY is challenging due to its non-persistent mode of transmission by over 66 species of aphids (Pelletier et al. 2012; Pitt et al. 2022; Sigvald 1987, 1990). The virus particles translocate to progeny tubers, and seed potatoes harvested from the aphid-infected plants give rise to infected plants. Management is further complicated by PVY's genetic diversity with many newly emerged strains and recombinants, and the introduction of cultivars displaying mild foliar symptoms (Davidson et al. 2013; Gray and Power 2018). Infection reduces yield and tuber quality, with every 1% PVY incidence reducing yield by 1.6 cwt/A (Nolte et al. 2004). PVY levels are responsible for most of the rejection of seed lots for recertification (Chikh-Ali et al. 2020). Given that PVY is incurable once present in a field, standard practices involve preventive measures such as genetic resistance, hindering virus transmission, using certified seed tubers, and removing infected plants. The success of these phytosanitary actions hinges on early, precise, and strain-specific detection

methods. Despite recent progress in detection and molecular techniques (Davidson et al. 2013; Gray and Power 2018), current procedures have not proven effective in significantly reducing PVY incidence.

The genome of PVY is a positive single-stranded RNA genome of about 9.7 kb encoding one large polyprotein and one small protein that is processed into ten functional proteins, including the viral coat protein. Potyviruses are well-known for their evolutionary adaptability, evolving gradually through mutation accumulation and more rapidly through recombination. During recombination, different PVY strains within a single host can exchange substantial genome segments, some spanning several hundred nucleotides. This ability has plausibly granted them several selective advantages in specific hosts and diverse environments (Gibbs and Ohshima 2010; Gibbs et al. 2017). PVY has certainly benefitted from this ability and currently exists in nature, as a complex of strains and genetic variants, designated and classified based on genome sequences and recombination patterns (Glais et al. 2017; Green et al. 2018). These strains exhibit significant variability at biological, serological, and molecular levels. Some cause severe foliar symptoms, while others induce mild foliar symptoms but severe tuber necrosis (Lacomme et al. 2017). Sequencing efforts have shed light on PVY's phylogenetic lineages and the structures of recombinant genotypes (Green et al. 2017; Green et al. 2018; Karasev et al. 2010; Lacomme et al. 2017). Notably, genetic variability within strains and the identification of 'rare' recombinant structures add complexity to the PVY landscape (Green et al. 2017; Green et al. 2018). Currently, over 39 defined strains of PVY exist, with five non-recombinant strains (PVY^{O}, PVY^{N}, PVY^{Na-N}, PVY^{O5}, and PVY^{C}). The remaining exhibit

unique recombinant genomes formed by combinations of two or more parental genomes (Green et al. 2018; Karasev et al. 2010; Lacomme et al. 2017). Not all PVY strains can infect potato, and most potato-adapted strains belong to two lineages, PVY^{O} and PVY^{N}, or their recombinants (Green et al. 2017; Green et al. 2018; Karasev and Gray 2013; Quenouille et al. 2013). These recombinants are particularly concerning because they are linked to the potato tuber necrotic ringspot disease (PTNRD) in susceptible potato cultivars (Chikh-Ali et al. 2010; Glais et al. 2002; Gray et al. 2010; Romancer et al. 1994).

There has been a major shift in PVY strain composition in potato fields over the past few decades. Previously, the non-recombinant PVY^{O} strain, causing primarily foliar symptoms without tuber necrosis was prevalent (Ellis et al. 1997). State-run seed potato certification systems successfully managed its spread by rogueing out symptomatic plants (Gray et al. 2010). However, this practice inadvertently contributed to the shift in PVY strain composition, leading to an increase in prevalence of strains that cause mild to no foliar symptoms but can cause PTNRD (Funke et al. 2017; Gray et al. 2010; Gray and Power 2018). Consequently, some recombinant strains, especially PVY^{NTNa} and $PVY^{N\text{-}Wi}$, which were previously a minor part of the PVY complex, have become dominant in the western US and Canada (Funke et al. 2017; MacKenzie et al. 2018; Tran et al. 2022). These two strains now account for 98% of all PVY positives from 2011 – 2021 in the Pacific Northwest (Tran et al. 2022), mirroring trends in other regions like the San Luis Valley in Colorado (Colorado Potato Certification Service Annual Reports) and Canada (MacKenzie et al. 2018). Host plant resistance to PVY exists in the potato germplasm and is being used by breeding programs as a tool to manage PVY. Three genes, Ry_{chc}, Ry_{adg}, and Ry_{sto}, derived

from wild potato confer extreme or broad-spectrum resistance to multiple strains and several commercial cultivars are now available that carry the *Ry* genes and have proven PVY resistance (Novy et al. 2017; Kasai et al. 2000; Heldak et al. 2007). Additionally, resistance conferred by hypersensitive response-inducing *Ny* genes offer strain-specific resistance (Valkonen et al. 2017) The strain-dependence, as seen with Ny_{tbr} effectiveness against PVY^{O} (Funke et al. 2017), may have inadvertently selected for the emergence of diverse recombinant strains. Therefore, accurate knowledge of prevalent PVY strains in a region is critical for effective management. It guides the choice of diagnostic tools, selection of potato cultivars with appropriate resistance genes, and breeding efforts for developing new resistant varieties (Funke et al. 2017; Karasev and Gray 2013).

The varying symptomology caused by PVY strains makes in-season field inspections based on foliar symptoms unreliable. Serological techniques such as ELISA (enzyme-linked immunosorbent assay) have been widely used for PVY detection (Rosenman et al. 2019) and can distinguish between some major serotypes (PVY^{O} and PVY^{N}) but lack strain-specificity. To address these limitation, molecular methods, such as RT-PCR, targeting recombination junctions have been developed and incorporated into multiplex RT-PCR reactions and immunocapture assays (Chikh-Ali et al. 2013; Chikh-Ali and Karasev 2015; Funke et al. 2017; Mallik et al. 2012; Tran et al. 2022). While these methods increase throughput (MacKenzie et al. 2018; Whitworth et al. 2012), they rely on accurate sequence information for primer design and can miss novel recombinations due to potential sequence changes in evolving PVY strains (Green et al. 2018). High throughput sequencing (HTS) offers a powerful alternative. It goes beyond strain typing by enabling whole-genome

sequencing of PVY (Della Bartola et al. 2020; Gallo et al. 2021; Glasa et al. 2021; Gutiérrez et al. 2021). HTS also allows for the identification of novel recombinants, confirmation of mixed infections and even the detection of multiple or mixed additional viruses within a single plant (Gutiérrez et al. 2021; Jo et al. 2020; Lai et al. 2022) which could impact the evolution and epidemiology of multiple viruses (Tollenaere et al. 2016).

The shift to predominantly recombinant strains and the possibility of new recombinants prompted a comprehensive investigation of PVY diversity in the San Luis Valley, Colorado. Between 2017 and 2022, we characterized PVY strains and mixed infections HTS. Samples from post-harvest testing (2017 – 2022) were tested using ELISA and multiplex RT-PCR, revealed $PVY^{N\text{-}Wi}$ as the dominant strain and increasing trend of mixed infection (from 2% to 6%) highlighting the need for robust strain typing. HTS analysis not only confirmed these findings but also identified mixed infections and polymorphisms in primer binding regions that hindered accurate strain typing by conventional methods. Additionally, HTS data suggested a high potato virus S (PVS) titer in 2021, potentially contributing to exacerbated PVY symptoms that year. This study underscores the value of HTS for comprehensive PVY detection and the importance of continuous monitoring to provide information needed to adapt management strategies against evolving viral threats.

MATERIALS AND METHODS

Seed potato certification data

Data from winter post-harvest tests was collected by the Colorado Potato Certification Service (CPCS), located at the San Luis Valley Research Center of Colorado State University. Each year, between 2017-2022, the number of lots and plants that were visually inspected, and the number of plants with mosaic symptoms were recorded. For the winter post-harvest tests, seed tubers (a typical sample size is 400 tubers per seedlot; for seed lots <1 acre, a smaller sample is required) were kept warm and treated with Rindite to break dormancy before they were planted in Oahu, Hawaii, around the first week of November. The plants were visually inspected in December or January. A subsample of the plants showing mosaic symptoms was analyzed using ELISA and PCR (as described below) to identify the PVY strain causing the mosaic symptoms.

Field Sample Collection

For the field portion of this study, a total of 600 potato leaf samples were collected from various locations in the San Luis Valley (SLV), Colorado, during the first week of August in both the 2021 and 2022 growing seasons. A randomized zig-zag pattern was used to select the uppermost five leaflets from 20 individual potato plants per field to ensure representative sampling. In 2021, nine fields were sampled, while in 2022, 20 fields were sampled for a more extensive assessment. After collection, the harvested samples were placed in small plastic bags and immediately transferred to ice containers in the field, ensuring minimal deterioration. Within the hour, the samples were flash-frozen in liquid

nitrogen and stored at -80°C to maintain their integrity for subsequent processing. In addition to the samples collected in Colorado, 103 symptomatic samples were received from the Colorado Potato Certification Service's post-harvest test conducted in Hawaii and placed in the -80°C. For subsequent analysis, 5 gm of each sample was homogenized using blue pestles in liquid nitrogen to ensure thorough and consistent grinding of the samples. The resulting ground samples were utilized for ELISA assays (described below) or stored at -80°C for later investigations. This ensured the samples retained their RNA integrity and inherent characteristics to enable subsequent analysis and verification.

ELISA

For all samples, i.e. those collected during the post-harvest tests and those collected from fields during the growing season, the original 5 gm sample was divided into two 200 mg of plant tissue, of which one was used for the ELISA assay. The polyclonal TAS-ELISA for PVY was conducted according to the instructions provided by the manufacturer for both in-field and grow-out samples (Agdia, Inc, Elkhart, IN, SRA 20001). For both tests, 200 mg of homogenized ground sample was used in the initial step. Negative controls using healthy plant tissue were placed at the beginning and end of the plate. Positive controls were provided from the kit and used below negative controls. After the final incubation period, absorbance was measured at 405 nm on a plate reader (Synergy 2, BioTek®), and samples were considered positive if their absorbance reading was double that of the negative control on the plate. Positive results were calculated and recorded.

RNA extraction and cDNA synthesis

A 100 mg aliquot of ELISA-positive samples was used for subsequent RNA extraction using the Zymo Direct-zol™ RNA Miniprep Kit (Zymo, Irvine, California). Extraction was performed as per the manufacturer's protocol. Briefly, 600 µl of TRI Reagent® was added to the ground leaf tissue, centrifuged, and 600 µl of the supernatant was transferred to a new tube. An equal volume of 95% ethanol was added to the supernatant, and the resulting mixture was transferred to a Zymo-Spin™ IICR Column. The column was centrifuged at 12,000 rpm for 1 minute, and the resulting flow-through was discarded. Following this step, 400 µl of wash buffer was added to the column, and another round of centrifugation was performed. DNase treatment was carried out by incubating the column for 15 minutes at room temperature. Subsequently, the column was washed twice with 400 µl of pre-wash buffer, followed by a final wash with 700 µl of buffer. The tubes were centrifuged for an additional 2 minutes to ensure the complete removal of residual wash buffer. RNA was eluted from the column matrix using 50 µl of DNase/RNase-Free water. cDNA synthesis was performed using 1 µg of RNA using the Promega GoScript Reverse Transcription System kit per the manufacturer's protocol. RNA and 1 µl of Oligo(dT)15 primer, and 1 µl of random hexamers (0.5 µg) were incubated at 70°C for 5 minutes. Following this step, the tubes were placed on ice for 5 minutes. Additional components required for cDNA synthesis, including 4 µl of GoScript™ 5X reaction buffer, 1.2 µl of $MgCl_2$, 1 µl of dNTPs, 20 units of Recombinant RNasin® ribonuclease inhibitor, and 1 µl of GoScript™ reverse transcriptase, were added. Nuclease-free water was subsequently added to achieve a final volume of 15 µl. The thermocycling protocol for cDNA synthesis was as follows: initial

incubation at 25°C for 5 minutes, reverse transcription at 42°C for 60 minutes, and finally, an inactivation step at 70°C for 15 minutes.

Multiplex polymerase chain reaction assays for PVY strain identification

The strain identity of ELISA PVY-positive PVY samples was determined using the multiplex RT-PCR methods developed by Chikh-Ali et al. (2013) and Lorenzen et al. (2006) for all samples mentioned in this study. To determine the PVY strain present in the ELISA-positive samples, multiplex RT-PCR assays were performed using the methodology outlined in Chikh-Ali 2013. cDNA was amplified to target the primary recombination sites across the PVY genome. The primers used in our study are listed in **Table 2.1** All primers were included in the PCR reaction mix at a concentration of 100 µM. The multiplex RT-PCR was carried out with GenScript 2x Taq Master Mix. The final reaction volume of 20 µl included 3 µl of cDNA. The temperature protocol was: 94 °C for 4 min, followed by 10 cycles of 94 °C for 30 s, 64 °C for 30 s, 72 °C for 90, and 10 cycles of 94 °C for 30 s, 62 °C for 30 s, 72 °C for 90 s, and 10 cycles of 94 °C for 30 s, 60 °C for 30 s. 72 °C for 90 s and a final extension at 72 °C for 5 min. PCR results were analyzed on a 2% agarose gel and visualized under UV light. The size of the PCR products was examined and compared to expected strain-specific sizes. Samples that showed non-specific or no banding patterns were subjected to Illumina sequencing.

Table 2.1 List of primers used in the study for the multiplex PCR assays and to confirm strain identity.

Primer name	Sequence-5 →3	Location	Source
n156	GGGCAAACTCTCGTAAATTGCAG	160f	Chikh-Ali et al. (2013)
o514	GATCCTCCATCAAAGTCTGAGC	515f	Chikh-Ali et al. (2013)
n787	GTCCACTCTCTTTCGTAAACCTC	770r	Chikh-Ali et al. (2013)
n2258	GTCGATCACGAAACGCAGACAT	2260f	Lorenzen et al. (2006)
o2172	CAACTATGATGGATTTGGCGACC	2169f	Lorenzen et al. (2006)
n2650c	TGATCCACAACTTCACCGCTAACT	2169r	Lorenzen et al. (2006)
o2700	CGTAGGGCTAAAGCTGATAGTAG	2678r	Chikh-Ali et al. (2013)
S5585m	GGATCTCAAGTTGAAGGGGAC	5578f	Lorenzen et al. (2006)
o6400	GTAACTCCTAAACAAATGGTGGTTC	6405r	Chikh-Ali et al. (2013)
n7577	ACTGCTGCACCTTTAGATACTCTA	7582f	Chikh-Ali et al. (2013)
YO3-8648	CTTTTCCTTTGTTCGGGTTTGAC	8635r	Schubert et al. (2007)
SeroN	GTTTCTCCTATGTCGTATGCAAGTT	8864r	Chikh Ali et al. (2008)
o2439c	CCCAAGTTCAGGGCATGCAT	2467r	Lorenzen et al. (2006)
n5707	GTGTCTCACCAGGGCAAGAAC	5707f	Lorenzen et al. (2006)
A6032m	CTTGCGGACATCACTAAAGCG	6013r	Nie and Singh (2002)
5-17of	GCAGAAAGTGCACGAGTTCG	1820f	This study
5-17or	ATGTATGCTCCAAATCCAAAACCG	6357r	
5-20ntnf	GAGTTGGAGTCTGACATTAAGCAC	2358f	
5-20ntnr	GACATCCAAGAGCGCTTTAGTG	5909r	
5-20nwif	GACCATGACACTCAAACGTG	2248f	
5-20nwir	GACGTCGCAGCTTGGTGATG	5537r	
3-6ntnf	GTTTGCTAATGATGAGTTGGAGTCT	2363f	
3-6ntnr	CAATTTGAGCTCCAGTGAGCG	5272r	
3-6of	ATAGGCAATAGCGGTGACCAA	1952f	
3-6or	CCCATACATGTTGATGAACCTCC	5665r	
20-20nwi	CTCCAATGGGACAGAGCAGAC	6723f	
20-20ntnr	GCTGTGGCGACAATGAAGAG	8931r	
20-20ntnf	GCTGTGGCGACAATGAAGAG	6789f	

High Throughput Sequencing and Virome analysis

A total of 15 samples from 2021 (12 from in-field and 7 from the grow-out test) and 17 samples (in-field) from 2022 were sent to Novogene for Illumina sequencing. RNA (1 µg) was submitted to Novogene (Novogene Corporation Inc., Sacramento, CA), where library preparation, quality measurements, and sequencing were performed. Constructed libraries were sequenced on an Illumina NovaSeq600 instrument. Bioinformatic analysis was performed on Qiagen's CLC Genomics workbench (23.0.2) following the pipeline shown in **Figure 2.2**. First, the adapters were trimmed, and low-quality reads were removed using Illumina pipeline 1.5-1.7. Then, reads aligning with the *Solanum tuberosum* genome (SolTub_3.0, INSDC Assembly GCA_000226075.1) were removed. *De novo* assembly was performed on the remaining reads. All *de novo* assembled contigs were blasted against a PVY database containing 500 isolates extracted from NCBI (February 16, 2023). A database had to be created because the standard NCBI database only uses one complete genome for each virus. Therefore, a special PVY database was created to cover all strains of the virus and to account for the similarity of many isolates. This allowed the pipeline to detect the most similar sequence already in GenBank. Stringent coverage (>50%) and percent identity (>70%) cutoffs were used to identify PVY presence in the samples. A percent identity cutoff of 70 and not 80 was chosen to allow for some flexibility in detection of possible recombinants. PVY strains tend to be between 80-85% similar when compared to each other. Accession numbers were based on bit score and used in an additional mapping step (Starchevskaya et al. 2023). Trimmed reads were mapped back to specific accession numbers identified from the PVY database Blast. Visual examination was used to determine the coverage, depth (number of mapped reads) and amino acid

similarity of the PVY genome (Starchevskaya et al. 2023). Amino acid similarity was of particular interest to visually examine for possible recombination, meaning consistency in percent identity throughout the entire genome was important, it also confirmed samples with a full-length PVY initially detected with a 9 kb contig. Further analysis was performed on samples with smaller contigs that aligned to PVY. If smaller contigs aligned to different PVY isolates or strains, these contigs were manually extracted and alignment was examined in NCBI. Gaps between contigs and poorly sequenced contigs were detected in this alignment. Primers were then designed based on these gaps to investigate the missing or incomplete sequences (**Table 2.1**).

The initial steps used to examine the potato virome were the same as outlined for PVY above. Adaptors and low-quality reads were trimmed using Illumina pipeline 1.5-1.7 in CLC Workbench. The trimmed reads were aligned to the host genome, *Solanum tuberosum* (SolTub_3.0, INSDC Assembly GCA_000226075.1) using CLC's alignment tool. The subsequent steps to determine the potato virome in the samples was as per Starchevskaya et al. (2023). Unaligned reads were subjected to *de novo* assembly, forming contigs. Contigs longer than 350 bp and coverage greater than 40% were used in BLASTn and BLASTx searches against NCBI databases. BLAST results annotated with the term "virus" and exhibiting a percent identity greater than 80% to a known virus were extracted. To determine viruses in each sample, targets were identified based on redundancy. Bit score accession numbers were used to further identify and visualize the detected viruses in each sample. The trimmed reads from the sample of interest were mapped to the specific accession numbers. Visual examination was also used to assess coverage and depth.

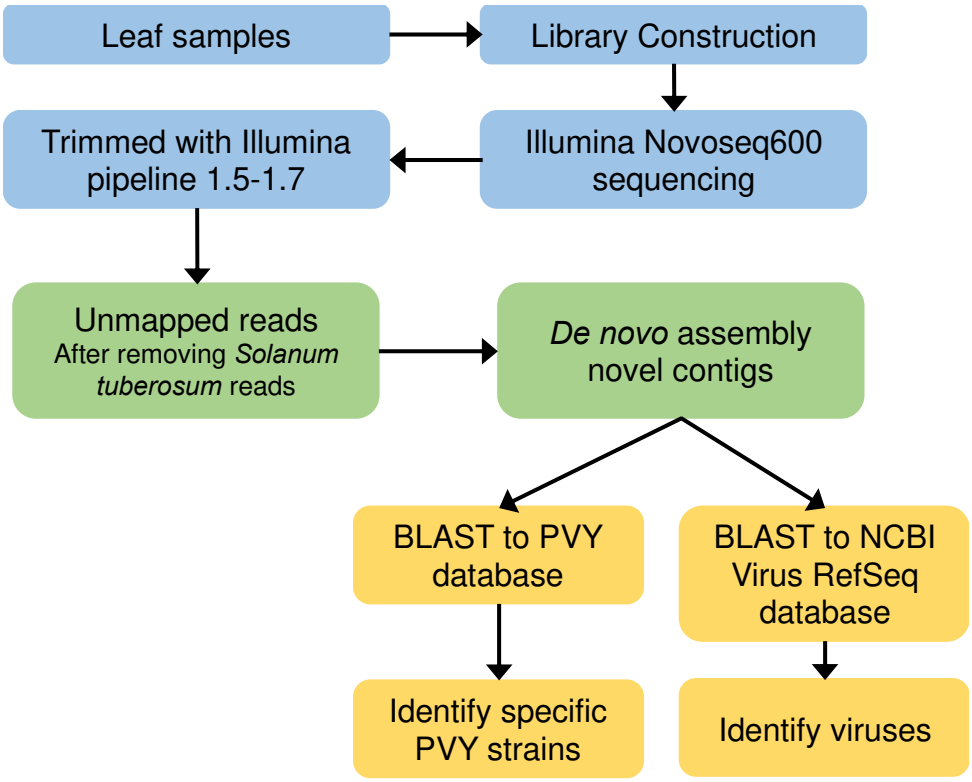

Figure 2.2 Pipeline for PVY and potato virome detection using bioinformatics analysis. Potato leaf samples were subjected to RNA extraction, short-read library construction, and Illumina Novoseq600 sequencing. Preprocessing and trimming of the raw sequencing data were performed using Qiagen's CLC workbench and Illumina's pipeline (version 1.5-1.7). De novo assembly and BLAST search against the PVY reference database (accessed February 2023) enabled the identification of specific accession numbers associated with reads exhibiting greater than 98% identity. Read mapping to the identified accession numbers allowed for sample correlation and assessment of PVY prevalence. Thorough inspection of alignments ensured full-length sequence, coverage, and depth, confirming PVY identification and providing insights into the virus's genetic diversity. A similar approach was employed for potato virome analysis, utilizing de novo assembled contigs and a BLAST (nr and nt) search against the NCBI Virus RefSeq database (accessed March 2023). Virus identity confirmation followed the same steps as described for PVY strains.

RT-PCR on Mixed Infections

RT-PCR reactions were performed to fill the gaps between contigs and poorly sequenced regions in four samples. The poorly sequenced regions in these samples was problematic because it prevented assembly and we could not determine if there were two separate strains and/or a recombinant strain present. To address this region, primers were created to span the gap which accounted for the two strains separately. Q5® High-Fidelity DNA Polymerase (New England Biolabs) was used as per manufacturer's recommendations and prescribed thermocycling conditions. Extension times varied based on the length of the intended product at 30s/kb. PCR products were run on a 2% agarose gel. The presence and size of products were examined to determine the missing sequences. Furthermore, the presence of PCR products in combination with primers was used to determine if the sample contained mixed infections or indicated the presence of novel recombination events.

Recombinant and phylogenetic analysis

We used RDP4.101 (Martin and Rybicki 2000) to perform recombinant analysis of the full-length isolates identified in this study to accurately identify isolate strain types and cross-verify them against previously established phylogenies. Each full-length isolate was subjected to individual analysis, comparing them with the genomes of potential parental strains, including at least one PVY isolate previously identified as belonging to each of the 5 non-recombinant genotypes (strains PVY^{O}, $PVY^{Eu\text{-}N}$, $PVY^{Na\text{-}N}$, PVY^{C} and PVY^{O5}). To detect recombination events, we used a set of six tests available in RDP4.101, including RDP, GENECONV, Chimaera, MaxChi, Bootscan and SiScan, all operating with default settings

(Gibbs et al. 2000; Martin and Rybicki 2000; Padidam et al. 1999; Posada and Crandall 2001; Salminen et al. 1995; Smith 1992). Additionally, whole genome UPGMA trees were generated in RDP4.101 to provide a general overview of how the new isolates clustered in relation to the reference PVY genomes from GenBank. Isolates classified as recombinants based on all six tests underwent further scrutiny. This included comparing the predicted recombination sites identified by RDP4.101 with known-recombinant sites to ensure no new recombination sites were present (Green et al. 2017).

The full-length isolates identified in the study and 26 reference genomes were aligned with MUSCLE in MEGAX (Della Bartola et al. 2020; Green et al. 2017). To confirm the UPGMA tree and validate the genotypes of the full-length isolates a phylogenetic tree was constructed using the maximum likelihood method. The alignment of the isolated and references genomes created using MUSCLE was used to create the tree in MEGAX. A Bootstrap value of 1000 and the GTR+G+I substitution model was used.

RESULTS

PVY incidence and strain identity in post-harvest testing

Data records of seed potato certification from the CPCS were evaluated to get an overview of PVY incidence in the SLV between 2017 – 2022, revealing that mosaic disease, commonly caused by PVY, was a persistent problem. The acreage represented in the post-harvest tests ranged from 8,744 in 2017 to 6,587 acres in 2023. The mosaic percentage for all lots in the SLV each year also showed fluctuations; for instance, the mosaic percentage increased from 2.6% in 2017 to 5.4% in 2018. In 2019, the mosaic percentage dropped to 2% before progressively increasing to 6.5% in 2022 (**Figure 2.3A**).

$PVY^{N\text{-}Wi}$ was the dominant strain in all years of post-harvest data, fluctuating, with a peak of 68% in 2019 and a low of 38% in 2020 (**Figure 2.3B**). The second most prevalent strain was PVY^{O}, which demonstrated a variable pattern, starting at 20% in 2017, reaching a maximum of 32% in 2021, and ending at 16% in 2022. PVY^{NTN} prevalence remained relatively low in all years, fluctuating between 5% and 10%. $PVY^{N:O}$ had minimal representation, with sporadic occurrences of 1% and 2% in 2017 and 2018, respectively. 2021 it exhibited a slight increase to 3%, but it was absent in 2020 and 2022. Strain $PVY^{NE\text{-}11}$ was generally absent, except for a minor presence of 1% in 2021. Dual infections with two different strains of PVY were detected in several samples. Such infections significantly rose from 2% in 2017 to 19% in 2020. The combinations of strains found were PVY^{O} with $PVY^{N\text{-}Wi}$, $PVY^{N:O}$ with $PVY^{N\text{-}Wi}$, PVY^{NTN} with $PVY^{N\text{-}Wi}$, and PVY^{O} with PVY^{NTN}. In 2% of the samples in 2017, 1% in 2019, 4% in 2020, and 6% in 2022, three different strains of PVY were detected in individual plant samples (**Figure 2.3B**). Two different combinations of PVY

strains were detected: (1) PVY^{O} + $PVY^{N\text{-}Wi}$ + PVY^{NTN}, and (2) PVY^{NTN} + $PVY^{N:O}$ + $PVY^{N\text{-}Wi}$. Furthermore, one sample from 2019 contained four different strains: PVY^{NTN} + $PVY^{N\text{-}Wi}$ + $PVY^{N:O}$ + PVY^{O}. In most years, several samples were detected where strain identity could not be determined using the multiplex RT-PCR (**Figure 2.3B**, inconclusive samples). In 2017, strain identity could not be identified in 4% of the samples, which declined to 3% and 1% in 2018 and 2019, respectively, and then increased to 7% in 2022 before dropping again to 2% in 2022. A subset of samples displayed mosaic symptoms and showed a positive ELISA test result; however, in these samples, no banding patterns were observed after the multiplex RT-PCR **(Figure 2.3B**, negative samples) (Chikh-Ali and Karasev 2015; Lorenzen et al. 2006). A steady increase was observed in their presence from 3% in 2018 to 25% in 2022, with significant growth observed in 2019 (7%), 2020 (13%), and 2021 (13%). The findings from the 5-year post-harvest data highlight the dynamic nature of strain prevalence over the years, emphasizing the importance of continuous monitoring and surveillance efforts to understand the evolving landscape of strains.

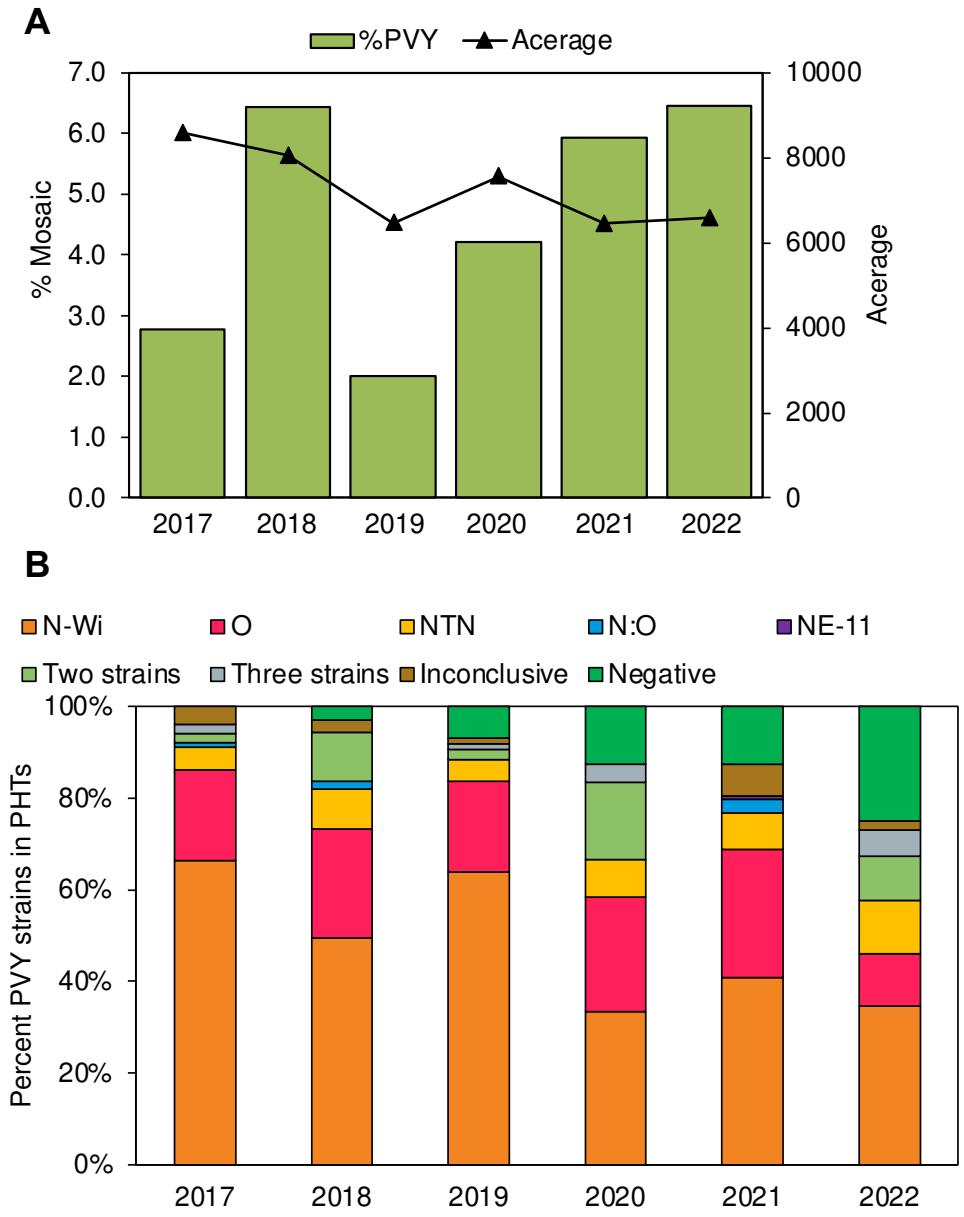

Figure 2.3 PVY prevalence and strain dynamics in the San Luis Valley over five years as viewed from post-harvest testing of certified seed.
(A) Prevalence of mosaic symptoms (bars) and the acreage of seed potatoes (line graph) represented by post-harvest testing during winter testing. (B) The incidence of PVY strains identified via multiplex PCR in a subset of samples from post-harvest testing.

In-season PVY Serological Detection and Strain-Typing

Given that PVY is transmitted by aphids during the potato growing season, we collected samples from commercial fields in 2021 – 2022 to identify prevalent strains, complementing the data records from the post-harvest testing and provide a comprehensive overview of the PVY strains present in the region. During in-season sample collection, it is important to note that we did not focus on only plants showing foliar symptoms in the field but instead prioritized random sample collection to account for in-season PVY presence post aphid-flight. First, ELISA was used to calculate the percent PVY incidence per field (**Table 2.2**). There was an overall higher incidence in 2021 compared to 2022. In 2021, field incidence ranged from 0-85%, averaging 26.5%. In 2022, field incidence ranged from 0-25%, with an average of 9.3%. The cultivars sampled between years differed, with 70% of the fields being Russet Norkotah selections in 2021 and only 30% in 2022.

Two multiplex RT-PCR methods were able to determine strain identity in 83.3% of the ELISA-positive samples (**Figure 2.4, Table 2.2**). In 2021, $PVY^{N\text{-}Wi}$ was the most prevalent strain identified in nearly half (49.1%) of the samples, followed by PVY^{O} (17%), $PVY^{N:O}$ (5.7%), PVY^{NTN} (7.5%), and PVY-NE-11 (1.9%). While $PVY^{N\text{-}Wi}$ remained the dominant strain in 2022, its incidence decreased to 39.5%. Similarly, PVY^{O} and $PVY^{N:O}$ incidence decreased to 13.2% and 5.2%, respectively. On the other hand, PVY^{NTN} and PVY-NE-11 incidence increased compared to 2021, with them being identified in 13.2% and 2.6% of the samples, respectively. Samples were identified as PVY-NE-11 based on specific banding patterns described by both the Chikh-Ali and Lorenzen multiplex assays. In both assays, PVY-NE-11

is differentiated by the absence of other bands, leaving only a single band at 633bp in the Chikh-Ali assay and 328bp band in the Lorenzen assay.

The number of inconclusive samples, samples in which multiplex PCR assays could not determine PVY strain identity, significantly increased from 18.9% in 2021 to 36.8% in 2022 (**Table 2.2, Figure 2.4**). In these cases, the observed banding patterns did not agree with the expected patterns derived from the genomic organization of PVY. Among the ELISA-positive samples tested in both years, 32 were categorized as 'inconclusive' due to either the complete lack of bands or unexpected banding patterns (**Figure 2.6A, Figure 2.5**). These inconclusive patterns could plausibly be attributed to low virus loads or due to the presence of novel recombinations. In 14 samples, no bands were observed. In 15 samples, a single band was observed at 441 bp, indicating the potential presence of strains such as E, $PVY^{N:O}$, $PVY^{N\text{-}Wi}$, PVY^{NTNa}, PVY^{NTNb}, $PVY^{NTN\text{-}NW}$, SYR-III, or 261-4. One sample showed a band at 633 bp, indicating the possible existence of the strains PVY^{N}, PVY^{E}, $PVY^{N:O}$, or PVY^{NTNa}. Another sample showed a solitary band at 853 bp, indicating the potential presence of PVY^{O}, PVY^{O5}, $PVY^{N\text{-}Wi}$, $PVY^{N:O}$, or 261-4 strains. Lastly, a single sample exhibited three bands at 398, 532, and 853 bp, indicating the potential presence of multiple PVY strains.

Table 2.2 Percent incidence of various PVY strains.
PVY strains were detected using multiplex RT-PCR (Lorenzen et al., 2006; Chikh-Ali et al., 2010). Leaf tissue samples were collected from commercial fields in 2021 and 2022.

Year	Overall Incidence	PVY Strain					
		PVY^{N-Wi}	**PVY^{O}**	**$PVY^{N:O}$**	**PVY^{NTN}**	**PVY-NE-11**	**Inconclusive**
2021	26.5% (0 -85%)	49.1	17.0	5.7	7.5	1.9	18.9
2022	9.3% (0 – 25%)	39.5	13.2	5.3	13.2	2.6	36.8

Note: In 2021, 5 out of the 10 and in 2022, 6 out of the 20 fields were planted with Norkotah.

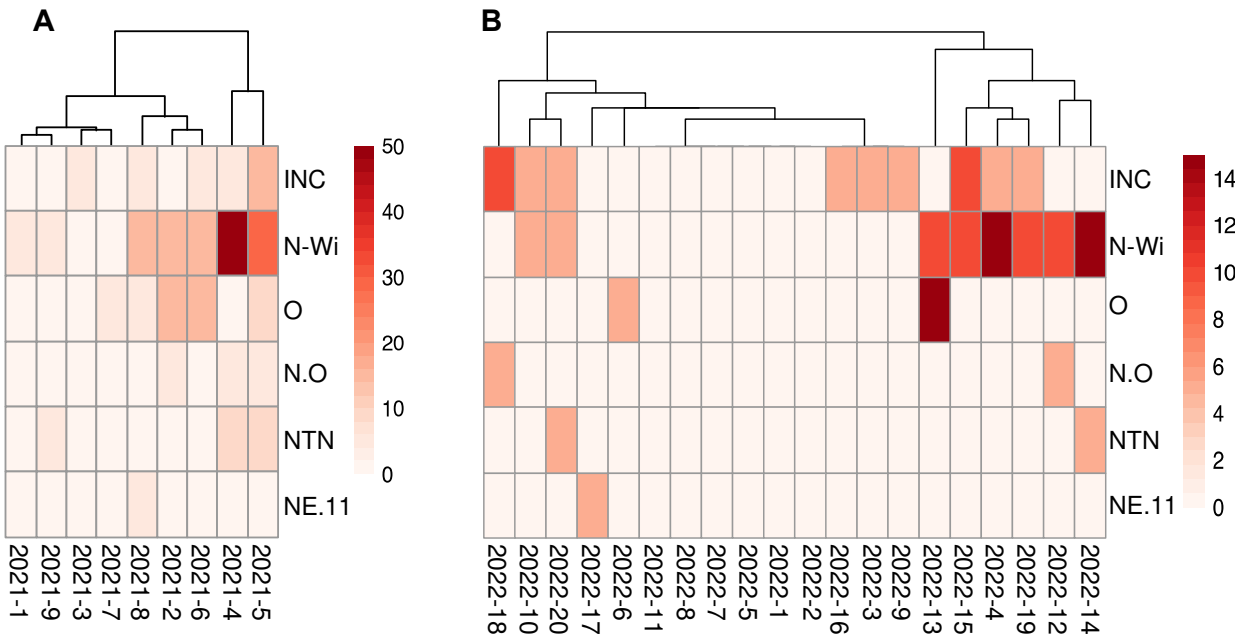

Figure 2.4 In-season strain diversity and prevalence of PVY. (A) 2021 and (B) 2022. Heatmap was generated based on the results from the multiplex-RT-PCR. The y-axis is the percent incidence, and the x-axis is the field. The fields are clustered based on similarity in prevalent strains. The intensity of the color in each cell represents the relative incidence of the corresponding strain, providing a clear illustration of strain prevalence within each field. “Inc” stands for inconclusive samples.

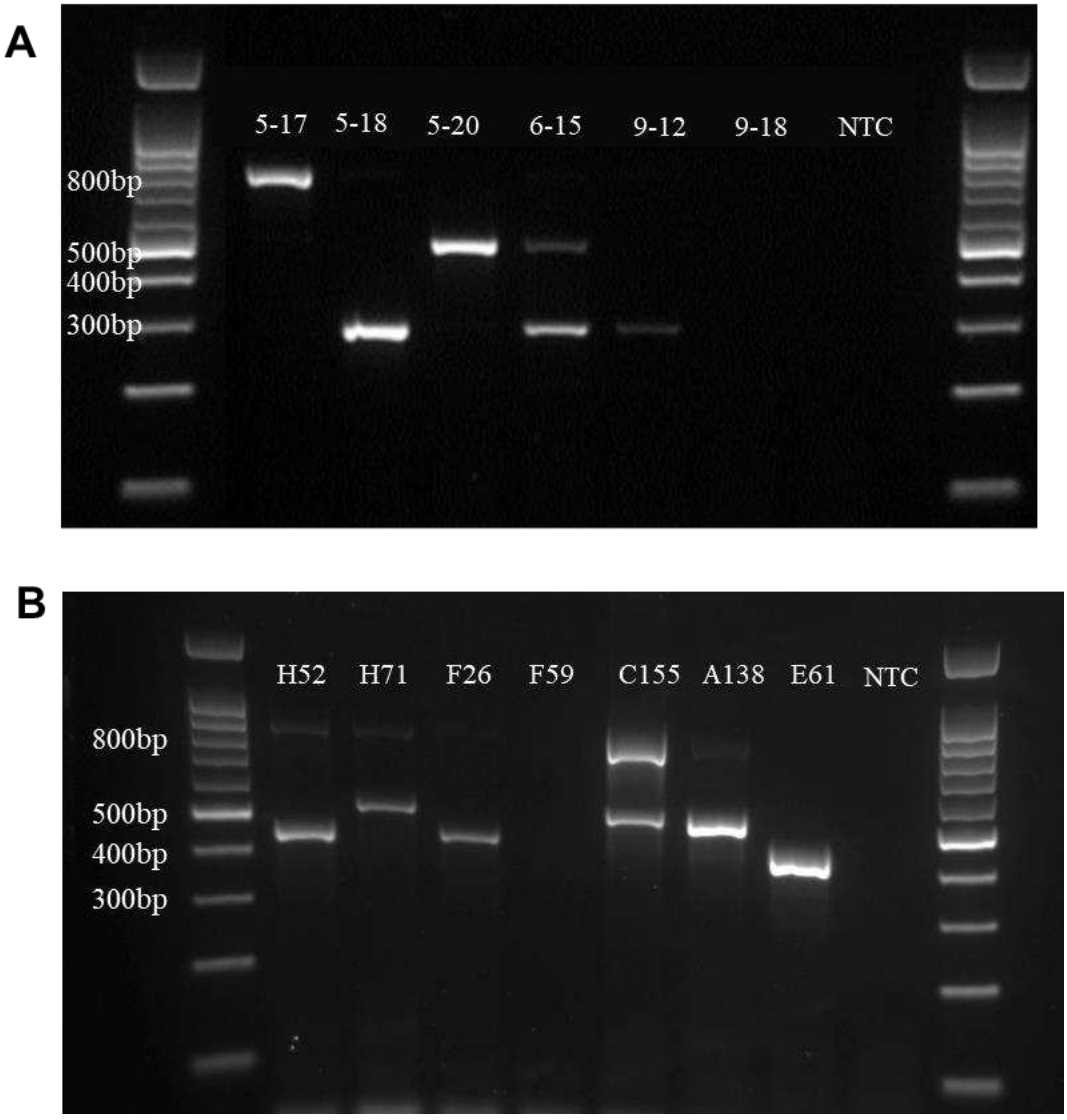

Figure 2.5 Results of Multiplex-RT-PCR on 2021 samples. A) Agarose gel showing the results of a multiplex-RT-PCR performed using primers from Lorenzen et al. (2006). B) Agarose gel showing results from a multiple-RT-PCR using primers from Chikh-Ali and Karasev (2015). The banding patterns represent inconclusive banding. In most cases there is a single band at 441bp which correlates to various PVY strains. Samples in this figure were subjected to HTS sequencing to investigate the unique banding patterns.

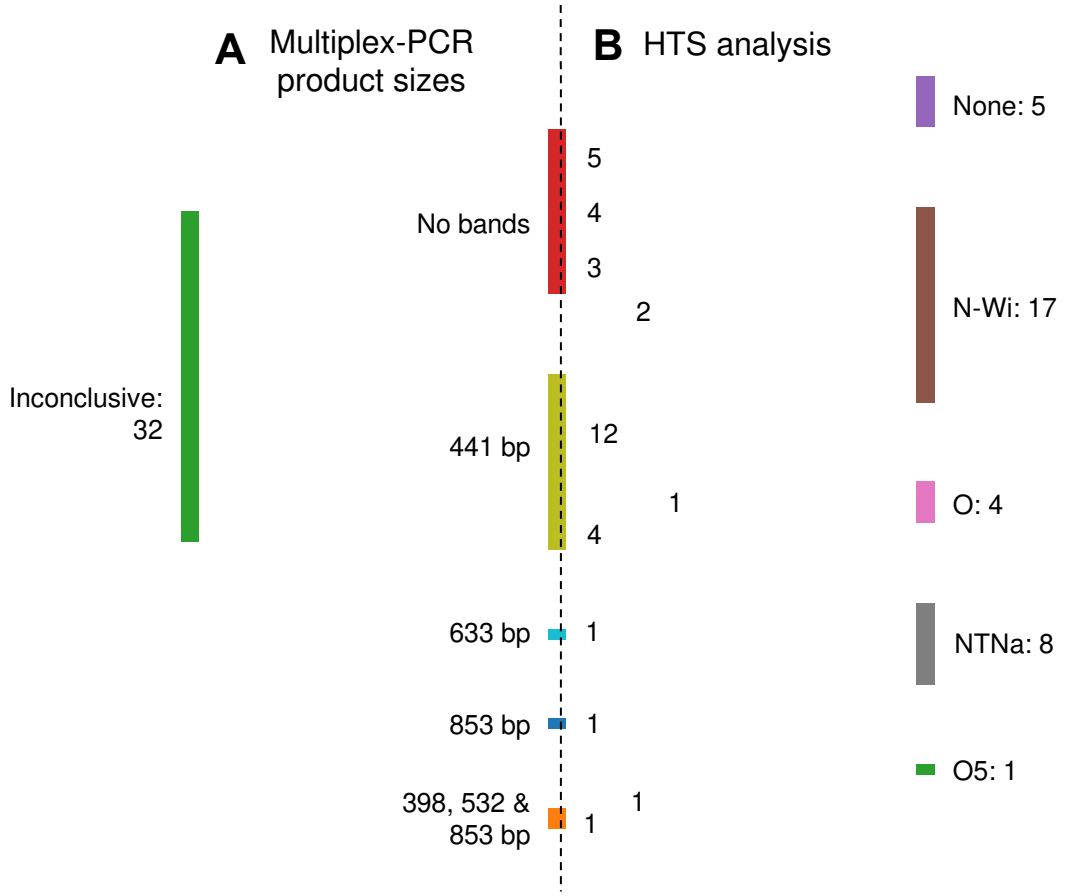

Figure 2.6 Sankey diagram showing strain identity of inconclusive samples. (A) The bars on the right represent the total number of inconclusive samples, and the bars on the dashed line represent the different band sizes observed in the 32 samples after multiplex-RT-PCR. (B) The bars on the right represent the strain identity determined after HTS analysis. The numbers next to the bars in the middle of the diagram represent the samples showing the specific strain associated with that specific banding pattern.

Complete genomes of PVY

To reconcile the discrepancies between the ELISA positive results and the lack of accurate strain identification by the multiplex RT-PCRs, nucleic acids from these thirty-two samples were sequenced. Potential factors such as novel recombinants, single or multiple nucleotide polymorphisms (SNPs or MNPs), or mixed infections may have influenced primer binding, leading to unexpected banding patterns. On average, 44 million paired-end reads were generated per sample (ranging from 39.6 to 56.5 million). Utilizing the bioinformatics pipeline (Figure 2.1), we determined that there was no PVY in five samples, confirmed by the lack of bands observed in the multiplex RT-PCRs (**Figure 2.5**), indicating potential false positives in the ELISA test. Among the remaining 27 samples, we identified four PVY strains: $PVY^{N\text{-}Wi}$, PVY^{O}, PVY^{O5}, and PVY^{NTNa}. Twenty three contained a single PVY strain, while four contained mixed infections with two strains (**Tables 2.2 & 2.3**).

In 18 samples, we successfully assembled a single large contig covering the full-length, or nearly full-length PVY genome. BLASTn analysis of the *de novo* assembled contigs against the custom NCBI PVY database revealed PVY strains with an identity of 99% or greater (**Table 2.3**). Of these, 14 samples were $PVY^{N\text{-}Wi}$, three were PVY^{NTNa}, and one was PVY^{O}. For $PVY^{N\text{-}Wi}$, the Chikh-Ali multiplex RT-PCR produces bands of 441 bp (primer combination n2258 + o2700) and 853 bp (primer combination of S5585m + o6400) (Chikh-Ali et al. 2013). However, 12 of the 14 $PVY^{N\text{-}Wi}$ samples exhibited only the 441 bp, while two had no bands (**Figure 2.5, 2.3).** Further analysis of the primer binding regions for S5585m and o6400 revealed the presence of an SNP in o6400, T/C_{6356} (**Figure 2.8**) in all 14 samples. Similarly, the banding patterns for the three PVY^{NTNa} samples were unexpected. Bands are

expected at 633, 441, and 1307 bp for PVY^{NTNa} (Chikh-Ali et al. 2013). However, samples 2021-4-18 and 2021-F26 displayed only a single 441 bp band and sample 2022-15-18, which has a SNP (T→ C transversion) in the o2700 primer binding region (**Figure 2.9**). Lastly, sample 2021-5-17, identified as containing PVY^{O}, did not exhibit any bands despite the expected bands of 532 and 853 bp, plausibly due to polymorphisms in the o6400 and o2700 primer binding regions.

Table 2.3 Analysis of high throughput sequencing (HTS) data from potato leaflet samples with complete genome sequence of PVY.

	Sample ID	Multiplex-RT-PCR band size	# reads mapping to PVY	% reads mapping to PVY	# contigs	Mean contig length	Mean Sequence Depth (coverage)[1]	Percent Genome Coverage[a]	Consensus Length or HSP*	Closest BLAST Relative (% of nt Identity)	PVY Strain	Isolate
1	2021-4-18	441 bp	1,160,855	9.7	6	3,529	94.48	100%	9,685	FJ204165.1 (USA) 99.9%	NTN	L26
2	2021-5-17	No bands	1,589,968	5.2	1	9,686	95.45	100%	9,686	EF026074.1 (USA) 99.7%	O	PVY-Oz
			346,514	1.4	14	1,608	98	77.8%	5987*	KY847950.1 (USA) 99.6%	N-Wi	ID125
3	2021-8-18	No bands	7,009,404	77.9	5	2,187	96.02	100%	9,098	KY847950.1 (USA) 99.8%	N-Wi	ID125
4	2021-9-12	441 bp	5,873,584	77.8	5	2,512	93.67	100%	9,810	KY847950.1 (USA) 99.8%	N-Wi	ID125
5	2021-3-8	441 bp	5,141,787	68.7	2	5,067	94.07	100%	9,768	KY847950.1 (USA) 99.9%	N-Wi	ID125
6	2021-A138	441 bp	2,890,910	14.0	9	1,530	95.28	100%	9,188	KY847950.1 (USA) 99.9%	N-Wi	ID125
7	2021-A143	441 bp	1,968,701	15.1	7	1,625	93.55	100%	9,021	KY847950.1 (USA) 99.8%	N-Wi	ID125
8	2021-A149	441 bp	505,628	12.9	3	3,542	94.63	100%	9,814	KY847950.1 (USA) 99.9%	N-Wi	ID125

9	2021-B109	441 bp	3,207,881	16.5	11	1,987	70.19	100%	9,798	KY847950.1 (USA) 99.8%	N-Wi	ID125
10	2021-B136	441 bp	1,539,451	3.9	13	1,135	94.79	100%	9,694	KY847950.1 (USA) 99.9%	N-Wi	ID125
11	2012-E61	441 bp	1,314,813	15.6	4	4,761	94.34	100%	8,528	KY847950.1 (USA) 99.8%	N-Wi	ID125
12	2021-F26	441 bp	4,091,986	31.2	7	1,329	94.28	100%	9,743	EF026075.1 (USA) 99.9%	NTN	PB312
13	2022-4-14	441 bp	398,686	29.2	1	9,721	94.53	100%	9,721	KY847950.1 (USA) 99.8%	N-Wi	ID125
14	2022-4-16	441 bp	90,498	7.6	1	9,702	94.71	100%	9,702	KY847950.1 (USA) 99.9%	N-Wi	ID125
15	2022-12-4	441 bp	7,572	0.1	5	3,892	94.7	100%	9,703	KY847950.1 (USA) 99.8%	N-Wi	ID125
16	2022-13-10	No Bands	2,620	0.1	1	9,696	94.77	100%	9,696	KY847942.1 (USA) 99.7%	N-Wi	ID125
17	2022-15-18	633 bp	528,129	20.1	1	9,710	94.6	100%	9,710	JF927763.1 (Poland) 99.7%	NTN	IUNG-15
18	2021-5-20	No bands	39,075	0.31	7	1723	99.32	100%	5125*	JQ969034.2 (Belgium) 99.8%	NTNa	GBVC-15
			95,695	0.93	8	1,871	99.77	98.9%	5118*	EF026074.1 (USA) 99.8%	O	PVY-Oz
19	2022-20-20	441 bp	2,030	0.1	1	2,561	100	100%	8,488	KY847950.1 (USA) 100%	N-Wi	ID125

			422	0.1	8	732	70.89	72.20%	1209*	FJ204164.1 (USA) 99.7%	NTN	N4
20	2022-3-6	398, 532 & 953 bp	886,741	2.7	3	4581	99.72	100%	4001*	KY847938.1 (USA) 98.9%	O5	CO86
			379,335	1.3	2	5121	98.2	100%	3580*	FJ204164.1 (USA) 99.7%	NTN	N4

[a] Percent genome coverage is based on the number of reads that map to the genome and not the contigs.

Table 2.4 Analysis of high throughput sequencing (HTS) data from potato leaflet samples containing partial genome sequences of PVY.

	Sample ID	Multiplex-RT-PCR band size	# reads mapping to PVY	% reads mapping to PVY	# contigs	Mean contig length	Contigs length range	Mean Sequence Depth (coverage)	% Genome coverage	Greatest HSP (contig length)	Closest BLAST Relative (% of nt Identity)	PVY Strain	Isolate
1	2021-5-18	No Band	1,367	0.07	10	678	384 - 1721	100	75.5	1721	HQ912890 .1 (USA) 99.5%	O	ID1_5_62A
2	2021-6-10	No Bands	2,020	0.12	7	1251	418 - 2707	95.5%	83.3	4409	HQ912896 .1 (USA) 100%	N-Wi	LR
3	2022-15-17	441 bp	2,405	0.09	4	1563	524-2602	94.8	33.3	2289	KY847946. 1 (USA) 99.8%	NTNa	ID12_102IC3
4	2022-16-3	No Bands	367	0.03	2	863	406-1321	94.6	15.2	1321	KY847946. 1 (USA) 99.7%	NTNa	ID12_102IC3
5	2022-17-3	No Bands	1,108	0.05	1	789	789	93.3	11.1	789	HQ912863 .1 (USA) 100%	N-Wi	N1
6	2022-18-12	853 bp	1,974	0.14	2	635	239-1032	95.6	22.3	1032	HQ912863 .1 (USA) 98.8%	N-Wi	N1
7	2022-19-2	No Bands	598	0.04	1	524	524	94.5	11.1	524	HQ912863 .1 (USA) 100%	N-Wi	N1

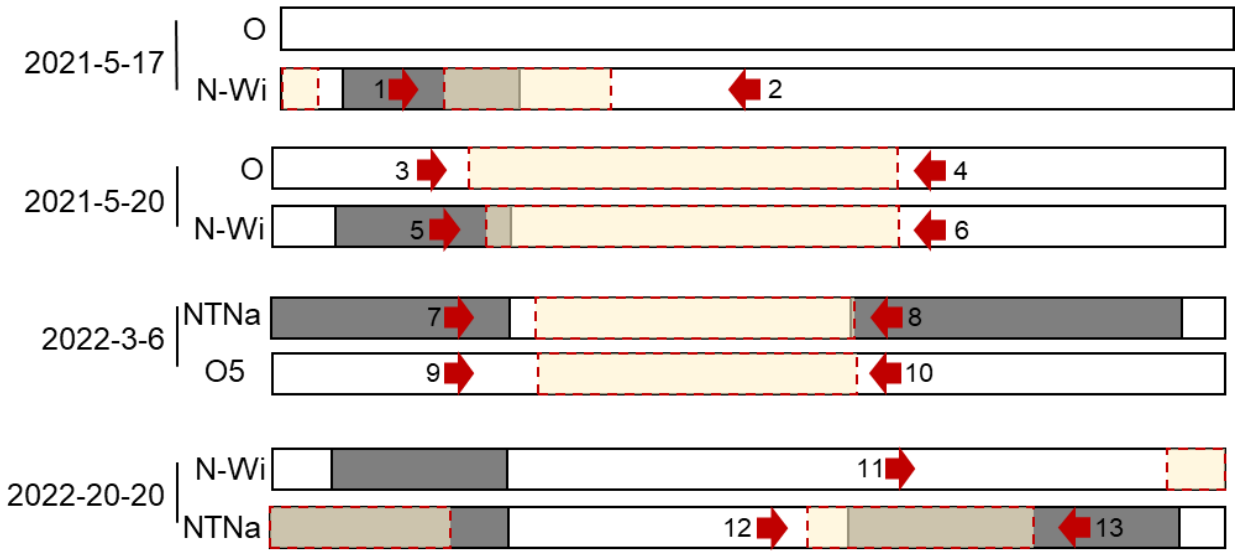

Figure 2.7 Samples with Mixed infections and the location of primers used to confirm strain identities.

Each sample is depicted with the two strains suspected in the sample. The genome of the strain is depicted with shaded and white areas to represent parent genomes that have created recombinants. The highlighted yellow regions are where contigs are missing—although reads may be present in that region. The remaining solid regions have contigs coverage and were ultimately used for primer design. Arrows show the locations of forward and reverse primers. A combination of primers was used for each sample to rule out recombination and to confirm mixed infections. Primers labeled are as follows: (1) 5-17of (2) 5-17or (3) 5-20ntnf (4) 5-20nwir (5) 5-20ntnf (6) 5-20ntnr (7) 3-6ntnf (8) 3-6ntnr (9) 3-6of (10) 3-6or (11) 2020-nwif (12) 2020-ntnf (13) 2020ntn-r. Sequences of these primers are presented in Table 2.1.

Mixed infections and partial genomes

In four samples (**Table 2.3**), analysis of HTS data revealed potential mixed infections or novel recombination events. These samples exhibited two sets of contigs: one set mapping to a full-length genome, and another set mapping to a different strain (**Figure 2.7**). For instance, in sample 2021-5-17, PVY^{O} (100% genome coverage) and $PVY^{N\text{-}Wi}$ (77.8% genome coverage) were identified. Notably, $PVY^{N\text{-}Wi}$ is a recombinant primarily composed of a PVY^{O}-derived genome, except for the region between RJ1 (502-652) and RJ2 (2395-2419) which is derived from the parental strain N (reference) and for here, only a few contigs mapped to this region of the $PVY^{N\text{-}Wi}$ assembly. In sample 2021-5-20, both PVY^{NTNa} (100% genome coverage) and PVY^{O} (98.9% genome coverage) were detected. However, full-length consensus sequences for both strains could not be assembled, with high-scoring segment pairs (HSPs) around 5000 bp (**Table 2.3**). Similarly, in samples2022-3-6 and 2022-20-20 two strains were identified PVY^{O5} (100% genome coverage) and PVY^{NTNa} (100% genome coverage) and $PVY^{N\text{-}Wi}$ (100% genome coverage) and PVY^{NTNa} (72.2% genome coverage), respectively.

RT-PCR and visual mapping confirmed the presence of multiple strains in the four samples. To achieve this, primers were designed to target regions with polymorphisms between contigs or strains (**Figure 2.7**). For instance, if PVY^{O} and $PVY^{N\text{-}Wi}$ were identified as the two strains in a sample, four primers were designed: a forward and reverse primer targeting PVY^{O}, and another set targeting $PVY^{N\text{-}Wi}$. Subsequently, four different RT-PCR reactions were performed, using various combinations of these four primers (**Figure 2.7**, **Table 2.1**). The analysis of PCR results involved assessing band sizes and the specific

combination of primers used. This approach validated that samples 2021-5-20 and 2022-3-6 indeed had mixed infections. However, sample 2022-20-20 remained inconclusive due to the absence of bands, likely caused by incomplete contig sequences used in primer design.

Additionally, seven samples did not yield full-length PVY genomes, with genome coverage ranging from 11.1 - 83.3%. The number of reads mapping to PVY in these samples ranged from 367 to 2405. Despite fewer reads, $PVY^{N\text{-}Wi}$ was identified in four samples, PVY^{NTNa} in two, and the PVY^{O} in one sample. Given that full-length sequences of the strains could not be assembled in these samples, we were unable to ascertain the presence of SNPs in the primer binding regions in these samples.

Phylogenetic Analysis of PVY Based on Complete Genomes

Recombinant and phylogenetic analysis play a crucial role in characterizing virus isolates and strains. Sequence alignment of the 18 samples with complete genomes and 26 reference sequences MUSCLE revealed few nucleotide polymorphisms compared to reference sequences. Notably, significant polymorphisms were observed in the primer binding regions (**Figure 2.9**). Subsequent genome analysis using MEGA X and RDP4.101, focused on phylogenetic relationships and recombination. A maximum likelihood tree demonstrated that 14 out of the 18 samples formed a cluster with the previously characterized $PVY^{N\text{-}Wi}$ strain (KY847996.1), indicating minimal distance between them (**Figure 2.8A**). Three samples clustered with PVY^{NTN}, while one sample clustered with PVY^{O} (AJ585196.1). Further examination revealed that the three samples clustering with the

PVY^{NTN} strains, namely 2021-F26, 2022-15-18 and 2021-4-18, revealed that these samples belong to the PVY^{NTNa} genotype due to the presence of three recombinant junctions (RJs). These RJs distinguish them from the PVY^{NTNb} strain, which has only two (Rodriguez-Rodriguez et al. 2020). UPGMA trees created by RDP4.101 corroborated these findings, with the RJs aligning with known recombination sites described in literature (Green et al. 2017). No novel recombination sites were found in our samples, and all the recombinant sites were within 80 nucleotides from established boundaries. An ANI matrix (**Figure 2.8B**) highlights similarities among the $PVY^{N\text{-}Wi}$ isolates from our study, with nearly 100% nucleotide similarity for all 14 $PVY^{N\text{-}Wi}$ isolates. The sole difference between the 14 samples and the reference $PVY^{N\text{-}Wi}$ strain (KY847996.1) was the presence of the SNP in primer o6400, $T/C_{6356.}$

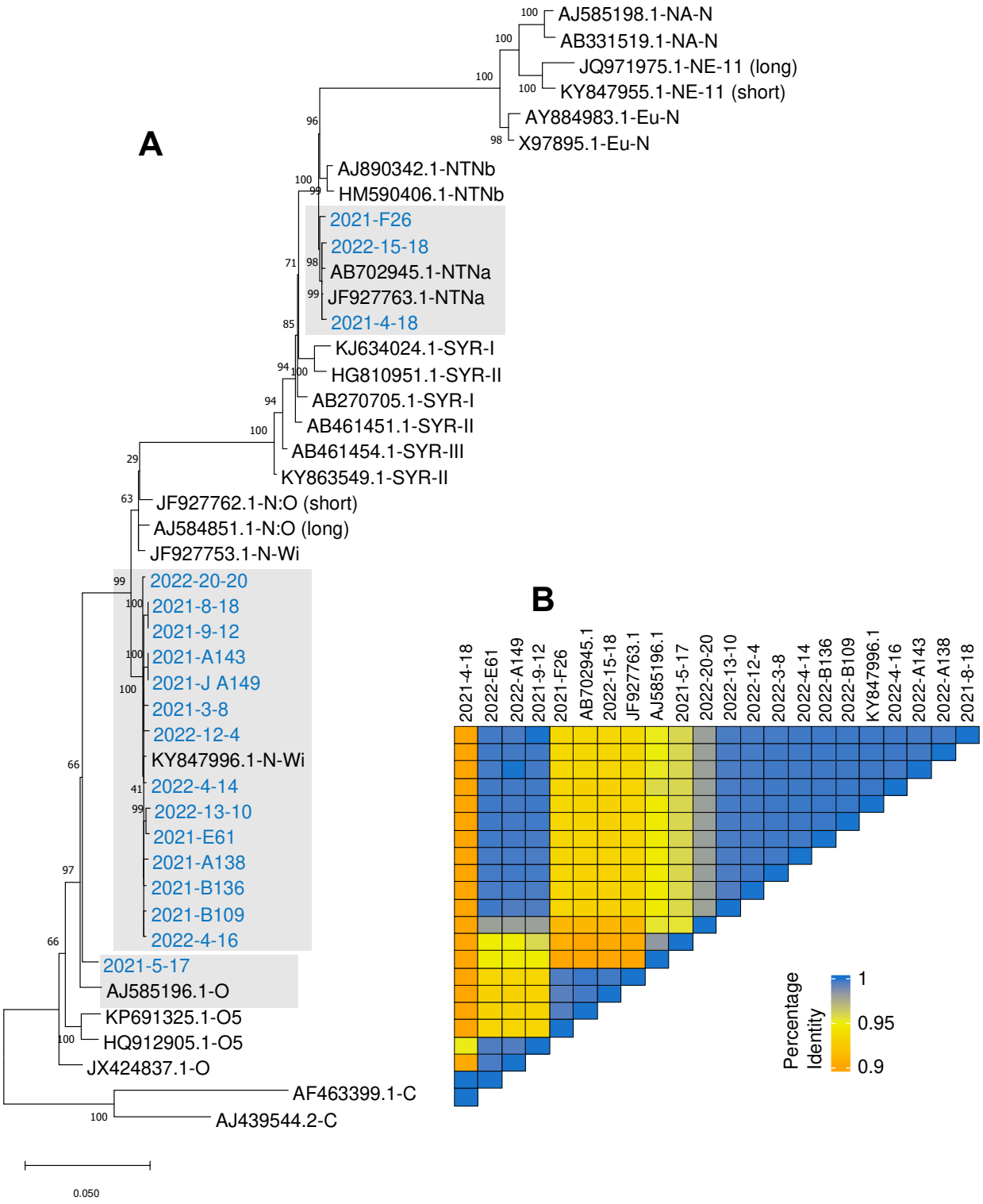

Figure 2.8 Phylogenetic Analysis and Complete Genome Nucleotide Identity of PVY Isolates. (A) Maximum likelihood phylogenetic tree constructed using the PVY isolates from this study and 26 reference PVY genomes retrieved from GenBank. The tree was generated using MEGAX following alignment with MUSCLE. The GTR+G+I substitution

model and a bootstrap value of 1000 were employed. Sequences from this study are written in blue and designated with the year of isolation (2021 or 2022). Branch lengths represent the number of substitutions per site. Reference genome strains are identified by the GenBank accession number. Grey panels indicate samples that are highly similar to each other with minimal distance between them. (B) ANI matrix depicting nucleotide similarities among full-length isolates identified in this study. Genome-scale nucleotide identity was calculated using Pyani v0.2 in Python. The results are visualized in this matrix, with dark red representing 99% or greater identity. Blue shades represent 80% to 90% similarity.

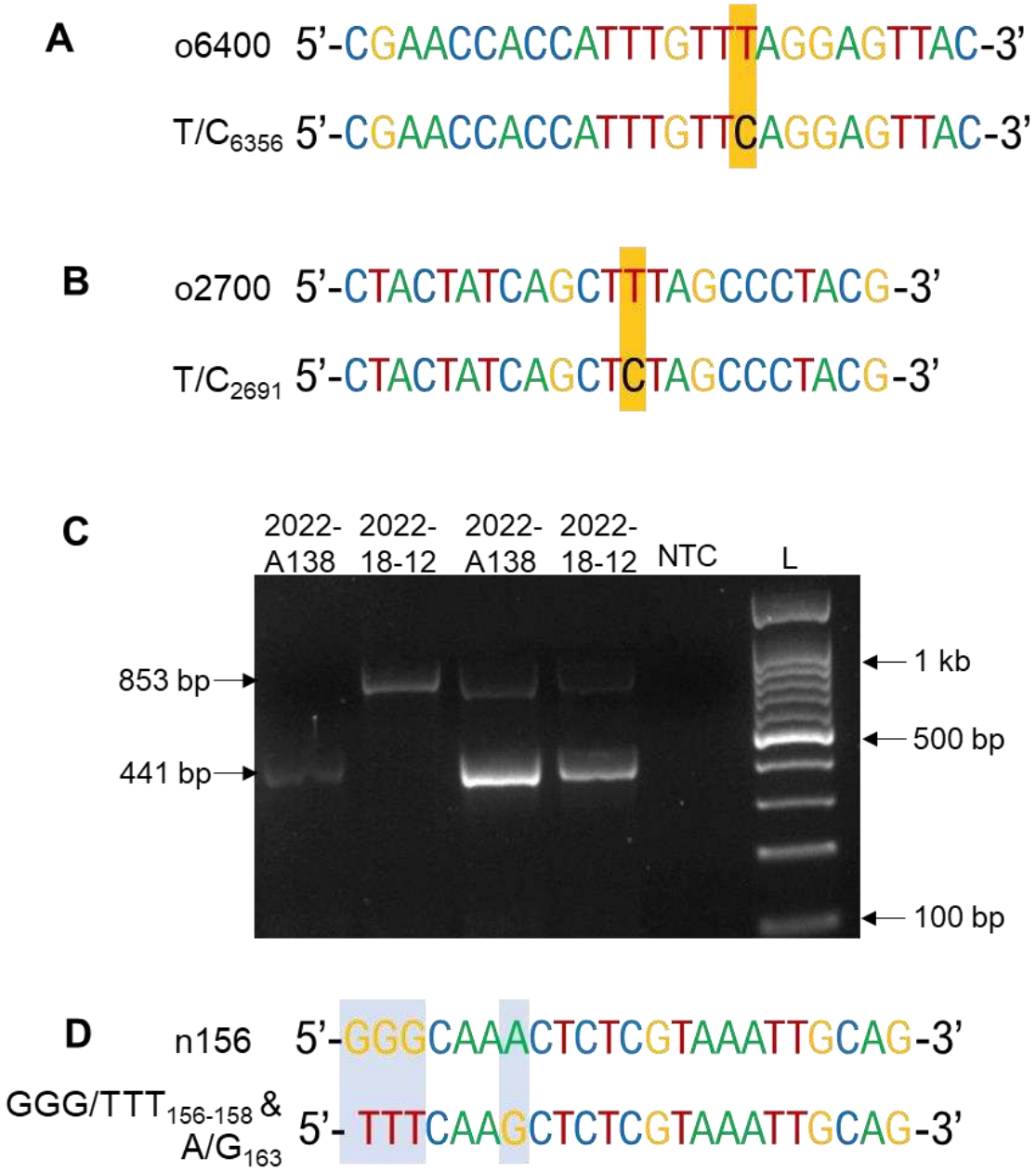

Figure 2.9 Nucleotide polymorphisms identified in three primers used in the Multiplex-RT-PCR. (A) Reverse primer o2700 in the 5'→3' direction produces a band at 441 bp that has an SNP, a T to C, in the middle of the primer that prevents no priming. (B) Reverse primer o6400 in the 5'→3' direction is required for the 441 bp amplicon, harbors a T to C SNP within the primer sequence, disrupting primer binding. (C) Agarose gel image showing the results of a PCR using the original multiplex primers in samples A138 and 18-12 in the first two wells on the left. Sample A138 carries a SNP in the o6400 primer binding region and only produces a band at 441 bps. Conversely, sample 18-12 possesses an SNP in the o2700 primer, resulting in the solo amplification of the 853 bp band. The final two wells demonstrate the effectiveness of modified primers with corrected T to C substitutions. Both bands are successfully amplified in both samples, indicating the restoration of primer

functionality and accurate strain detection. (D) Forward primer n156 (5’→ 3’) direction, which produces a band at 633 bp, has a MNP (GGG/TTT) and a SNP (T/C) that results in no banding in the multiplex.

Potato Virome Analysis

The potato virome in SLV Colorado was assessed using HTS data from field-collected samples. First, we determined the proportion of identified viruses across all sample libraries. Virome analysis revealed two plant viruses and eight viruses of non-plant hosts. Apart from PVY, the other major plant virus detected was PVS. Notably, PVS was the major virus in both years, with 37.5 million reads in 2021 and 11.6 million reads in 2022 (**Table 2.5)**. PVS was present in 12 out of the 25 samples, showing a higher prevalence in 2021. Three samples (2021-4-18, 2021-5-17, 2022-12-4) exhibited > 79% of reads mapping to the PVS genome, indicating a high level of infection, while the remaining nine samples displayed lower infection rates, with reads mapping to 0.001 – 3% of the PVS genome. The genome coverage of PVS in the 12 samples ranged from 64.7% to 100%.

Overall, viral diversity was higher in 2022 compared to 2021, with a higher prevalence of PVS and PVY. Additionally, a diverse range of viruses of non-plant hosts were identified, spanning various families, including *Fusarivirdae*, *Hypoviridae*, some that were unclassified, and a few associated with insects (**Figure 2.10, Table 2.5**). While genome coverage for these viruses was lower than that of PVS and PVY, ranging from 34.4% to 98.7%, all viruses exhibited sequencing depths of > 70% and percent nucleotide identities above 93.3%. The reads mapping to the respective reference genomes was much lower (**Table 2.5**). Among these, three were mycoviruses belonging to the family *Fusariviridae* (Gao et al. 2021): Botrytis cinerea fusarivirus 5 (BcFV5), Botrytis cinerea hypovirus 1 (BcHV1), and Plasmopara viticola lesion-associated fusarivirus 3 (PvlaFV3). BcFV5 and BcHV1 are associated with *Botrytis cinerea*, the causal agent of Botrytis blight (grey mold

and tan spot) on potatoes. PvlaFV3, associated with the oomycete *Plasmopara viticola*, the causal agent of grape downy mildew, was found in three samples. Furthermore, one hypovirus, Leptosphaeria biglobosa hypovirus 1 (LbHV1) (Shah et al. 2018), typically associated with *Leptosphaeria biglobosa*, the causal agent of phoma stem canker (blackleg) in oilseed rape, was also detected. Additionally, two mycoviruses associated with *Erysiphe necator*, the causal agent of powdery mildew on various plants, including potatoes, Erysiphe necator-associated virus 5 (EnaV5) and Erysiphe necator-associated flexivirus 1 (EnaFV1), were identified. Both viruses are not yet classified. Four other virus sequences typically associated with insects/arthropods were also detected in the samples. Dali Fusar tick virus 1 (DFtV1), an unclassified fusarivirus, was detected in five samples across both years as a complete sequence or in various forms. Two unclassified adenoviruses, Acyrthosiphon pisum adenovirus, and Frankliniella occidentalis adenovirus, associated with the pea aphid and western flower thrips, were detected in two and one samples, respectively. Finally, a virus associated with the green peach aphid, Myzus persicae influenza virus, was detected in one sample. This study represents the first comprehensive assessment of the potato virome in the San Luis Valley, providing valuable insights into the prevalence and diversity of other viruses affecting PVY positive potato crops in this region. Virus identity was verified using BLAST searches against the NCBI nr and nt databases. Cutoffs for virus identification were >70% sequencing depth and percent nucleotide identity over 93%. Different colors indicate different viruses and the bar height represent the number of reads mapped to each specific virus. PVY - Potato virus Y, PVS - Potato virus S, BcFV5 - Botrytis cinerea fusarivirus 5, BcHV1 - Botrytis cinerea hypovirus 1,

LbHV1 - Leptosphaeria biglobosa hypovirus 1 (LbHV1), EnaFV1 - Erysiphe necator associated flexivirus 1, EnaV5 – Erysiphe necator associated virus 1, PvlaFV3 – Plasmopara viticola lesion associated flexivirus 3. The other viruses have not been abbreviated as they have not been characterized or classified into families.

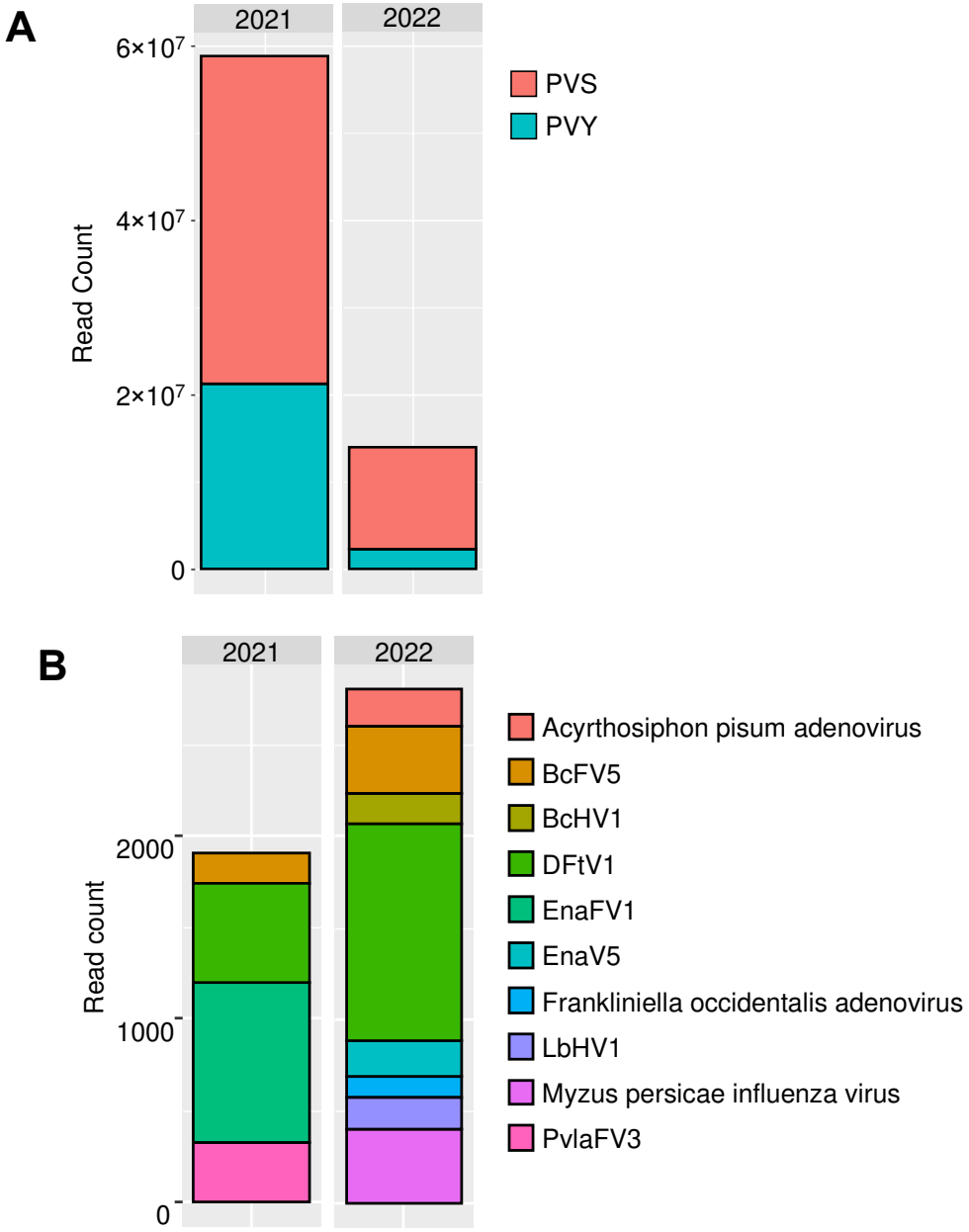

Figure 2.10 Viruses and their read counts in 2021 and 2022. (A) Stacked bar graph showing the relative abundance of PVY and PVS reads in field-collected samples from SLV. (B) Comparison of viral diversity and read counts of secondary viruses identified in the field-collected samples.

Table 2.5 Summary of raw sequence data for RNA-seq (bioproject PRJNA1058533)

	Sample ID	Total Reads	# reads mapped to Potato Genome	# unmapped reads	SRA Accession number
1	2021-4-18	42,018,908	30,081,441	11,937,467	SRR27373714
2	2021-5-17	43,193,070	12,773,648	30,419,422	SRR27373713
3	2021-8-18	43,216,898	34,217,440	8,999,458	SRR27373719
4	2021-9-12	40,627,618	33,073,635	7,553,983	SRR27373718
5	2021-3-8	46,754,950	39,268,421	7,486,529	SRR27373717
6	2021-A138	43,254,262	22,602,392	20,651,870	SRR27373716
7	2021-A143	40,286,090	27,228,828	13,057,262	SRR27373715
8	2021-A149	43,929,068	40,024,195	3,904,873	SRR27373712
9	2021-B109	47,815,366	28,371,911	19,443,455	SRR27373711
10	2021-B136	42,101,636	2,199,755	39,901,881	SRR27373742
11	2021-E61	43,034,204	34,622,235	8,411,969	SRR27373741
12	2021-F26	43,458,900	30,341,588	13,117,312	SRR27373740
13	2022-4-14	39,856,704	38,489,966	1,366,738	SRR27373737
14	2022-4-16	48,856,306	47,663,748	1,192,558	SRR27373735
15	2022-12-4	56,598,508	43,963,371	12,635,137	SRR27373731
16	2022-13-10	42,017,032	40,118,089	1,898,943	SRR27373730
17	2022-15-18	40,811,560	37,476,735	2,629,237	SRR27373728
18	2022-20-20	46,485,752	38,777,665	2,033,895	SRR27373721
19	2021-5-18	54,078,424	52,347,222	1,731,202	SRR27373734
20	2021-5-20	43,351,632	32,852,842	10,498,790	SRR27373723
21	2021-6-10	44,472,106	42,891,545	1,580,561	SRR27373720
22	2022-3-6	47,307,654	45,074,749	2,232,905	SRR27373739
23	2022-15-17	40,105,972	37,476,735	2,629,237	SRR27373729
24	2022-16-3	40,772,258	39,770,206	1,002,052	SRR27373727
25	2022-17-3	39,614,130	37,326,757	2,287,373	SRR27373726
26	2022-18-12	43,450,756	42,084,547	1,366,209	SRR27373725
27	2022-19-2	48,311,738	46,941,208	1,370,530	SRR27373722
28	2022-4-6	41,204,200	40,353,309	850,891	SRR27373738
29	2022-4-15	42,606,082	38,489,966	1,366,738	SRR27373736
30	2022-9-7	44,462,472	43,535,019	927,453	SRR27373733
31	2022-10-2	46,796,606	45,820,177	976,429	SRR27373732
32	2022-18-18	41,352,654	40,333,513	1,019,141	SRR27373724

Table 2.6 Summary of potato viromes from San Luis Valley, Colorado

Sample	Virus	# reads mapping to Virus	% reads mapping to Virus	# contigs	Mean contig length	Mean Sequence Depth (coverage)	Percent Genome coverage	Consensus Length or HSP*	Closest BLASTRelative	Closest accession number
2021-4-18	Potato virus S	9,390,557	78.7	1	8472	99.9	100	8472	99.7	FJ813513
2021-5-17	Potato virus S	28,101,266	92.4	15	739	99.9	72	5981	99.5	FJ813512
2021-5-18	Potato virus S	965	0.06	5	662	98.3	64.7	1201	99.2	FJ813512
2021-5-20	Potato virus S	952	0.01	3	3224	99.7	100	8375	99.5	FJ813513
2021-6-10	Dali Fusar tick virus 1	875	0.01	2	491	99	46.5	491	100	ON746394
	Erysiphe necator associated flexivirus 1	536	0.01	2	434	99.5	86.7	434	99.1	MN627466
2021-9-12	Potato virus S	345	0.00	3	635	100	22.8	1825	99.3	FJ813513
	Botrytis cinerea fusarivirus 5	105	0.00	2	384	78.4	48.2	602	93.3	MN617764
2021-8-18	Botrytis cinerea fusarivirus 5	64	0.00	1	416	93.3	32.6	416	93.7	MN617764
2022-4-6	Botrytis cinerea fusarivirus 5	55	0.01	2	448	70.6	34.88	546	96.4	MN617764
	Potato virus S	770	0.09	3	1696	99.9	92.3	3848	100	FJ813513
2022-3-6	Botrytis cinerea fusarivirus 5	233	0.01	7	897	96.6	98.2	1942	97.4	MN617764
2022-4-15	Erysiphe necator associated virus 5	60	0.0	1	1,021	96.3	94.9	1,021	100	MN611685

	Dali Fusar tick virus 1	97	0.0	2	675	97	69.2	675	100	ON746394
2022-12-4	Potato virus S	11,546,480	91.4	6	1724	100	100	8462	100	MT107560
2022-4-16	Frankliniella occidentalis adenovirus	79	0.0	2	404	93.8	43.1	555	100	XM_026428773
2022-13-10	Botrytis cinerea hypovirus 1	168	0.01	2	1092	100	47.4	1255	96.8	NC_037659
	Leptosphaeria biglobosa hypovirus 1	169	0.01	3	799	77.9	91.9	930	100	OP441692
2022-16-3	Dali Fusar tick virus 1	32	0.0	1	605	82.4	53.5	605	100	ON746394
2022-15-18	Leptosphaeria biglobosa hypovirus 1	22	0.0	3	434	100	70.6	651	100	OP441692
	Plasmopara viticola lesion associated fusarivirus 3	64	0.0	2	472	94.1	53.9	577	100	MN551107
2022-20-20	Potato virus S	40,243	3.1	1	8453	97.7	100	8453	99.7	FJ813513
	Botrytis cinerea fusarivirus 5	78	0.01	4	922	100	66.5	922	96.6	MN617764
2022-19-2	Plasmopara viticola lesion associated fusarivirus 3	60	0.0	2	820	96.5	34.3	1019	96.6	MN551107
2022-15-17	Plasmopara viticola lesion associated fusarivirus 3	234	0.01	1	6605	96.8	98.7	6605	6340	MN551107
	Myzus persicae influenza virus	403	0.02	5	2921	100	100	2903	99.8	XM_022320503

	Acyrthosiphon pisum adenovirus	199	0.01	1	1869	98.8	87	1869	100	XM_001951458
	Erysiphe necator associated virus 5	37	0.0	1	677	100	94.2	677	100	MN611685
2022-17-3	Erysiphe necator associated virus 5	93	0.0	3	598	93.8	73.4	598	96.7	MN611685
	Frankliniella occidentalis adenovirus	42	0.0	1	814	96.4	43.5	814	100	XM_026428773
	Dali Fusar tick virus 1	542	0.01	1	4511	100	88.7	3970	93.8	ON746394

DISCUSSION

Our study comprehensively assesses PVY strains in the SLV. Initial data from grow-out testing, beginning in 2019, revealed a dynamic strain landscape characterized by mosaic symptoms and varying strain compositions varying across the years. PVY^{N-Wi} is the predominant strain in SLV potatoes, followed by PVY^{O}. Significantly, mixed infection rates and intra-plant strain diversity increased throughout the study period. HTS of samples that showed inconclusive results in the multiplex RT-PCRs revealed the presence of mixed infections and the presence of SNPs in primers used to detect common strains PVY^{N-Wi}, PVY^{O}, and PVY^{NTN}. The observed changes in strain composition and increased mixed infection rates underscore the necessity for continuous monitoring of PVY strains in the SLV. Although, not detected in this study, such shifts could lead to the emergence of new recombinants, posing a greater threat to potato production in the region. Therefore, ongoing surveillance and using robust strain identification methods are crucial for effective PVY management. Additionally, our characterization of the potato virome in the SLV revealed PVS and PVY as the most prevalent viruses within PVY positive plants, alongside a diverse range of non-plant host and insect-associated viruses.

Understanding the prevalent PVY strains within a potato production area is crucial, as their relative diversity and abundance can impact various factors, including diagnostic tool selection by certification agencies, the choice of potato cultivar to plant by growers, and the direction of breeding efforts for virus-resistant varieties (Karasev and Gray 2013). Here, we monitored PVY strain diversity over five years (2017-2022) in post-harvest tests and two growing seasons (2021-2022). The persistent presence of PVY-associated mosaic

symptoms in post-harvest tests underscores the continuous threat posed by the virus to potato production in SLV. While PVY^{N-Wi} is the dominant strain, PVY^{O} and PVY^{NTN} are present in moderate proportions, with $PVY^{N:O}$ and PVY-NE-11 being rare. A similar trend was found in the field samples tested from the two growing seasons. Notably, the population structure of PVY in SLV differs from other regions. For instance, in Canada, the PVY^{O} strain has been largely replaced by PVY^{NTN} and $PVY^{N:O}$ (MacKenzie et al. 2018), while the Pacific Northwest experienced a significant decline in PVY^{O}, replaced by PVY^{N-Wi} and PVY^{NTNa} (Funke et al. 2017; Tran et al. 2022). In contrast, the SLV region experienced no such decline, with PVY^{O} remaining the second most prevalent strain. This stability can be partly attributed to the extensive planting of Russet Norkotah, which lacks resistance to the PVY^{O} strain, in SLV. A recent study using tobacco bait plants also identified PVY^{N-Wi}, PVY^{O}, and PVY^{NTN} as the major strains in SLV but reported a higher PVY^{O} incidence (Daniel and Chikh-Ali 2024) than our findings. This discrepancy can be explained in part by the higher transmission efficiency of PVY^{O} by aphid vectors when compared to PVY^{N-Wi} (Mondal et al. 2016). Despite these variations, both studies underscore the complexity of PVY spread and incidence within agroecosystems.

The occurrence of dual and triple infections highlight the dynamic nature of the PVY population structure in SLV. While mixed infections by different PVY strains are commonly reported (Della Bartola et al. 2020; Glasa et al. 2021; Tran et al. 2022), their increase in presence during the study period raises concerns. Interestingly, the SLV appears to exhibit higher incidences of mixed infections compared to previous studies. As a potyvirus PVY, like other members of this family has a propensity for recombination (Gibbs et al. 2017).

This process along with mutations, increases genetic variability, potentially enhancing the fitness or survival within the viral population and facilitating the emergence of new resistance-breaking strains (Gibbs et al. 2020; Gibbs et al. 2017; Nigam et al. 2019). Moreover, mixed infections can result in synergistic or antagonistic interactions between viral strains, with diverse biological, epidemiological, and economic implications (Syller 2012). Specifically, evidence suggests an antagonistic interaction between PVY strains within plants. For instance, PVY^{NTN} accumulates to higher titers than PVY^{O} or $PVY^{N\text{-}Wi}$ as infection progresses (Mondal et al. 2023; Syller and Grupa 2014). Additionally, the presence of multiple strains within a plant can influence their transmission efficiency by aphids (Syller and Grupa 2014). Another layer of complexity includes the presence of potato cultivars resistant to specific strains. Symptoms can present themselves in the plant differently depending on cultivar and the strain infecting the host.

Mixed PVY infections add complexity to strain characterization. Our study revealed a concerning increase in mixed infections since 2017 in the post-harvest testing and during in-season testing, posing challenges for accurate PVY strain identification. Illumina short-read sequencing proved insufficient for *de novo* assembly of full length contigs when dealing with multiple strains within a sample. The high similarity of PVY genomes, particularly at recombinant junctions, hampered accurate alignment. Consequently, traditional methods like RT-PCR and manual assembly were necessary to confirm mixed infections (Glasa et al. 2021), albeit with increased time and labor. Combining short-read sequencing with long-read sequencing technologies like the MinION (Oxford Nanopore Technologies) offers a promising solution. Della Bartola et al. (2020) demonstrated near

full-length contig recovery for different strains in a single sample using this approach. Ultimately, mixed infections pose obstacles for multiplex PCRs and Illumina short-read sequencing-based PVY strain characterization. While time-intensive methods offer solutions, long-read sequencing can overcome these challenges and enhance our understanding of PVY strain diversity.

HTS detected mixed infections and identified a possible cause of the inconclusive results obtained with the two multiplex RT-PCR assays. Full-length PVY sequences were recovered from the samples previously classified as inconclusive after the two multiplex RT-PCRs. Full-length PVY sequences were recovered from previously inconclusive samples, revealing nucleotide polymorphisms in primer binding regions. It is plausible that these polymorphisms hindered the amplification of specific bands, resulting in ambiguous results. Polymorphisms in primer binding regions, particularly in the o6400, o2700, and n156 primers may have affected the accurate identification of $PVY^{N\text{-}Wi}$, PVY^{NTN}, and PVY^{O} strains. It is important to note that other factors such as low virus titers and RNA quality can affect PCR efficiency (Pallás et al. 2018), although the samples that underwent HTS required stringent quality checks. Notably, all the $PVY^{N\text{-}Wi}$ identified from the multiplex RT-PCR inconclusive samples in this study had mismatches in the o6400 primer binding regions. However, it will be important to obtain a larger dataset of $PVY^{N\text{-}Wi}$ samples before further conclusions can be drawn. The detection of polymorphisms underscore the critical need for comprehensive identification and sequencing of PVY strains prevalent within a given potato production area. This will ensure that detection tools remain effective and

accurate, enabling timely interventions and mitigating the impact of PVY on potato production.

This study presents the first report of the potato virome in the U.S. from field-collected samples in the SLV. While potatoes can be infected by over 40 viruses (Chikh-Ali and Karasev 2023; Kreuze et al. 2020), our analysis detected only two plant viruses, PVY and PVS, in our samples. Although PVS is not typically considered a major threat to potato production in the U.S. and is not routinely monitored, it can cause yield losses of 10-20% under severe conditions (Chikh-Ali and Karasev 2023). In the 12 samples with PVS, genome coverage was > 96%, with millions of reads mapping back to the genome, indicating a high level of infection. While PVS alone may not cause significant yield losses, co-infections with PVY can exacerbate symptoms (Chikh-Ali and Karasev 2023; Valkonen et al. 2017), potentially complicating PVY detection during field inspections. The higher incidence of both viruses in 2021 compared to 2022 suggests a potential correlation between their incidence and co-infection rates. Our findings contrast with other studies that have identified more viruses in potatoes from different growing regions. For example, Lai et al. (2022) identified three major viruses, PVY, PVS, and potato virus M (PVM), and eight other plant viruses, including common potato viruses like potato virus H (PVH), potato mop-top virus PMTV) and potato leafroll virus (PLRV) in China. Similarly, Gutiérrez et al. (2021) detected PVY, PVS, PMTV, potato virus X (PVX), potato virus V (PVV), and potato yellow vein virus (PYVV) in samples containing a mixture of foliar and sprout tissue from Colombian potatoes. Finally, Elwan et al. (2023) identified three potato viruses, alfalfa mosaic virus (AMV), PLRV, and PVY, in addition to three other viruses, PVV, Andean potato latent virus

(APLV), and tomato chlorosis virus (ToCV) in Egyptian potatoes. A possible reason for the low virus diversity observed in our study could be that we included only PVY-positive samples in our virome analysis. It is plausible that this strategy likely resulted in the exclusion of other economically important viruses, such as PMTV, that are known to be found in the San Luis Valley (Mallik and Gudmestad, 2014). The limited diversity of potato-infecting viruses detected in our study may also be attributed to the number of samples used and the tissue sampled. For instance, PMTV and tobacco rattle virus (TRV) were not detected, but since they are largely confined to the tubers, their detection in foliar tissue may require sampling of tuber tissues or additional sequencing depth.

The detection of viruses typically infecting fungal plant pathogens and insect pests in potato plant virome is notable. Elwan et al. (2023) and Rivarez et al. (2023) also documented mycoviruses and insect-associated viruses in crop and weed viromes, highlighting the understudied diversity of plant viromes in agroecosystems. Mycoviruses, capable of infecting and replicating within their fungal and oomycete hosts, hold promise as biocontrol agents by attenuating fungal virulence (Sutela et al. 2019). For instance, BcHV1 infection has been shown to diminish the virulence of *B. cinerea* (Hao et al. 2018). Further research is needed to ascertain the impact of other identified mycoviruses on their respective pathogen hosts. The identification of insect-associated viruses raises questions about their source. Alternatively, these viruses might reside within the plants themselves. Recent discoveries are rapidly expanding our understanding of the "virosphere," revealing a previously unseen level of viral diversity that extends far beyond the traditional focus on just pathogenic and plant-infecting viruses (Maclot et al. 2020; Shi et al. 2016). As more

HTS datasets from field-collected plant virome studies emerge, we are encountering viruses in plant tissues that were not previously considered present in such locations.

We propose several important avenues for effectively managing PVY and mitigating its impact on potato production. First, we recommend prioritizing planting potato cultivars that carry any of the *Ry* resistance genes whenever feasible. One key area is continued monitoring of PVY strain diversity and abundance over time. Utilizing HTS will provide crucial insights into the emergence of new strains. This knowledge will facilitate the development of more robust and accurate PVY strain identification methods. Additionally, unraveling the intricate interactions between multiple PVY strains and co-infecting viruses within the potato virome is vital for understanding their synergistic and antagonistic effects on disease severity and potato yield. Addressing these research priorities will assist breeding programs in developing potato varieties with appropriate host plant resistance.

REFERENCES

Chikh-Ali, M., and Karasev, A. V. 2015. Immunocapture-Multiplex RT-PCR for the Simultaneous Detection and Identification of Plant Viruses and Their Strains: Study Case, Potato Virus Y (PVY). Pages 177-186. Springer New York.

Chikh-Ali, M., and Karasev, A. V. 2023. Virus diseases of potato and their control. Pages 199-212 in: Potato Production Worldwide. Elsevier.

Chikh-Ali, M., Gray, S. M., and Karasev, A. V. 2013. An improved multiplex IC-RT-PCR assay distinguishes nine strains of Potato virus Y. Plant Dis. 97:1370-1374.

Chikh-Ali, M., Maoka, T., Natsuaki, T., and Natsuaki, K. 2010. PVY^{NTN-NW}, a novel recombinant strain of Potato virus Y predominating in potato fields in Syria. Plant pathol. 59:31-41.

Chikh-Ali, M., Tran, L. T., Price, W. J., and Karasev, A. V. 2020. Effects of the age-related resistance to Potato virus Y in potato on the systemic spread of the virus, incidence of the potato tuber necrotic ringspot disease, tuber yield, and translocation rates into progeny tubers. Plant Dis. 104:269-275.

Daniel, J., and Chikh-Ali, M. 2024. Dynamics of Potato Cirus Y Infection Pressure and Strain Composition in The San Luis Valley, Colorado. Plant Dis. TBD.

Davidson, R., Houser, A., Sather, K., and Haslar, R. 2013. Controlling PVY in seed: what works and what does not. Am. J. Potato Res. 90:28-32.

Della Bartola, M., Byrne, S., and Mullins, E. 2020. Characterization of Potato Virus Y Isolates and Assessment of Nanopore Sequencing to Detect and Genotype Potato Viruses. Viruses 12:478.

Devaux, A., Goffart, J.-P., Petsakos, A., Kromann, P., Gatto, M., Okello, J., Suarez, V., and Hareau, G. 2020. Global food security, contributions from sustainable potato agri-food systems. Pages 3-35 in: The potato crop. Springer, Cham.

Ellis, P., Stace-Smith, R., and De Villiers, G. 1997. Identification and geographic distribution of serotypes of potato virus Y. Plant Dis. 81:481-484.

Elwan, E. A., Rabie, M., Aleem, E. E. A., Fattouh, F. A., Kagda, M. S., and Zaghloul, H. A. 2023. Exploring virus presence in field-collected potato leaf samples using RNA sequencing. Journal of Genetic Engineering and Biotechnology 21:106.

Funke, C. N., Nikolaeva, O. V., Green, K. J., Tran, L. T., Chikh-Ali, M., Quintero-Ferrer, A., Cating, R. A., Frost, K. E., Hamm, P. B., and Olsen, N. 2017. Strain-specific resistance to

Potato virus Y (PVY) in potato and its effect on the relative abundance of PVY strains in commercial potato fields. Plant Dis. 101:20-28.

Gallo, Y., Marín, M., and Gutiérrez, P. 2021. Detection of RNA viruses in *Solanum quitoense* by high-throughput sequencing (HTS) using total and double stranded RNA inputs. Physiol. Mol. Plant Pathol. 113:101570.

Gao, Z., Cai, L., Liu, M., Wang, X., Yang, J., An, H., Deng, Q., Zhang, S., and Fang, S. 2021. A novel previously undescribed fusarivirus from the phytopathogenic fungus *Setosphaeria turcica*. Arch. Virol. 166:665-669.

Gibbs, A., and Ohshima, K. 2010. Potyviruses and the digital revolution. Annu. Rev. Phytopathol. 48:205-223.

Gibbs, A. J., Hajizadeh, M., Ohshima, K., and Jones, R. A. 2020. The potyviruses: an evolutionary synthesis is emerging. Viruses 12:132.

Gibbs, A. J., Ohshima, K., Yasaka, R., Mohammadi, M., Gibbs, M. J., and Jones, R. A. 2017. The phylogenetics of the global population of potato virus Y and its necrogenic recombinants. Virus Evolution 3:vex002.

Gibbs, M. J., Armstrong, J. S., and Gibbs, A. J. 2000. Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics 16:573-582.

Glais, L., Tribodet, M., and Kerlan, C. 2002. Genomic variability in Potato potyvirus Y (PVY): evidence that $PVY^{N}W$ and PVY^{NTN} variants are single to multiple recombinants between PVY^{O} and PVY^{N} isolates. Arch. Virol. 147:363-378.

Glais, L., Bellstedt, D. U., and Lacomme, C. 2017. Diversity, Characterisation and Classification of PVY. Pages 43-76. Springer International Publishing.

Glasa, M., Hančinský, R., Šoltys, K., Predajňa, L., Tomašechová, J., Hauptvogel, P., Mrkvová, M., Mihálik, D., and Candresse, T. 2021. Molecular characterization of Potato Virus Y (PVY) using high-throughput sequencing: Constraints on full genome reconstructions imposed by mixed infection involving recombinant PVY strains. Plants 10:753.

Gray, S., De Boer, S., Lorenzen, J., Karasev, A., Whitworth, J., Nolte, P., Singh, R., Boucher, A., and Xu, H. 2010. Potato virus Y: an evolving concern for potato crops in the United States and Canada. Plant Dis. 94:1384-1397.

Gray, S. M., and Power, A. G. 2018. Anthropogenic influences on emergence of vector-borne plant viruses: The persistent problem of Potato virus Y. Current opinion in virology 33:177-183.

Green, K. J., Brown, C. J., and Karasev, A. V. 2018. Genetic diversity of potato virus Y (PVY): sequence analyses reveal ten novel PVY recombinant structures. Arch. Virol. 163:23-32.

Green, K. J., Brown, C. J., Gray, S. M., and Karasev, A. V. 2017. Phylogenetic study of recombinant strains of Potato virus Y. Virology 507:40-52.

Gutiérrez, P., Rivillas, A., Tejada, D., Giraldo, S., Restrepo, A., Ospina, M., Cadavid, S., Gallo, Y., and Marín, M. 2021. PVDP: A portable open source pipeline for detection of plant viruses in RNAseq data. A case study on potato viruses in Antioquia (Colombia). Physiol. Mol. Plant Pathol. 113:101604.

Hao, F., Ding, T., Wu, M., Zhang, J., Yang, L., Chen, W., and Li, G. 2018. Two novel hypovirulence-associated mycoviruses in the phytopathogenic fungus *Botrytis cinerea*: Molecular characterization and suppression of infection cushion formation. Viruses 10:254.

Jo, Y., Kim, S.-M., Choi, H., Yang, J. W., Lee, B. C., and Cho, W. K. 2020. Sweet potato viromes in eight different geographical regions in Korea and two different cultivars. Scientific Reports 10:2588.

Karasev, A. V., and Gray, S. M. 2013. Genetic diversity of Potato virus Y complex. Am. J. Potato Res. 90:7-13.

Karasev, A. V., Nikolaeva, O. V., Hu, X., Sielaff, Z., Whitworth, J., Lorenzen, J. H., and Gray, S. M. 2010. Serological properties of ordinary and necrotic isolates of Potato virus Y: a case study of PVY N misidentification. Am. J. Potato Res. 87:1-9.

Kerlan, C., and Moury, B. 2008. Potato Virus Y. In 'Encyclopedia of Virology'.(Eds, BWJ Mahy, MHV van Regenmortel) pp. 287–296. Academic Press: San Diego.

Kreuze, J. F., Souza-Dias, J., Jeevalatha, A., Figueira, A., Valkonen, J., and Jones, R. 2020. Viral diseases in potato. The potato crop: its agricultural, nutritional and social contribution to humankind:389-430.

Lacomme, C., Glais, L., Bellstedt, D. U., Dupuis, B., Karasev, A. V., and Jacquot, E. 2017. Potato virus Y: biodiversity, pathogenicity, epidemiology and management. Springer.

Lai, X., Wang, H., Wu, C., Zheng, W., Leng, J., Zhang, Y., and Yan, L. 2022. Comparison of potato viromes between introduced and indigenous varieties. Frontiers in microbiology 13:809780.

Lorenzen, J. H., Piche, L. M., Gudmestad, N. C., Meacham, T., and Shiel, P. 2006. A multiplex PCR assay to characterize Potato virus Y isolates and identify strain mixtures. Plant Dis. 90:935-940.

MacKenzie, T. D., Lavoie, J., Nie, X., and Singh, M. 2018. Differential spread of Potato virus Y (PVY) strains O, N: O and NTN in the field: Implications for the rise of recombinant PVY strains in New Brunswick, Canada. Am. J. Potato Res. 95:301-310.

Maclot, F., Candresse, T., Filloux, D., Malmstrom, C. M., Roumagnac, P., Van der Vlugt, R., and Massart, S. 2020. Illuminating an ecological blackbox: using high throughput sequencing to characterize the plant virome across scales. Frontiers in Microbiology 11:578064.

Mallik, I., Anderson, N. R., and Gudmestad, N. C. 2012. Detection and differentiation of Potato virus Y strains from potato using immunocapture multiplex RT-PCR. Am. J. Potato Res. 89:184-191.

Mallik, I., and Gudmestad, N. 2014. First report of Potato mop top virus causing potato tuber necrosis in Colorado and New Mexico. Plant Dis. 99:164-164.

Martin, D., and Rybicki, E. 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics 16:562-563.

Mondal, S., Wintermantel, W. M., and Gray, S. M. 2023. Infection dynamics of potato virus Y isolate combinations in three potato cultivars. Plant Dis. 107:157-166.

Mondal, S., Wenninger, E. J., Hutchinson, P. J., Whitworth, J. L., Shrestha, D., Eigenbrode, S. D., and Bosque-Pérez, N. A. 2016. Comparison of transmission efficiency of various isolates of Potato virus Y among three aphid vectors. Entomol. Exp. Appl. 158:258-268.

Nigam, D., LaTourrette, K., Souza, P. F., and Garcia-Ruiz, H. 2019. Genome-wide variation in potyviruses. Frontiers in Plant Science 10:1439.

Nolte, P., Whitworth, J. L., Thornton, M. K., and McIntosh, C. S. 2004. Effect of seedborne Potato virus Y on performance of Russet Burbank, Russet Norkotah, and Shepody potato. Plant Disease 88:248-252.

Padidam, M., Sawyer, S., and Fauquet, C. M. 1999. Possible emergence of new geminiviruses by frequent recombination. Virology 265:218-225.

Pallás, V., Sánchez-Navarro, J. A., and James, D. 2018. Recent advances on the multiplex molecular detection of plant viruses and viroids. Frontiers in microbiology 9:2087.

Pelletier, Y., Nie, X., Giguère, M. A., Nanayakkara, U., Maw, E., and Foottit, R. 2012. A New Approach for the Identification of Aphid Vectors (Hemiptera: Aphididae) of Potato Virus Y. J. Econ. Entomol. 105:1909-1914.

Pitt, W. J., Kairy, L., Villa, E., Nalam, V. J., and Nachappa, P. 2022. Virus Infection and Host Plant Suitability Affect Feeding Behaviors of Cannabis Aphid (Hemiptera: Aphididae), a Newly Described Vector of Potato Virus Y. Environ. Entomol. 51:322-331.

Posada, D., and Crandall, K. A. 2001. Selecting the best-fit model of nucleotide substitution. Syst. Biol. 50:580-601.

Quenouille, J., Vassilakos, N., and Moury, B. 2013. Potato virus Y: a major crop pathogen that has provided major insights into the evolution of viral pathogenicity. Mol. Plant Pathol. 14:439-452.

Rivarez, M. P. S., Pecman, A., Bačnik, K., Maksimović, O., Vučurović, A., Seljak, G., Mehle, N., Gutiérrez-Aguirre, I., Ravnikar, M., and Kutnjak, D. 2023. In-depth study of tomato and weed viromes reveals undiscovered plant virus diversity in an agroecosystem. Microbiome 11:1-24.

Rodriguez-Rodriguez, M., Chikh-Ali, M., Johnson, S. B., Gray, S. M., Malseed, N., Crump, N., and Karasev, A. V. 2020. The recombinant potato virus Y (PVY) strain, PVY^{NTN}, identified in potato fields in Victoria, southeastern Australia. Plant Dis. 104:3110-3114.

Romancer, M. L., Kerlan, C., and Nedellec, M. 1994. Biological characterisation of various geographical isolates of potato virus Y inducing superficial necrosis on potato tubers. Plant pathol. 43:138-144.

Rosenman, J., McIntosh, C. S., Aryal, G. R., and Nolte, P. 2019. Planting a problem: Examining the spread of seed-borne Potato virus Y. Plant Dis. 103:2179-2183.

Salminen, M. O., Carr, J. K., Burke, D. S., and McCutchan, F. E. 1995. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. AIDS Res. Hum. Retrovir. 11:1423-1425.

Shah, U. A., Kotta-Loizou, I., Fitt, B. D., and Coutts, R. H. 2018. Identification, molecular characterization, and biology of a novel quadrivirus infecting the phytopathogenic fungus Leptosphaeria biglobosa. Viruses 11:9.

Shi, M., Lin, X.-D., Tian, J.-H., Chen, L.-J., Chen, X., Li, C.-X., Qin, X.-C., Li, J., Cao, J.-P., and Eden, J.-S. 2016. Redefining the invertebrate RNA virosphere. Nature 540:539-543.

Sigvald, R. 1987. Aphid migration and the importance of some aphid species as vectors of potato virus Y o (PVY o) in Sweden. Potato Research 30:267-283.

Sigvald, R. 1990. Aphids on potato foliage in Sweden and their importance as vectors of Potato virus Y0. Acta Agriculturae Scandinavica 40:53-58.

Smith, J. M. 1992. Analyzing the mosaic structure of genes. J. Mol. Evol. 34:126-129.

Smith, K. M. 1931. Composite nature of certain potato viruses of the mosaic group. Nature 127:702-702.

Starchevskaya, M., Kamanova, E., Vyatkin, Y., Tregubchak, T., Bauer, T., Bodnev, S., Rotskaya, U., Polenogova, O., Kryukov, V., and Antonets, D. 2023. The Metagenomic Analysis of Viral Diversity in Colorado Potato Beetle Public NGS Data. Viruses 15:395.

Sutela, S., Poimala, A., and Vainio, E. J. 2019. Viruses of fungi and oomycetes in the soil environment. FEMS Microbiol. Ecol. 95:fiz119.

Syller, J. 2012. Facilitative and antagonistic interactions between plant viruses in mixed infections. Mol. Plant Pathol. 13:204-216.

Syller, J., and Grupa, A. 2014. The effects of co-infection by different P otato virus Y (PVY) isolates on virus concentration in solanaceous hosts and efficiency of transmission. Plant pathol. 63:466-475.

Tollenaere, C., Susi, H., and Laine, A.-L. 2016. Evolutionary and epidemiological implications of multiple infection in plants. Trends Plant Sci. 21:80-90.

Tran, L. T., Green, K. J., Rodriguez-Rodriguez, M., Orellana, G. E., Funke, C. N., Nikolaeva, O. V., Quintero-Ferrer, A., Chikh-Ali, M., Woodell, L., Olsen, N., and Karasev, A. V. 2022. Prevalence of Recombinant Strains of Potato Virus Y in Seed Potato Planted in Idaho and Washington States Between 2011 and 2021. Plant Disease 106:810-817.

Valkonen, J. P. T., Gebhardt, C., Zimnoch-Guzowska, E., and Watanabe, K. N. 2017. Resistance to Potato virus Y in Potato. Pages 207-241 in: Potato virus Y: biodiversity, pathogenicity, epidemiology and management. C. Lacomme, L. Glais, D. U. Bellstedt, B. Dupuis, A. V. Karasev and E. Jacquot, eds. Springer International Publishing, Cham.

Whitworth, J. L., Hamm, P. B., and Nolte, P. 2012. Distribution of Potato virus Y strains in tubers during the post-harvest period. Am. J. Potato Res. 89:136-141.

CHAPTER 3 – POTATO MOP-TOP VIRUS GENOME DIVERSITY IN THE UNITED STATES

INTRODUCTION

Plasmodiophorids, soil-borne protists are known to vector four economically important viruses (Sarwar et al., 2020). Understanding the interactions between the plasmodiophorid vector and the virus is limited due to the biological challenges of studying this system. Additionally, plasmodiophorid-vectored viruses often persist in the soil for decades, and the lack of effective control strategies makes them difficult to manage (Merz 1995). The potato mop-top virus (PMTV) is a rapidly growing threat to commercial potato production in the United States and is vectored by the plasmodiophorid, *Spongospora subterranea* f. sp. *subterranea* (*S. subterranea*), the causal agent of powdery scab disease (Jones and Harrison 1969; Lambert et al. 2003; Santala et al. 2010; Whitworth and Crosslin 2013). Synergistic infection with the pathogens causes qualitative damage to the potato host tuber. *S. subterranea* infection not only damages the tuber but also potato roots. The life cycle of *S. subterranea* is complex, involving the production of zoospores that initially cause root galls early in the growing season and scab-like lesions on the surface of the tuber later in the season (Merz 2008). PMTV-induced necrotic rings or arcs inside the tuber are characteristic of virus infection. Additionally, a yellowing chevron pattern may appear on the leaves of tuber-infected potato plants as an indication of PMTV infection (Calvert and Harrison 1966). PMTV incidence and/or symptoms are not always uniform within a field, presenting challenges for detection and management (Kirk 2008).

PMTV is a member of the *Pomovirus* genus, and its tripartite genome consists of RNA-RdRP, RNA-CP, and RNA-TGB (Savenkov et al. 1999; Reavy et al. 1998; Scott et al.

1994). RNA-RdRP is responsible for replication and contains the RNA-dependent replicase protein (Savenkov et al. 1999). RNA-CP encodes the coat protein and a coat-protein read-through (RT) domain which is involved in vector transmission (Reavy et al. 1998). RNA-TGB contains three triple gene bock proteins (TGB1, TGB2, TGB3), which play a crucial role in virus movement within the plant (Scott et al. 1994; Torrance et al. 2009). The PMTV genome shows low genetic variability globally, with more diversity found in regions like Peru and Colombia where potato was first domesticated (Beuch et al. 2015; Hu et al. 2016; Latvala-Kilby et al. 2009; Ramesh et al. 2014; Zhai et al. 2020; Kalayandurg et al., 2017). Studies have shown close relation amongst United States isolates and with European isolates, with amino acid sequence identity ranging from 96-99% (Ramesh et al. 2014; Zhai et al. 2019). Additionally, New Zealand isolates have been shown to come from the same lineage and with limited diversity (Frampton et al. 2022). Low genetic variability is uncommon in viruses due to the nature of replication which excludes a proofreading RNA polymerase (Pfeiffer and Kirkegaard 2005). Mutations are essential to virus fitness and virulence. It is speculated that PMTV's low genetic variability could be a result of the tripartite nature of PMTV's genome and the virus-vector relationship, but further investigation is needed (Kalyandurg et al. 2017). Despite low genetic variability, there are two strains of PMTV, severe (S) and mild (M). Mild PMTV is only found in Peru and Colombia (Kalyandrug et al. 2017). Strain type is characterized by a motif in the RT region of RNA-CP which creates a transmembrane protein and correlates to symptom development in the host. The seven conserved amino acids are different for S (RRSIGAV) and M (KKNVREM). Recent research discovered the CPRT gene frequently experiences deletions during mechanical

transmission, which can impact vector transmission due to disruptions in the associated transmembrane protein (Reavy et al. 1998; Sandgren et al. 2001; Kalyandrug et al 2017). Additionally, a separate study identified positive selection acting on codon 689 within the RT domain (Zhai et al. 2020). Given these findings, the CPRT gene remains an intriguing area to investigate. With only six full-length PMTV genomes available from the US, we aimed to expand this dataset by sequencing additional PMTV genomes.

To address a knowledge gap, our study focused on the genomic characterization of PMTV in the United States. Specifically, we investigated the RNA-CP and the RT domain. We conducted a comprehensive study capturing PMTV genomes from potato growing regions across the United States. Our objectives were to examine PMTV diversity and analyze the CPRT gene. We investigated PMTV's genome using next generation technologies which allowed for the exploration of PMTV's conserved 5' and 3' ends. A tomato-bait assay was successfully used to increase pathogen loads. Ultimately, we aspired to obtain a deeper understanding of the pathosystem using molecular methods.

MATERIALS AND METHODS

Soil sample collection/spore count

The soil samples used in this study were acquired from a larger effort in collaboration with the Charkowski lab (CSU). The collaborate soil sample collection was performed across five potato-growing states including Colorado, Oregon, North Dakota, Minnesota, and Maine. For proximity, North Dakota and Minnesota were counted as one state when sampling. Within those states, four locations were tested with four plots – totaling 64 sites. The soil was collected in gallon-sized bags at six-time points throughout the potato growing season in 2021, from May to September. The timepoint, in August or September 2021 (T4), before vine-kill, was used in this study. From the soil samples collected at T4 only 11 were used, including one or more plots from each location. More than one plot from each location was used if more than one cultivar was planted at that site. For example, if a location had different cultivars in its plots, we collected from multiple plots, whereas in locations with a singular cultivar in all plots, soil from only one plot was used. After sample collection in gallon sized-bags, soil was stored in boxes in a cool room. The spores within the soil was counted by Dr. Yuan Zheng (Charkowski lab) using a hemocytometer. The spore count was recorded and used for our study in later analysis.

Tomato bait-assay

A tomato bioassay was conducted in biological triplicates using 11 soil samples from different states and sites. The experimental design also consisted of *S. subterranea-*

free controls, three per state, totaling 12 controls. The control soils were autoclaved twice to ensure the removal of *S. subterranea* and other pathogens (Alaryn et al. 2023). The bioassay was performed in a hydroponic-like set up using a nutrient solution (NS) for tomatoes (Merz 1989). The set up consists of 45 autoclaved 4oz mason jars, one per tomato. The jar lids contain a ¼ inch hole drilled in the top for stem insertion, allowing the roots to suspend in 90 ml of NS. The NS consists of a 4-18-38 (NPK) tomato formula from Masterblend® Premium Fertilizers. The solution was made fresh every two weeks by mixing the soluble components with di water as directed on the package. Tomatoes (cv. Alaskan Fancy) were germinated on a petri dish (moist) for five days in the dark. On day five, 3 ml of NS was aseptically pipetted on the petri dish. The petri dishes were returned to the dark, where the tomatoes grew for two more days. After a total of seven days, once cotyledon is established, the tomatoes were transferred to the hydroponic set-up with fresh NS. The tomatoes were grown in the system for two weeks in a growth chamber at 16°C on a 16 h light regiment (Yu et al. 2023. After two weeks, the tomatoes were inoculated by changing the original NS solution with fresh NS and soil inoculum. To prepare inoculum, the soil was sieved through two kitchen sieves (4 and 2 mm) to remove contaminates (Alaryn et al. 2023). Sieved soil was suspended in 1 ml to 5 ml of water and spores were recounted with a hemocytometer to calculate dilution for uniform inoculations. The resulting soil was used to produce inoculum consisting of approximately 3,000 sporosori suspended in 12 ml of nutrient solution in a 15 ml conical tube. The tube was placed on a rocking vortexter for one hour. The tubes were put in the dark for 24 hours and were inverted several times by hand before transferring to the jars with 78 ml of fresh NS. Tomatoes were baited in the

chamber for two weeks before tissue collection. Tomato roots were washed with di water and patted dry. A total of 4-6 inches of the root tissue per tomato was collected in a 1.5 ml tube and immediately placed in liquid nitrogen to deter degradation of RNA. Samples were stored at -80°C until further processing.

RNA extraction

Using liquid nitrogen, the collected roots were homogenized in the 1.5 ml tube with blue pestles. After homogenization, the samples were weighed to ensure 100-150 mg of tissue was processed. The RNA extraction was performed using the Zymo Direct-zol™ RNA miniprep plus per the manufacturer's protocol (Zymo, Irvine, California). Ground tissue was combined with 600 µl of TRI Reagent® and thawed on ice. The tubes were centrifuged at maximum speed for 3 minutes, and 600 µl of the supernatant was transferred to a new tube. An equal volume of 95% ethanol was added to the supernatant, and the resulting mixture was transferred to a Zymo-Spin™ IIICG column. The column was centrifuged at 12,000 rpm for 1 minute, and the resulting flow-through was discarded. Following this step, 400 µl of wash buffer was added to the column, and another round of centrifugation was performed. DNase treatment was carried out by incubating the column for 15 minutes at room temperature. Subsequently, the column was washed twice with 400 µl of pre-wash buffer, followed by a final wash with 700 µl of buffer. The tubes were centrifuged for an additional 2 minutes to ensure the complete removal of residual wash buffer. RNA was eluted from the column matrix using 50 µl of DNase/RNase-Free water. RNA concentration and quality were checked in the lab using nanodrop values and suitable 260/280 scores.

The final 50 µl of RNA was stored in two separate tubes (25 µl each); one was used for cDNA synthesis, and the other for sequencing.

cDNA Synthesis

cDNA synthesis was performed using 1 µg of RNA in the GoScript™ Reverse Transcription System (Promega, Fitchburg, WI) according to the manufacturer's protocol. RNA, 1 µl of Oligo(dT)15 primer (500 µg/ml), and 1 µl of random hexamers (500 µg/ml) were incubated at 70°C for 5 minutes, followed by incubation on ice for 5 minutes. Additional components required for cDNA synthesis, including 4 µl of GoScript™ 5X reaction buffer, 1.2 µl of MgCl2, 1 µl of dNTPs, 20 units of Recombinant RNasin® ribonuclease inhibitor, and 1 µl of GoScript™ reverse transcriptase, were then added. The final volume was brought to 15 µl by adding Nuclease-free water. cDNA synthesis was carried out using the following thermocycling protocol: an initial incubation at 25°C for 5 minutes, reverse transcription at 42°C for 60 minutes, and an inactivation step at 70°C for 15 minutes.

Quantitative Polymerase chain reaction (qPCR)

Root sample cDNA was processed by qPCR to verify infection of both virus and vector before proceeding with sequencing. An initial qPCR was done using tomato housekeeping primers (EF1) to ensure RNA quality (Lacerda et al. 2015). PMTV primers described and tested by Pandey et al. 2020, detecting the coat protein (CP) of RNA-CP were used. The primer set designed by Nei et al. 2021 was also used to confirm the presence of *S. subterranea* by detecting the ITS region of the genome. qPCR reactions were

performed individually for each primer with SsoAdvanced™ Universal SYBR® Green Supermix (BioRad, Hercules, CA). The reactions were run with a total reaction volume of 20 µl composed of 10 µl of SsoAdvanced™ Universal SYBR® Green Supermix, 2 µl of cDNA, and 0.5 µl of each primer (250 nM); the total volume of 20 µl was achieved with nuclease-free water. The thermocycling protocol was as follows: 95°C for 30 sec, followed by 40 cycles of 95°C for 15 sec, annealing at 60°C for 30 sec. Each tomato root sample was qPCRed with each primer set in technical triplicates, and Cq values were recorded.

Whole-transcriptome sequencing and analysis

A total of eight PMTV positive samples were sent for sequencing, two samples per state. Samples with high Cq values, determined by the qPCR, were chosen for whole transcriptome sequencing. RNA (1 µg) was submitted to Novogene (Novogene Corporation Inc., Sacramento, CA), where RNA quality was further examined, and only RIN values above 4.0. were accepted. Libraries were constructed and sequenced using an Illumina NovaSeq. Bioinformatic analysis was performed on Qiagen's CLC Genomics workbench (23.0.2) following the pipeline shown in Figure 3.1. First, the adapters were trimmed, and low-quality reads were removed using Illumina pipeline 1.5-1.7. Then, the reads aligning with the tomato genome (SL4.0) were removed. *De novo* assembly was performed on the remaining reads. All *de novo* assembled contigs were blasted against NCBI's nucleotide (nt) database (BLASTn, November 16, 2023). Stringent coverage (>50%) and percent identity (>80%) cutoffs were used in the BLAST search. Results annotated with "Potato mop-top virus," exhibiting a percent identity greater than 80%, were extracted to create a

consensus sequence for each RNA. Specific PMTV-positive contigs were successfully identified using this method. Contigs varied in length; some were full-length, while others were partial (Table 3.1). Only one genome per virus is available in the BLAST search using this method (GCF_000850685.1). Therefore, the recovered contigs were individually blasted on NCBI's website. The individual BLAST allowed a more precise search, giving us the closest BLAST relative. The closest BLAST relative was recorded, and the accession number was used in an additional read alignment step. After the host genome removal, the remaining reads aligned to the closest BLAST relative of each RNA of PMTV's multipartite genome. The number of reads mapping to the genome was recorded, and a visual examination was then performed to assess coverage and depth of read alignment with the PMTV RNA genome. In some cases, reads could fill gaps where contigs were missing. A consensus read sequence was extracted from the reads alignment step. Combining contigs and reads allowed us to extract a consensus sequence for seven of the eight genomes with three RNAs per sample (Table 3.1).

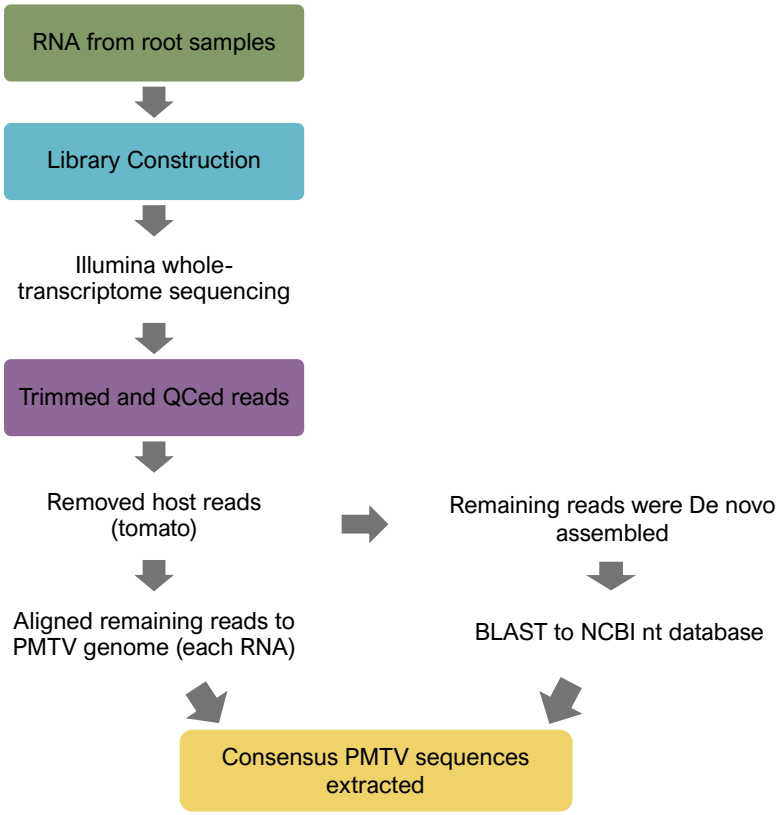

Figure 3.1 Pipeline used to extract PMTV genomes from Whole Transcriptome Sequencing. RNA extracted from the tomato bait assay was subjected to library construction and Illumina NovaSeq. Preprocessing and trimming of the raw sequencing data were performed using Qiagen's CLC workbench and Illumina's pipeline (version 1.5-1.7). Tomato host reads were removed before alignment to the PMTV genome and De novo assembly of remaining reads. De novo assembled reads were subjected to a BLAST search to recover reads similar to a PMTV isolate from NCBI. Reads acquired from the BLAST and consensus sequences were cross-examined to create full-length PMTV genomes.

Alignment and phylogenetic trees

Read and contig-derived consensus sequences were used in phylogenetic analysis. Consensus sequences from this study were aligned with 31 reference sequences in MEGAX using MUSCLE for each RNA of the tripartite genome. The alignment was exported as a fasta file and used in IQtree to create a phylogenetic tree of RNA-CP. The fasta file was checked in IQtree using the iqtree2 function and a partition file to ensure symmetry homogeneous sequence evolution assumptions (Naser-Khdour et al. 2019). A maximum likelihood tree with a bootstrap value of 1000 was created using iqtree2, where the substitution model was determined, and a tree was created. The tree was then exported as a .ufboot file and transferred to Geneious Prime. The tree was visualized in Geneious Prime and exported to be generated as a figure. To further investigate RNA-CP, the amino acids of the CPRT gene were aligned. First, consensus sequences from this study for RNA-CP were aligned in MEGAX using MUSCLE along with three reference genomes. Reference sequences from suspected recombinant C61 isolates and noted severe and mild strains were used in the alignment. After the sequences were aligned, the 5' and 3' UTR regions were manually removed, and the nucleic acids were translated to amino acids in MEGA. The amnio acids were extracted from MEGA and saved in a fasta format for better visualization. The fasta file was then uploaded to SnapGeneViewer (v 7.2.0) to make viewing codons sites easier. The RT domain, specifically the seven notable codons previously described by Kalyandrung et al. 2017, were examined and amino acids were recorded at codons 289, 464, 482, 592, 604, 686, and 757.

RESULTS

WTS extracts PMTV genomes from tomato-baited roots

PMTV genomes were constructed from contigs and reads recovered from whole-transcriptome sequencing. ME32 exhibited insufficient virus coverage among the eight samples and was excluded from all phylogenetic analyses. Four of the remaining seven samples had complete contig coverage of the tripartite genome, meaning 12 full-length contigs were extracted. The samples with complete continuous coverage include CO14, OR738, ND41, and ME31. The other three samples had variations of full-length and partial contigs (Table 3.1). Parameters used in the BLAST search allowed us to recover contigs with high coverage and identity to the PMTV genome available in NCBI's nt database. Given only one genome per virus is available in the BLAST search, an individual BLAST was performed to acquire the closest BLAST relative (100% nucleotide identity), which was used in an additional read alignment step. The additional read alignment step allowed coverage to be visualized, and more PMTV reads were collected. Consensus read sequences made up for the lack of contigs for samples OR741, ND24, and CO24, thus accounting for the other three genomes that initially lacked contig consensus. Our method successfully acquired seven new PMTV isolate sequences in the United States.

Overall, Maine had the lowest number of reads while Colorado had the highest; North Dakota/Minnesota and Oregon had a relatively similar number of reads (Table 3.1). An ANOVA test was performed, and no statistical difference in the number of reads was found amongst states. Every sample had the highest number of reads aligning to RNA-TGB and the lowest amount for RNA-RdRP.

Table 3.1 Summary of WTS analysis and PMTV census sequences.

Sample ID	RNA	# of total reads	# reads mapping to RNA	% reads mapping to RNA	# contigs	Mean contig length	Consensus Length or HSP*	Percent Genome coverage	Closest BLASTRelative (% of nt Identity)+	Accession Number of BLAST result
CO14	RdRP	45814394	281567	0.61	2	3336.5	4570	100	100	KU955473
	CPRT	45814394	305058	0.67	4	749	1775	100	100	KR857359
	TGB	45814394	690370	1.51	1	1775	2564	100	100	KR857357
CO24	RdRP	45316860	26	0.00	1	536	540	65.4	100	MK539859
	CPRT	45316860	32	0.00	1	367	367	82.5	100	MK539873
	TGB	45316860	47	0.00	2	636	713	74.3	100	KY284862
OR738	RdRP	43146716	239767	0.56	1	5918	5901	100	100	KU955473
	CPRT	43146716	218069	0.51	0	0	0	100		
	TGB	43146716	682375	1.58	1	3130	3100	100	100	OP221273
OR741	RdRP	45124294	112	0.00	4	596	726	63.5	100	MK539859
	CPRT	45124294	84	0.00	0	0	0	76.6		
	TGB	45124294	316	0.00	1	2012	2012	100	100	MK539868
ND24	RdRP	49419144	110	0.00	2	629	614	88.9	100	KR857358
	CPRT	49419144	196	0.00	2	628	711	100	100	MK539873
	TGB	49419144	441	0.00	2	1784	2944	100	100	KR857357
ND41	RdRP	45764084	170214	0.37	1	5721	5721	100	100	KR857349
	CPRT	45764084	242237	0.53	1	3121	3119	100	100	KR857359
	TGB	45764084	755894	1.65	1	2622	2601	100	100	KR857357
ME13	RdRP	54324504	94	0.00	3	571	459	100	100	KY275269
	CPRT	54324504	44	0.00	1	433	356	100	100	KY284863
	TGB	54324504	154	0.00	2	1439	506	100	100	KY284862
ME32	RdRP	50479640	16	0.00	0	0	0	23.3		
	CPRT	50479640	6	0.00	0	0	0	17.6		
	TGB	50479640	40	0.00	0	0	0	70		

Genome alignment and phylogenetic analysis reveal an amino a change in a conserved motif

To determine the relationship between the newly identified isolates and known PMTV strains, a phylogenetic tree was constructed. The read and contig derived consensus sequences for RNA-CP were used for alignment and phylogenetic analysis (Figure 3.2). The maximum likelihood tree showed distinct clades for severe, mild, and recombinant strains. All isolates found in this study grouped with severe isolates. The mild strain isolates were from Peru or Colombia, whereas severe strains were mostly from Europe and the North America.

To further investigate RNA-CP, nucleotide sequences were translated to amnio acid sequences to explore the transmembrane protein composed of seven amino acids in the CPRT region. The alignment revealed almost all the isolates found in this study contained the conserved amino acids (RRSIGAV) in PMTVs severe strains, besides sample ND24 (Figure 3.3). At codon 289, sample ND24 has a K (lysine) instead of a R (arginine). When comparing this substitution to mild and recombinant (mild/severe) strains, both strain types also have a K at this codon. When considering this non-synonymous substitution in our sample of interest, ND24's position within our phylogenetic tree was cross examined. ND24 grouped with severe strains, specifically creating a subclade with a Maine isolate. Although ND24 has a K (indicative of mild strains) instead of an R (indicative of severe strains) at codon 289, the phylogenetic analysis does not show close relation to recombinant isolates.

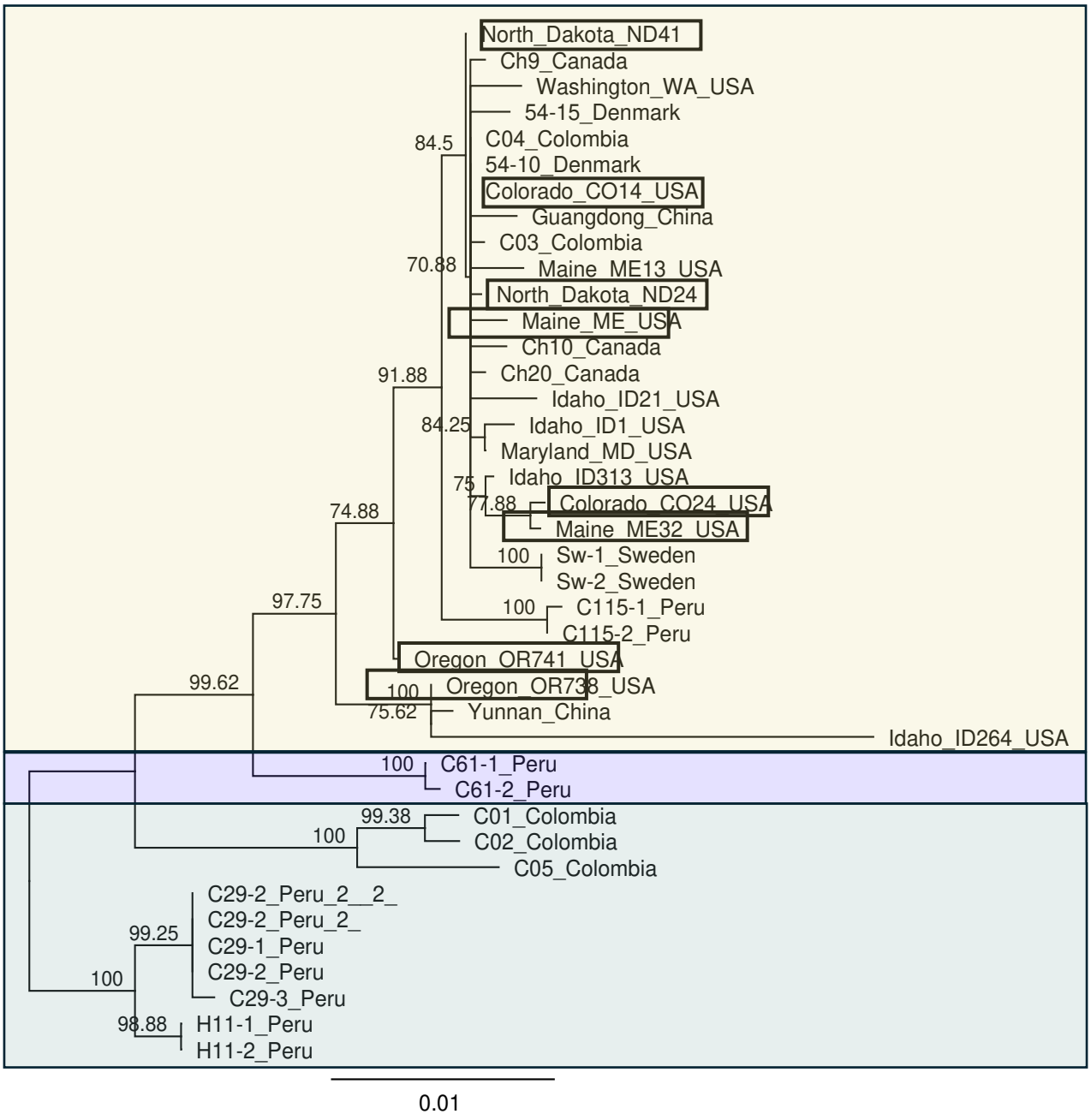

Figure 3.2 Phylogenetic tree and strain analysis of new PMTV isolates. A maximum likelihood tree was created using full-length RNA-CP sequences from this study and references sequences. The tree was generated using IQtree following alignment in MEGAX. The GTR+G+I substitution model and a bootstrap value of 1000 was employed. Branch lengths represent the number of substitutions per site. Sequences from this study are outlined by a black box and strains are highlighted by color. Isolates that fall within the yellow box are severe, purple indicates recombinant severe/mild, and green represents mild strains.

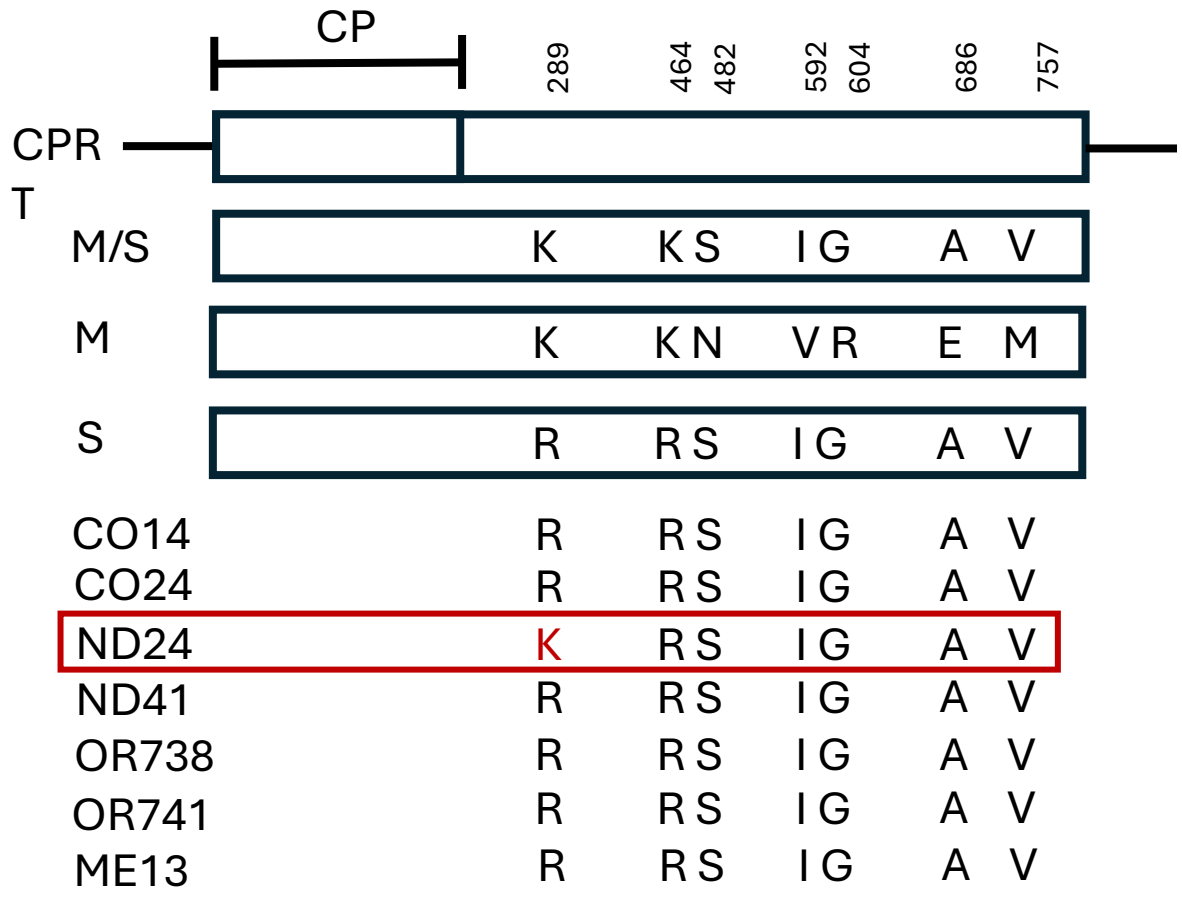

Figure 3.3 Alignment of CPRT shows one polymorphism. New isolate sequences were aligned using MUSCLE in MEGAX with references sequences before being converted to amino acids. Nucleic acids where recorded at codons of interest and observed for polymorphisms.

Table 3.2 Primers used to amplify PMTV genome and the mismatching nucleotides

RNA	Primer		Position
RNA-RdRP	R1R1	AAATCGCGCATAATCTCCAC	2762
RNA-RdRP	R1F2	GTCGAAGACTCCGAGGTCAG	2552
RNA-RdRP	R1R2	AGCTCGTCAGACGTCATCCT	4769
RNA-CP	R3F2	GATTATTTTGATATTATTGAAGCGAAG	1604
RNA-TGB	R2R1	GCGTAGACTCTATACCGCACTCA	1540

Tomato bait assay and reads counts show pathogen loads

The plant bioassay was successful in amplifying PMTV viral loads for sequencing apart from sample ME32. Although all tomato root RNA was tested with qPCR before sequencing, there was not enough viral reads acquired during sequencing for sample ME32. Successful baiting was based on Cq values, higher than 30. For all locations, negative controls resulted in no amplification or very high Cq values (>32) during qPCR analysis. Three different primer sets were used to assess tomato root RNA, one to amplify the housekeeping gene of tomato, and the other two looking at the virus-vector system evaluating the ITS region of *S. subterranea* and the coat-protein of RNA-CP. qPCR was performed on each sample in biological triplicates and technical triplicates, totaling nine data points per sample to ensure pathogen infection and RNA quality.

DISCUSSION

Potato mop-top virus exhibits unusually low genetic variability, which is atypical for successful viruses (Beuch et al., 2015; Ramesh et al., 2014; Zhai et al., 2020). Previous studies have emphasized the scarcity of diversity in PMTV populations outside potato-originating regions, the only location where both mild and severe strains coexist. Our findings from newly characterized PMTV isolates align with existing research. Phylogenetic analysis categorized all isolates as severe and grouped them with other North American and European isolates. The seven amino acid motif in the RT domain of RNA-CP was identified in six of the seven isolates and was conserved in all samples but one, ND24, where one amino acid difference was present. Our study was able to close a knowledge gap by uncovering more PMTV genomes in the United States.

Whole transcriptome sequencing (WTS) successfully sequenced seven new PMTV isolates from the United States, including Colorado, North Dakota/Minnesota, Oregon, and Maine. Illumina technology allowed an unbiased approach to sequencing new PMTV isolates by avoiding traditional primers and the well-conserved 5- and 3-termini of the PMTV genome (Gil et al., 2016). Although our study supported well-conserved ends of the genome, we found mismatches in primers that traditionally amplified inner regions of the genome (Table 3.1). Specifically, primers used to amplify RNA-RdRP in United States isolates favor identifying severe strains over mild due to 2-3 mismatches in the primer sequence and genome (Ramesh et al., 2014). Although only severe strains have been detected in the United States, this could prevent future detection of mild isolates. Designing primers for unknown sequences presents the risk of failing to amplify the

intended region. Mis-priming is one of many challenges of using traditional cloning methods. A study from New Zealand divided the genome into several 600-800 bp sections to assemble the genome and construct a full-length sequence (Frampton et al., 2022). This method excludes a cloning step, which other groups have used to ensure complete, accurate sequences. Both methods are labor-intensive and require numerous PCR steps (Ramesh et al., 2014; Scott et al., 1994). The steps include cDNA synthesis and several PCRs, risking the introduction of unintended mutations. As an alternative to traditional methods, our study pursued Whole Transcriptome Sequencing (WTS) to reduce lab time for more rapid results. WTS was chosen over other NGS methods to address the distinctive genome feature of PMTV, which lacks poly-A tails. Our study showed how NGS technologies can sequence a challenging plant virus. With this method, we increased the genome's coverage and depth to ensure our consensus sequences' confidence. Coverage and depth made it easy to explore the CPRT gene, our primary RNA of interest.

RNA-CP characterizes PMTV strain and, more importantly, plays a role in virus transmission (Adams et al., 2001). The conserved motif responsible for the pathobiological differences in strains was discovered by Kalyandurg et al. 2017, and was, therefore, our main area of interest. We found that six of our seven samples have the conserved severe motif (RRSIGAV), but one sample, ND24, has a modification in the motif with a K instead of an R at codon 289 (KRSIGAV). A recent study found codon 689 in the RT domain to be under positive selection based on ω, or dN/dS (dN: rate of non-synonymous substitutions and dS: rate of synonymous substitutions) (Zhai et al., 2020). Interestingly, we found no changes at codon 689 and instead found the substitution at codon 289. Although Zhai et

al. 2020 and Kalyandurg et al. 2017 did not find the entire CP-RT protein in RNA-CP under diversifying selection, they noted a lack of available sequences to be a limitation in their conclusions. Positive selection is critical in virus adaptation and, in some cases, is essential to virus/host interactions. In the case of PMTV, the RT domain is crucial for vector transmission, likely making it a region opposed to change because the virus relies on the vector for host infection. Furthermore, deletions in the CP-RT gene have been observed in various studies affecting transmissibility, specifically during mechanical inoculations (Reavy et al., 1998; Sandgren et al., 2001; Torrance, 2008). We detected no deletions in the CP-RT protein, likely due to the effects deletions can have on pathogenic abilities, which would be detrimental to PMTV's success in nature.

Our goal was to characterize new isolates and examine their relationship with other United States isolates. We found all new isolates could be categorized as severe and did not create a unique clade. Notably, there were distinct clades for Peruvian and Colombian isolates. In contrast, the isolates from North America and Europe were more sporadic within the severe clade, showing close relations amongst those isolates. Our phylogenetic tree also showed a distinct clade for the possible C61 recombinant isolates. Although there are conflicting conclusions about whether C61 is a true recombinant, it is still important to note the recombinant-like amino acid motif in the RT domain (Kalyandurg et al., 2017; Zhai et al., 2020). Our sample of interest, ND24, which has a K instead of an R at codon 289, like C61, does not show close relations to the recombinant clade based on phylogenetic analysis.

Due to *S. subterranea*'s obligate pathogen nature, it is common to use a plant bioassay to overcome the challenge of culturing the pathogen. We decided to use a hydroponic system based initially on a protocol published in 1989 (Merz, 1989). Various groups have modified the protocol, and our modifications included using a growth chamber with a controlled setting and soil instead of hydroponics (Alaryan et al., 2023; Davey, 2009 ; Van De Graaf et al., 2003). The hydroponic system was chosen to increase virus titer and look at *S. subterranea*, while excluding soil factors that could affect baiting. The plants were grown in a chamber at 16°C, which has been proven to be the optimal temperature for *S. subterranea* root attachment; the cold temperature also helps reduce unwanted contamination (Yu et al., 2023). We found our modified tomato bait assay based on qPCR effective in increasing virus titers.

Conclusion

Phylogenetic data was generated from this study using new PMTV genomes to better understand the PMTV genomes in the United States. Further investigation of PMTV isolates confirmed a lack of variation and discarded the hypothesis that it could be due to a knowledge gap. Additionally, the genomes' 5' and 3' terminus were highly conserved, suggesting the use of primers amplifying the ends of the genome is sufficient when attempting to sequence the genome.

REFERENCES

Adams, M. J., Antoniw, J. F., & Mullins, J. G. L. (2001). Plant virus transmission by plasmodiophorid fungi is associated with distinctive transmembrane regions of virus-encoded proteins. Archives of Virology, 146(6), 1139–1153. https://doi.org/10.1007/S007050170111/METRICS

Alaryan, M. M., Zeng, Y., Fulladolsa, A. C., & Charkowski, A. O. (2023). Brassica Cover Crops and Natural Spongospora subterranea Infestation of Peat-Based Potting Mix May Increase Powdery Scab Risk on Potato. Plant Disease, 107(9), 2769–2777. https://doi.org/10.1094/PDIS-04-22-0863-RE

Beuch, U., Berlin, S., Åkerblom, J., Nicolaisen, M., Nielsen, S. L., Crosslin, J. M., Hamm, P. B., Santala, J., Valkonen, J. P. T., & Kvarnheden, A. (2015). Diversity and evolution of potato mop-top virus. Archives of Virology, 160(5), 1345–1351. https://doi.org/10.1007/S00705-015-2381-7/FIGURES/2

Calvert, E. L., & Harrison, B. D. (1966). POTATO MOP-TOP, A SOIL-BORNE VIRUS. Plant Pathology, 15(3), 134–139. https://doi.org/10.1111/J.1365-3059.1966.TB00333.X

Davey, T. (2009). An accelerated soil bait assay for the detection of potato mop top virus in agricultural soil. Methods in Molecular Biology (Clifton, N.J.), 508, 259–265. https://doi.org/10.1007/978-1-59745-062-1_20/TABLES/1

Frampton, R. A., Addison, S. M., Kalamorz, F., & Smith, G. R. (2022). Genomes of Potato Mop-Top Virus (Virgaviridae: Pomovirus) Isolates from New Zealand and Their Impact on Diagnostic Methods. Plant Disease, 106(10), 2571–2575. https://doi.org/10.1094/PDIS-01-22-0192-SC

Gil, J. F., Adams, I., Boonham, N., Nielsen, S. L., & Nicolaisen, M. (2016). Molecular and biological characterization of Potato mop-top virus (PMTV, Pomovirus) isolates from the potato-growing regions of Colombia. Plant Pathology, 65(7), 1210–1220. https://doi.org/10.1111/ppa.12491

Hu, X., Dickison, V., Lei, Y., He, C., Singh, M., Yang, Y., Xiong, X., & Nie, X. (2016). Molecular characterization of Potato mop-top virus isolates from China and Canada and development of RT-PCR differentiation of two sequence variant groups. Canadian Journal of Plant Pathology, 38(2), 231–242. https://doi.org/10.1080/07060661.2016.1189968

Jones, R. A. C., & Harrison, B. D. (1969). The behaviour of potato mop-top virus in soil, and evidence for its transmission by Spongospora subterranea (Wallr.) Lagerh. Annals of Applied Biology, 63(1), 1–17. https://doi.org/10.1111/j.1744-7348.1969.tb05461.x

Kalyandurg, P., Gil, J. F., Lukhovitskaya, N. I., Flores, B., Müller, G., Chuquillanqui, C., Palomino, L., Monjane, A., Barker, I., Kreuze, J., & Savenkov, E. I. (2017). Molecular and pathobiological characterization of 61 Potato mop-top virus full-length cDNAs reveals great variability of the virus in the centre of potato domestication, novel genotypes and evidence for recombination. Molecular Plant Pathology, 18(6), 864–877. https://doi.org/10.1111/mpp.12552

Kirk, H. G. (2008). Mop-top virus, relationship to its vector. American Journal of Potato Research, 85(4), 261–265. https://doi.org/10.1007/s12230-008-9021-7

Lambert, D. H., Levy, L., Mavrodieva, V. A., Johnson, S. B., Babcock, M. J., & Vayda, M. E. (2007). First Report of Potato mop-top virus on Potato from the United States. Https://Doi.Org/10.1094/PDIS.2003.87.7.872A, 87(7), 872–872. https://doi.org/10.1094/PDIS.2003.87.7.872A

Merz, U. (n.d.). Infectivity, inoculum density and germination of Spongospora subterranea resting spores: a soiution-culture test system'.

Merz, U. (2008). Powdery scab of potato - Occurrence, life cycle and epidemiology. American Journal of Potato Research, 85(4), 241–246. https://doi.org/10.1007/S12230-008-9019-1/FIGURES/2

Pfeiffer, J. K., & Kirkegaard, K. (2005). Increased Fidelity Reduces Poliovirus Fitness and Virulence under Selective Pressure in Mice. PLOS Pathogens, 1(2), e11. https://doi.org/10.1371/JOURNAL.PPAT.0010011

Ramesh, S. V., Raikhy, G., Brown, C. R., Whitworth, J. L., & Pappu, H. R. (2014). Complete genomic characterization of a potato mop-top virus isolate from the United States. Archives of Virology, 159(12), 3427–3433. https://doi.org/10.1007/S00705-014-2214-0/FIGURES/4

Reavy, B., Arif, M., Cowan, G. H., & Torrance, L. (1998). Association of sequences in the coat protein/readthrough domain of potato mop-top virus with transmission by Spongospora subterranea. Journal of General Virology, 79(10), 2343–2347. https://doi.org/10.1099/0022-1317-79-10-2343/CITE/REFWORKS

Sandgren, M., Savenkov, E. I., & Valkonen, J. P. T. (2001). The readthrough region of Potato mop-top virus (PMTV) coat protein encoding RNA, the second largest RNA of PMTV genome, undergoes structural changes in naturally infected and experimentally inoculated plants. In Arch Virol (Vol. 146). http://bioweb.pasteur.fr/seqanal/interfaces/mfold-simple.html

Santala, J., Samuilova, O., Hannukkala, A., Latvala, S., Kortemaa, H., Beuch, U., Kvarnheden, A., Persson, P., Topp, K., Ørstad, K., Spetz, C., Nielsen, S. L., Kirk, H. G., Budziszewska, M., Wieczorek, P., Obrepalska-Steplowska, A., Pospieszny, H., Kryszczuk, A., Sztangret-Wiśniewska, J., … Valkonen, J. P. T. (2010). Detection, distribution and control of Potato mop-top virus, a soil-borne virus, in northern Europe. Annals of Applied Biology, 157(2), 163–178. https://doi.org/10.1111/j.1744-7348.2010.00423.x

Sarwar, M., Aslam, M., Sarwar, S., & Iftikhar, R. (2020). Different nematodes and plasmodiophorids as vectors of plant viruses. In Applied Plant Virology: Advances, Detection, and Antiviral Strategies (pp. 275–290). Elsevier. https://doi.org/10.1016/B978-0-12-818654-1.00021-9

Savenkov, E. I., Sandgren, M., & Valkonen, J. P. T. (1999). Printed in Great Britain Complete sequence of RNA 1 and the presence of tRNA-like structures in all RNAs of Potato mop-top virus, genus Pomovirus. In Journal of General Virology (Vol. 80).

Scott, K. P., Kashiwazaki, S., Reavy, B., & Harrison, B. D. (1994). The nucleotide sequence of potato mop-top virus RNA 2: a novel type of genome organization for a furovirus. In Journal of General Virology (Vol. 3561).

Torrance, L. (2008). Pomovirus. In Encyclopedia of Virology: Volume 1-5 (Vols. 1–5, pp. V4-282-V4-287). Elsevier. https://doi.org/10.1016/B978-012374410-4.00543-4

Whitworth, J. L., & Crosslin, J. M. (2012). Detection of Potato mop top virus (Furovirus) on potato in southeast Idaho. Https://Doi.Org/10.1094/PDIS-08-12-0707-PDN, 97(1), 149. https://doi.org/10.1094/PDIS-08-12-0707-PDN

Yu, X., Tegg, R. S., Eyles, A., Wilson, A. J., & Wilson, C. R. (2023). Development and validation of a novel rapid in vitro assay for determining resistance of potato cultivars to root attachment by Spongospora subterranea zoospores. Plant Pathology, 72(2), 392–405. https://doi.org/10.1111/ppa.13659

Zhai, Y., Mallik, I., Hamid, A., Tabassum, A., Gudmestad, N., Gray, S. M., & Pappu, H. R. (2020). Genetic diversity in potato mop-top virus populations in the United States and a global analysis of the PMTV genome. European Journal of Plant Pathology, 156(2), 333–342. https://doi.org/10.1007/s10658-019-01836-6

CHAPTER 4 – POTATO MOP-TOP VIRUS DETECTION AND CORRELATION TO ITS VECTOR, *SPONGOSPORA SUBTERRANEA*

INTRODUCTION

Potato (*Solanum tuberosum*) is the fourth most important food crop worldwide and hosts many pathogens (Ezekiel et al., 2013). In particular, tuber necrotic viruses pose a significant economic burden on potato quality and production. One of the most notable potato necrotic viruses is the *Potato mop-top virus* (PMTV), a single-stranded positive-sense RNA virus belonging to the *Pomovirus* genus. PMTV is vectored by a pathogenic plasmodiophorid, *Spongospora subterranea* f. sp. *subterranea,* the causal agent of potato powdery scab disease (Calvert & Harrison, 1966). Both virus and vector infections result in symptoms problematic to the potato industry (Lambert et al., 2007; Santala et al., 2010; Whitworth & Crosslin, 2012). The complex life cycle of *S. subterranea* includes the production of zoospores, which initially cause the formation of root galls early in the growing season and scabs on tubers later in the season. Management of these pathogens relies on controlling conditions favorable to the vector. However, irrigation practices and spore longevity make management difficult (Kirk, 2008). Furthermore, the difficulties of detecting the soil-borne pathogens and knowledge gaps regarding their relationship add additional layers of complexity. PMTV is one of several economically important viruses vectored by a protist (Adams et al., 2001). The relationship between the protist and the virus is understudied due to challenges paired with studying these soil pathogens.

Issues with the accurate detection of PMTV could be due to the tripartite nature of its genome and the genetically complex relationship between the virus, its plasmodiophorid vector, and its host (Sarwar et al., 2020). The virus consists of three RNAs: RNA-RdRP,

responsible for replication, RNA-CP which encodes for the coat protein (CP); and RNA-TGB, which contains three triple-gene-block (TGB) proteins, traditionally responsible for movement in the host, all of which assemble as separate rod-shaped virions (Savenkov, 2020; Savenkov et al., 1999; Scott et al., 1994). Detection of PMTV cannot be based on visual symptoms as foliar mosaic symptoms vary, and tuber necrotic symptoms could result from other viruses, such as potato virus Y or tobacco rattle virus. Therefore, methods of PMTV detection usually involve time and labor-intensive plant bioassays, where bait plants are grown and observed for several weeks (Davey, 2009). More recently, a droplet digital PCR (ddPCR) assay was used to detect PMTV (Pandey et al., 2020). Primer sequences were designed based on highly conserved regions in RNA-CP, , but PMTV's ability to move systemically without this RNA may show this detection method may need require modification to detect both RNA-RdRP and RNA-TGB (Kalyandurg et al., 2017; Ramesh et al., 2014). PMTV can move in two forms: by encapsulation via the CP or in a ribonucleoprotein (RNP) complex assembled by a triple gene block protein, TGB1 (Savenkov, 2020) . As a result, potato pathologists across Europe have suggested that the other PMTV RNAs should be looked for during detection, but no study has demonstrated how this recommendation will fare (Savenkov, 2020; Zhai et al., 2020).

The characteristics of PMTV's genome challenge traditional virus models regarding its ability to travel through a host systemically without needing a CP. The coat protein-read-through gene (CP-RT) in RNA-CP, an extended version of the CP, is essential for vector transmission and vascular movement when in combination with the CP (Adams et al., 2001; Torrance et al., 2009). RNA-CP is of particular interest based on PMTV’s ability to

move without the CP, leaving an area to be explored when looking at PMTV detection. In the United States (US), only the severe strain of the virus (PMTV-S) has been found (Gau et al., 2013). PMTV-S strain causes severe disease in potatoes compared to the mild strain and is further characterized based on the CP-RT gene in RNA-CP. A signature motif of seven amino acids differs between the severe and mild strains (Kalyandurg et al., 2017; Zhai et al., 2020). A recent analysis of this region in newly available PMTV genomes in the US found the motif almost completely conserved along with the rest of the PMTV genome.

PMTV genome has very low genetic variability in the US, with a difference of less than 2% between all characterized isolates (Zhai et al., 2020). It is speculated that part of this could be due to the virus-vector nature: PMTV's ability to infect its host relies on its vector. The close association between PMTV and its vector, *S. subterranea*, makes it challenging to study PMTV's pathogenic properties individually. Since the pathogens cannot be readily isolated and/or separated from each other, understanding the diversity of *S. subterranea's* populations in the US is essential but has yet to be explored. Although *S. subterranea*'s genome cannot be easily sequenced, microsatellite markers are available to explore the protist's population genetics (Ciaghi et al., 2018; Gau et al., 2013). The markers have been used to investigate the genetic structure of *S. subterranea* populations in regions worldwide (Muzhinji and van der Waals 2019). One distinct population is seen in Colombia, while the other population is encompasses other regions such as Europe, the United States, and the Oceanias. Population genetic structure describes how genetic diversity is distributed within and across populations because of various evolutionary forces such as natural selection, mutation, genetic drift, recombination, and migration (McDonald and

Linde 2002). The forces facilitate adaptation to the environment and changes in ecological niches.

PMTV has co-evolved with its vector, *S. subterranea,* and their worldwide distributions coincide (Kirk, 2008). In other virus-vector systems in nature, the distribution and incidence of a virus should correspond to its vector due to the nature of plant virus transmission (Hogenhout et al., 2008). For PMTV and its vector, this relationship or correlation has not been investigated, mainly due to technical difficulties in accurately detecting both pathogens in soil samples. However, advances in PCR technologies allow for the accurate detection of both pathogens, making the study of relationships between PMTV and *S. subterranea* possible. The severity of PMTV infections in relation to *S. subterranea* infections has yet to be studied.

Our study gives insight into the novel virus-vector relationship using accurate detection methods of PMTV. We designed primer-probe sets for each RNA segment of the PMTV genome using newly available isolates from the US in efforts to investigate the virus-vector relationship in soil and to improve detection methods. To further understand the complex pathogenicity of PMTV, we explored viral pathogen titers using a tomato bait assay with inoculum from various potato growing regions in the United States. In the interest of the virus-vector relationship, we aimed to understand the diversity of *S. subterranea's* populations which were used as inoculum. Ultimately, we aspired to obtain a deeper understanding of the pathosystem using molecular methods.

MATERIALS AND METHODS

Soil sample collection/spore count

The soil samples used in this study were acquired from a larger effort in collaboration with the Charkowski lab (CSU). Soil sample collection was performed across five potato-growing states: Colorado, Oregon, North Dakota, Minnesota, and Maine. For the purpose of simplifying this study and proximity, North Dakota and Minnesota were counted as one state. Within the four states, four different locations were tested, each consisting of four plots within – totaling 64 sites (Table 4.1). The soil was collected in gallon-sized bags at six-time points throughout the potato growing season in 2021, from May to September. The timepoint, in August 2021 (T4) prior to vine-kill, was used in this study. From the soil samples collected at T4, only 11 sites were used, including one or more plots from each location. More than one plot from each location was used if more than one cultivar was planted at that site. For example, if a location had different cultivars in its plots, we collected from multiple plots. In contrast, in locations that only had a singular cultivar in all plots, soil from only one plot was used. After sample collection in gallon-sized bags, soil was stored in boxes in a cool room. Dr. Yuan Zheng (Charkowski lab) counted the spores within the soil using a hemocytometer. The spore counts were recorded and used for our study in later analysis.

Table 4.1 Sample collection and sample names

State	Site	Plot	Sample Name	Timepoint	Cultivar at location
Colorado	CO-S1	plot 2	COP1P2	8/18	Yukon Gold
		plot 4	COP1P4	8/18	Austrain Crescent
	CO-S2	plot 4	COP2P4	9/22	Ruesset Nortkotah 3
	CO-S3	plot 2	COP3P2	10/5	Revellie Ruesset
	CO-S4	plot 1	COP4P1	9/7	Russet Norkotah 8
		plot 2	COP4P2	9/7	Rio Grande
		plot 4	COP4P4	9/9	Russet Nortkotah 296
Oregon	OR-738	plot 2	OR738	august	Umatilla Russet
	OR-741	plot 4	OR741	august	Umatilla Russet
	OR-745	plot 1	OR745	august	Umatilla Russet
	OR-746	plot 2	OR746	august	Umatilla Russet
North Dakota/Minnesota	ND-F1	plot 4	NDF1P4	august	Umatilla Russet
	ND-F2	plot 4	NDF2P4	august	Umatilla Russet
	ND-F3	plot 4	NDF3P4	august	Umatilla Russet
	ND-F4	plot 1	NDF4P1	august	Umatilla Russet
Maine	ME-F1	plot 1	MEF1P1	august	Mountain Gem
		plot 2	MEF1P2	august	Caribou Russet
		plot 3	MEF1P3	august	Russet Burbank
	ME-F2	plot 4	MEF2P4	august	Soraya
	ME-F3	plot 1	MEF3P1	august	Dark Red Norland
		plot 2	MEF3P2	august	Yukon Gold
		plot 3	MEF3P3	august	CC23-Corolla
		plot 4	MEF3P4	august	BR2-Baltic Rose
	ME-F4	plot 1	MEF4P1	september	Yukon Gold
		plot 2	MEF4P2	september	Katahdin

	plot 4	MEF4P4	september	Dark Red Norland

Primer-probe design

Primer-probe sets were adapted from Pandey et al. 2019 for the CP-RT domain of PMTV RNA-CP (401). In efforts to account for the tripartite nature of the PMTV genome, primers were designed for RNA-TGB and RNA-RdRP. The two new primer-probe sets were designed using PMTV isolates newly sequenced along with reference US isolates (Zhai et al., 2020). The sequences were obtained from NCBI GenBank and aligned in MEGAX using MUSCLE. The alignment allowed for visual inspection of conserved regions and identification of regions suitable for primer design. Primers were designed based on conserved regions, meaning 100% similarity amongst isolates. Conserved regions from TGB1 in RNA-TGB and the small replicase protein in RNA-RdRP were targeted. The GenScript real-time PCR (TaqMan) primer design tool was used to design primer/probe sets by importing conserved sequences greater than 100 bps. A set of primer-probes were tested for efficiency. The primer-probe sets were tested on plasmid controls by qPCR using SsoAdvanced™ Universal SYBR® Green Supermix and were annealed at 60°C (Bio-Rad, Herucles, CA). Full-length RNA-RdRP, RNA-CP, and RNA-TGB were previously cloned into pUC19 and pGWB402 and were used in a ten-fold serial dilution, 1 to 1×10^7. Primers were tested in technical triplicates on each dilution. Efficiency was evaluated based on the standard curve created with plasmids as templates and Cq values using new primers sets. We ensured the primer-probe set had reliable and reproducible detection, determined by low Cq values and R^2 value.

Soil RNA extraction

Soil was mixed in the plastic gallon bag to ensure even distribution. For the RNA extraction a total of 300 mg of soil was used. Extractions were performed in three biological replicates, totaling 900 mg of soil for each site. The Norgen BioteK Soil RNA Extraction kit (Norgen BioteK, ON, Canada) was used per manufacturer's instructions and tubes were centrifuged at 20,000 x g, unless specified otherwise. First, the soil was homogenized with 700 µl of Lysis Buffer in the Bead Tube. The Bead Tube was used in Qiagen's TissuelyzerII (Qiagen®, Germantown, MD) for 60 seconds at 30 HZ/second twice, with a 5 minute rest on ice in between. The tubes were then vortexed for 3 minutes for further lysis and centrifuged for 2 minutes. The supernatant was then transferred (450 µl) to an 1.7 ml tube. Solution BX and Binding Buffer E were added sequentially at 50 µl and mixed by inverting. After incubating on ice for five minutes, the lysate was spun for 1 minute to pellet proteins. The pellet was avoided when adding 450 µl of supernatant to the humic acid removal column. The column was spun at 4,000 x g for 30 seconds and the flow through (400 µl) containing RNA was transferred to a new tube. 300 µl of Lysis Buffer QP and 700 µl of 70% ethanol was added to the new tube and vortexed. The clarified lysate was added (650 µl) to the spin column and centrifuged. The flow-though was discarded and it was repeated. Next, on-column DNA removal was done with Norgen's RNase-Free DNase I kit (Norgen BioteK, ON, Canada). 500 µl of binding buffer B was added to the column and centrifuged for 1 minute. After flow through was discarded and 500 µl of Wash solution A was added and centrifuged for 1 minute. The DNase I solution (15 µl) was prepared with 100 µl of enzyme incubation buffer A per sample. The prepared DNase I solution was added to each column and centrifuged for 1 minute. The flowthrough was pipetted back into the column to incubate at

room temperature for 15 minutes. After incubation, 500 µl of wash solution A was added to the column and centrifuged for 2 minutes. RNA was eluted with 50 µl of Elution Solution A. The column was initially centrifuged for 2 minutes at 200 X g, this was followed by a 1 minute spin at 20, 000 x g for final RNA elution. The RNA was nano-dropped using 2 µl of sample, to determine concentration and quality.

cDNA Synthesis

cDNA synthesis was performed using 1 µg of RNA in the GoScript™ Reverse Transcription System (Promega, Fitchburg, WI) as per manufacturer's protocol. Briefly, RNA and 1 µl of Oligo(dT)15 primer (500 µg/ml), and 1 µl of random hexamers (500 µg/ml) were incubated at 70°C for 5 minutes, followed by incubation on ice for 5 min. The additional components required for cDNA synthesis, 4 µl of GoScript™ 5X reaction buffer, 1.2 µl of $MgCl_2$, 1 µl of dNTPs, 20 units of Recombinant RNasin® ribonuclease inhibitor and 1 µl of GoScript™ reverse transcriptase, were added. A final volume of 15 µl volume was achieved by adding Nuclease-free water to the reaction before incubation for cDNA synthesis. cDNA synthesis was carried out using the following thermocycling protocol: initial incubation at 25°C for 5 minutes, reverse transcription at 42°C for 60 minutes, and finally, an inactivation step at 70°C for 15 minutes.

ddPCR

ddPCR was performed on each biological replicate of soil sample separately. The Bio-Rad Supermix for Probes, No dTUP, was used with primers-probes designed above and

PMTV-401 (Bio-Rad, Herucles, CA). A 20 µl reaction was used, containing 4.5 µl of cDNA, 1 µl of each primer (final concentration of 900 nM) and 0.5 µl of the probe (final concentration of 250 nM). The remaining volume was achieved with nuclease-free water. The reaction was initially constructed in PCR strip tubes, before being transferred to a correlating well in a DG8 cartridge which was assembled in the DG8 cartridge holder (Bio-Rad, Herucles, CA). After the samples were added to the wells, 70 µl of QX200 droplet generation oil for probes (Bio-Rad) was added to the oil well for each sample. The cartridge was placed in the Bio-Rad's QX 200 droplet generator and droplets were generated, before being transferred into a 96-well plate. Approximately 40 µl of each sample was added to the plate in the same order. A pierceable foil heat seal was placed on top of the place and sealed using PX1 PCR plate sealer (Bio-Rad). The PCR assay was 96-well thermocycler under the following conditions: enzyme activation for 10 min at 95°C; 40 cycles of 30 sec at 95°C for denaturation, followed by 1 min at 60°C (ramp speed of 2.5°C/sec) for annealing/extension; and enzyme deactivation at 98°C for 10 min, followed by cooling at 4°C. The PCR assay was analyzed with a QX200 droplet reader and QuantaSoft Analysis Pro (Bio-Rad, Hercules, CA). Copy numbers were recorded from QuantaSoft Analysis Pro and used for later calculations. To test lower limits of detection of the primer-probe sets in ddPCR, plasmid controls (pENTR_RdRP, pGWB402_CP and pGWB402_TGB) were used in ten-fold serial dilutions. The use of plasmids allowed us to predict the exact copy number based on the size of plasmid and dilution factor. Efficiency was evaluated based on the expected copy number compared to the number determined by the QX200 droplet reader and QuantaSoft Analysis Pro. Once efficiency was verified, plasmids were used in the

experiment with concentrations ranging from 1 x 10^{-3} to 1 x 10^{4} in three technical replicates. This experiment allowed us to determine the lower detectable limit was 0.22 for RdRP, 0.2 for CP and 0.19 for TGB, approximately four viral copies per reaction.

Correlation Analysis

Correlation analysis was performed in RStudio by looking at the linear relationship between variables. Analysis was performed on field samples by looking at data collected by ddPCR, ELISA, and spore counts. These variables were subjected to the Pearsons correlation coefficient (r). Additionally, trends between spore count and viral load were visualized in RStudio using the 'ggplot2' package by making a dot plot, separating them by timepoint (T1 and T4). The tomato bait assay was also examined for correlation by analyzing copy numbers of the PMTV genome and *S. subterranea,* extracted from qPCR. The variables were used to calculate both Pearsons correlation coefficient and Spearman's rank correlation coefficient (p) due to the log scale.

Tomato bait-assay or Hydroponic assay

A plant bioassay was conducted in biological triplicates using 11 soils from different states and sites. The experimental design also consisted of *S. subterranea*-free controls, three per state, total 12 control jars. The controls were soils from the specific sites that had been autoclaved two times to ensure the removal of *S. subterranea* and other pathogens (Alaryn et al. 2023). The bioassay was performed in a hydroponic-like set up using a nutrient solution (NS) for tomatoes (Merz 1989). The set up consists of autoclaved 4 oz

mason jars, one per tomato. The lids of the jars have a ¼ inch hole drilled in the top for stem insertion, allowing the roots to suspend in 90 ml of NS. The NS consists of a 4-18-38 (NPK) tomato formula from Masterblend® Premium Fertilizers. The solution was made fresh every two weeks by mixing the soluble components with di water as directed on the package. Tomatoes (cv. Alaskan Fancy) were germinated on a petri dish (moist) for five days in the dark. On day five, 3 ml of NS was added to the petri dish. The dishes were moved back into the dark, where the tomatoes grew for 2 more days. After a total of 7 days (once cotyledon is established), the tomatoes were transferred to the hydroponic set-up with fresh NS. The system was kept in a growth chamber at 16°C with 16h of light (Yu et al. 2023). The tomatoes were grown in mason jars with NS for two weeks, before inoculation. To inoculate the tomatoes, the NS solution was swapped out with fresh NS and soil inoculum. The soil was sieved through two kitchen sieves (4 and 2 mm) to remove contaminates (Alaryn et al. 2023). Sieved soil was suspended in 1 ml to 5 ml of water and spores were recounted with a hemocytometer to calculate dilution for uniform inocolum. The resulting soil was used to produce inoculum consisting of approximately 3,000 sporosori suspended in 12 ml of nutrient solution in a 15 ml conical tube. The tube was placed on a rocking vortexter for 1 hour. The tubes were put in the dark for 24 hours and were inverted several times by hand before transferring to the jars with 78 fresh NS. Tomatoes were baited in the chamber for two weeks before tissue collection. Tomato roots were washed with di water and patted dry. A total of 4-6 inches of the root tissue per tomato was collected in a 1.5 ml tube and immediately placed in liquid nitrogen to deter degradation of RNA. Samples were stored at -80°C until further processing.

RNA extraction

Roots were collected from the plant assay after two weeks of inoculation. Roots were washed and patted dry before being placed in a 1.5 ml tube in liquid nitrogen to deter RNA degradation. Using liquid nitrogen, the collected roots were homogenized in the 1.5 ml tube with blue pestles. After homogenization, the samples were weighed to ensure the tissue sample weight was between 100-150 mg. The RNA extraction was performed using the Zymo Direct-zol RNA miniprep plus per the manufacturer's protocol (Zymo, Irvine, California). Ground tissue was combined with 600 µl of TRI Reagent® and thawed on ice. The tubes were centrifuged at 14,000 rpm for 3 minutes, and 600 µl of the supernatant was transferred to a new tube. An equal volume of 95% ethanol was added to the supernatant, and the resulting mixture was transferred to a Zymo-Spin™ IIICG Column. The column was centrifuged at 12,000 rpm for 1 minute, and the resulting flow-through was discarded. Following this step, 400 µl of wash buffer was added to the column, and another round of centrifugation was performed. DNase treatment was carried out by incubating the column for 15 minutes at room temperature. Subsequently, the column was washed twice with 400 µl of pre-wash buffer, followed by a final wash with 700 µl of buffer. The tubes were centrifuged for an additional 2 minutes to ensure the complete removal of residual wash buffer. RNA was eluted from the column matrix using 50 µl of DNase/RNase-Free water. RNA concentration and quality were checked in the lab using nanodrop values and suitable 260/280 ratios. The final 50 µl of RNA was stored at -80°and used for cDNA synthesis.

qPCR and analysis

qPCR was chosen as an alternative to ddPCR on tomato roots due to high virus titers in the sample, although a similar protocol was used. An initial qPCR was done using tomato housekeeping primers (EF1) to ensure RNA quality (Lacerda et al. 2015). PMTV primers described and tested by Pandey et al. 2020, detecting the coat protein (CP) of RNA-CP were used for pathogen testing. The newly designed primer-probe sets for RNA-TGB and RNA-RpRP were also used to detect PMTV. The primer set designed by Nei et al. 2021 was also used to detect the presence of *S. subterranea* by detecting the ITS region of the genome. qPCR reactions were performed separately for each primer with SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, Hercules, CA). The reactions were run with a total reaction volume of 20 µl composed of 10 µl of Supermix, 2 µl of cDNA, and 0.5 µl of each primer (250 nM); the total volume of 20 µl was achieved with nuclease-free water. The thermocycling protocol was as follows: 95°C for 30 sec, followed by 40 cycles of 95°C for 15 sec, annealing at 60°C for 30 sec. Using the qPCR protocol stated, a standard curve for primers was created using plasmids with a full-length viral insert for RNA-RdRP, RNA-CP, RNA-TGB, and the ITS region of *S. subterranea*. Plasmids pUC1-RdRP, pGWB402-CP, pGWB402-TGB, and TOPO-sss-729 were created for this study and used to create a standard curve to calculate the copy number. A set of serial dilutions ranging from 1 to 1 x 10^4 were used in qPCR to create a curve for each primer set. These standard curves produced an equation which was used to calculate copy numbers/g soil for *S. subterranea* and the PMTV genome. Each tomato root sample was qPCRed with each primer set in technical triplicates, and Cq values were recorded. After Cq values were converted to copy

number, results were visualized via a boxplot where an ANOVA was performed in RStudio using the 'ggplot2' and 'tidyverse' packages.

Soil DNA Extraction

DNA was extracted from the 11 soil samples previously mentioned. The DNeasy® PowerSoil® Pro Kit was used as per manufacturer's instructions (Qiagen®, Germantown, MD). A total of 250 mg of soil per sample was used and placed in a PowerBead Pro Tube® along with 800 µl of the CD1 solution. The tube was vortexed and placed in the adaptors for the Qiagen's TissuelyzerII (Qiagen®, Germantown, MD). The samples were shaken at 30 HZ/sec for 5 minutes, then placed on ice for 1 minute, before the process was repeated one time. The tubes were centrifuged at maximum speed for 3 minutes to separate the supernatant from the debris and beads. The supernatant was transferred to a new 2 ml tube and 200 µl of CD 2 solution was added. The tube was vortexed and centrifuged for 1 minute. Debris gathered in the bottom of the tube and the new supernatant (700 µl) was transferred to clean 2 ml tube where 600 µl of CD 3 solution was added. After vortexing, the lysate was loaded into the MB spin column and centrifuged for 1 minute (15,000 g). The flow-through was discarded and the previous step was repeated. The column was placed into a new 2 ml collection tube and 500 µl of solution EA was added. After centrifugation, the flow-through was discarded and 500 µl of C5 solution was added to the spin column. After centrifugation, the column was placed in a clean 2 ml tube and centrifuged for 2 more minutes to ensure the solution passed through. The column was placed in its final 1.5 ml tube and 50 µl solution C6 was added for final elution. After 1 minute of centrifugation, the

column was discarded, and the DNA was ready. DNA quality was checked on a nanodrop and stored at -20 °C.

Microsatellite marker analysis and population analysis

S. subterranea genome structure was examined using microsatellite markers published by Gau et al. 2013. Six loci were targeted by the designed primers and were tagged with different fluorescent markers to create unique PCR products. PCRs were performed using GenScript 2x Taq Master Mix with a total volume of 20 µl (Genscript, Piscataway, NJ). Other components added to the reaction include: 2 µl of cDNA, 20 nM labeled forward primer, 90 nM unlabeled forward primer, 100 nM unlabeled reverse primer, 0.2 mM dNTP, and nuclease-free water. The PCRs followed a normal thermocycling protocol: 95°C for 2:30 min initial denaturation; 35 cycles of 95°C for 40 sec, 55°C for 30 sec, 72°C for 30 sec; and 72°C for 7 min final elongation. Separate PCRs were carried out for each locus (primer set) before they were sent to GENEWIZ® for further analysis (GENEWIZ®, South Planfield, NJ). PCR products were examined by GENEWIZ® and the amplicons were separated on an ABI 3730 xl sequencer. Data processing and allele-size determination was performed using GeneScan LIZ600 standards and the raw data was uploaded to GENEMAPPER® software (Applied Biosystems®). The fluorescent peaks were examined and recorded via GENEMAPPER®. The microsatellite data was subjected to PCoA analysis in RStudio using the 'vegan' package. Additionally, a permutational multivariate analysis of variance (PERMANOVA) was conducted to examine the different *S. subterranea*

communities amongst states. Community dissimilarity was contrasted by location using Bray-Curtis distances and a PERMANOVA using the ‘vegan’ package (Oksanen et al. 2020).

RESULTS

Primer Efficiency and PMTV Detection

The new primers designed in this study for RNA-TGB and RNA-RdRP accurately identified viral copies in soils and tomato bait plants (Table 4.2). We found primer TGB-337 and RdRP-237 efficient in identifying PMTV based on standard curves obtained from qPCRs on viral plasmid DNA. Efficiency ranged from 97.3-100% with $R^2 > 0.98$ for all replicates (Figure 4.1). In ddPCR, the plasmids allowed us to accurately predict the viral copy number and compare it with the actual viral copy number detected: in all cases, the actual viral copy number detected was within 10% of the expected. The newly designed primers demonstrated high efficiency in qPCR and ddPCR, enabling accurate PMTV detection and quantification. New primer-probe sets and the previously described primers PMTV-401 were used to detect and account for the entire tripartite genome of PMTV in soils and tomato roots.

Both qPCR and ddPCR successfully identified PMTV-positive locations, with occasional discrepancies in detecting all three genome segments in soil samples. Results from soil and tomatoes were compared to cross-examine current PMTV detection methods. In Colorado, PMTV was detected in all locations besides CO12 No PMTV detection in this sample was consistent using both PCR methods. In samples CO14 and CO24, no RdRP was detected, while the other two RNAs were. In samples CO42, no TGB was detected, whereas in sample CO44, only TGB was detected. Although all three RNAs may not have been detected in every soil sample, the tomatoes tested positive for all. ddPCR and qPCR revealed PMTV in Oregon soil samples OR738, OR741 OR745 but none in

OR746. CP and TGB were detected in OR738 and OR745. Only TGB was detected in soil sample OR741 at a low level of four copies. Again, all three RNAs were detected using the tomato bait assay in the remaining samples. In North Dakota, CP and TGB were detected in all soil and tomato samples. Maine samples showed the most discrepancy in PMTV-positive samples. ddPCR had varying results for Maine. Some samples only had CP or only TGB. Soil samples ME24 and ME31 had the highest incidence of both RNAs in the soil, but no PMTV was detected in the samples ME31 in tomatoes. Samples ME12, ME13, AND ME44 only had TGB and no CP, while none were positive in soil. The remaining Maine samples tested positive in qPCR and ddPCR at some level. Ultimately, qPCR on the baited tomato consistently detected all three segments, suggesting the tomato bait assay has superior sensitivity for PMTV detection.

Table 4.2 Newly designed primer-probes sand primers used in this study

Primer name	Sequnce(5'-3')	Study
RdRP-237-F	TGTGCATTGTTGTTGTCCGA	This study
RdRP-237-R	CGCTCACCACTTTAGCATCA	This study
RdRP-237-pr	TCTCTGCGCATCACGAAAGTCGAGA	This study
PMTV-401-F	GAAGTTCACTTGTTGGTGATAAA	Pandey et al. 2020
PMTV-401-R	AAGCTTCTCTCGGACCTAATC	Pandey et al. 2020
401-CP-pr	TGTAGCGTTGCAGACAGTGAACGG	Pandey et al. 2020
TGB-337-F	ACTGGTGAGTCCGTAAAGCA	This study
TGB-337-R	AGTTCGCTGACCCTAAACCA	This study
TGB-337-pr	CTCACGCTTGAGCCTGCGGT	This study
728-Sss-F	CTTTGAGTGTCGGTTTCTATTCTCCC	Nie et al. 2021
729-Sss-R	GCACGCCAATGGTTAGAGACG	Nie et al. 2021
EF1-F	GATTGACAGACGTTCTGGTAAGGA	Lacerda et al. 2015
EF1-R	ACCGGCATCACCATTCTTCA	Lacerda et al. 2015

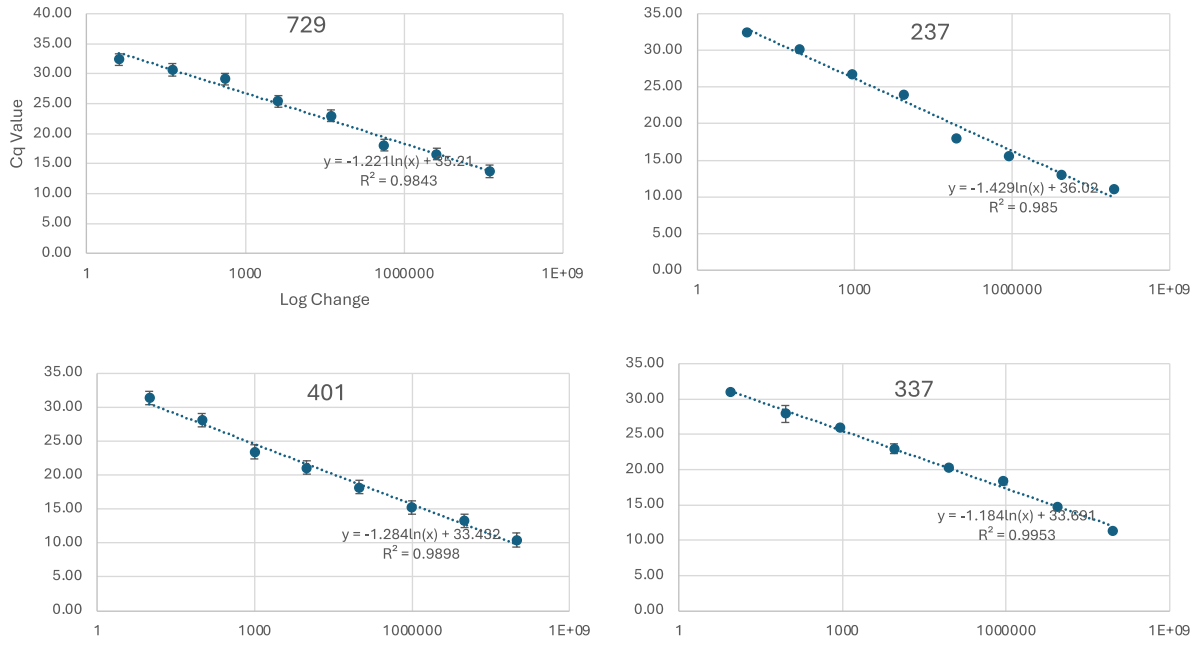

Figure 4.1 Plotted primer efficiencies for all primers used in this study. Primer efficiencies ranges from 98.43% to 99.45%.

Genome Segment Variability and TGB Dominance

ddPCR revealed each state had a different ratio of RNAs: In CO, ME and ND we found TGB present in the highest portion and RdRP in the lowest (Figure 4.2). The varying genome segments potentially complicates soil-based PMTV detection efforts and shows how soil presence may not reflect PMTV infections in the host. To further examine genome segment variability, we compared TGB and CP copy numbers results from ddPCR and qPCR. We found no statistical difference in copies numbers using ddPCR. In qPCR, there was statistical significance when comparing TGB and CP copy numbers in all states besides North Dakota. A one-tailed two sampled t-test, the null hypothesis being TGB is not found in higher ratios when compared to CP. P-values of 0.007, 0.02, and 0.05 rejected that hypothesis for Colorado, Oregon and Maine.

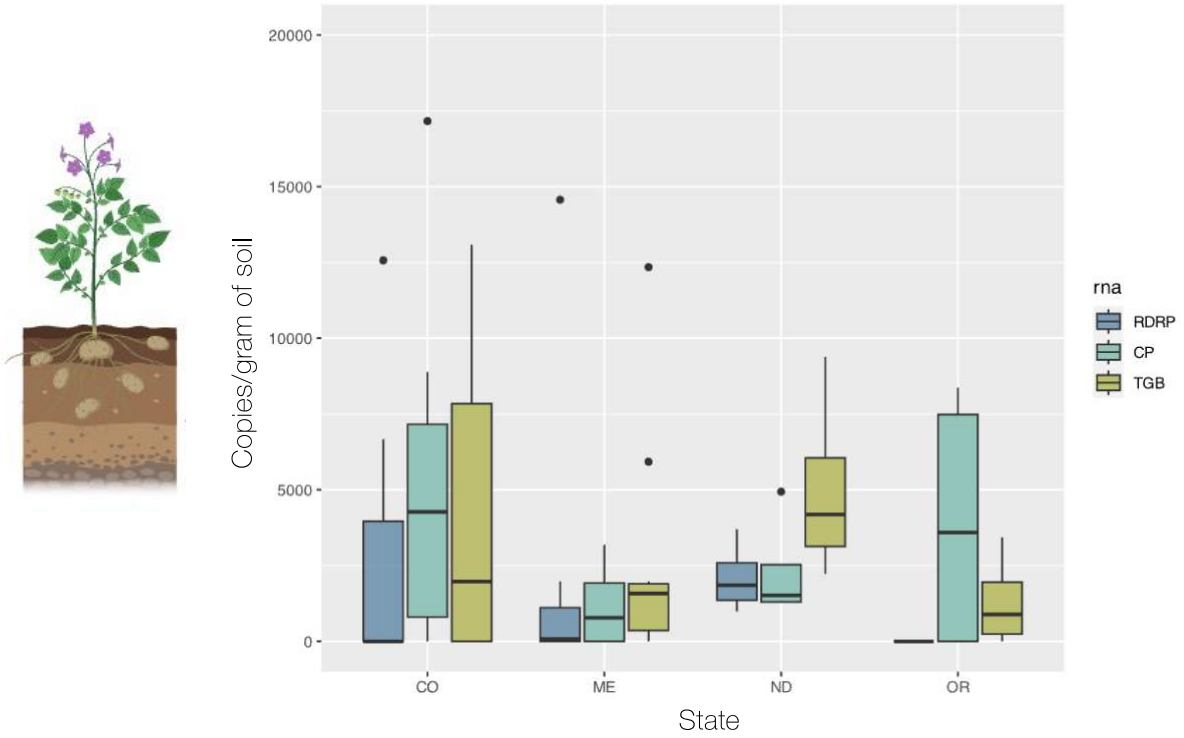

Figure 4.2 ddPCR on soil samples shows variability in RNA segments. RNA was extracted from soil samples and subjected to ddPCR with primers for RNA-RdRP, RNA-CP and RNA-TGB. Copy numbers were translated to grams per soil and grouped by state.

Comparison of Detection Methods and Field Correlations

Overall, PMTV was detected in three ways: tuber symptoms, presence in soils, and baited tomato plants to better understand detection and virus incidence. ELISA results on tubers showed relatively low PMTV incidence compared to the molecular detection methods, qPCR and ddPCR (Figure 4.3). To further investigate the relationship between PMTV RNA levels and other measured variables in this study, correlation analysis was performed, and the relationship was visualized using scatter plots fitted to a linear regression model (Figure 4.4). Sporosori/g of soil and CP copies were used to examine correlation, resulting in a Pearson's coefficient of 0.17 and a Spearman's coefficient of 0.24. Similarly, CP copies/g of soil was plotted on the x-axis with tuber incidence on the y-axis; this resulted in a Pearson's coefficient of 0.046 and Spearman's of 0.005. The relationship between pathogen copy numbers calculated from qPCR data was also examined. Correlation analysis resulted in a Pearson's coefficient of 0.66 and a Spearman's coefficient of 0.79. Spearman's coefficient was chosen as a more accurate coefficient to describe the relationship between *S. Subterranea* copy number and CP copy number due to the log scale used to fit the variables best.

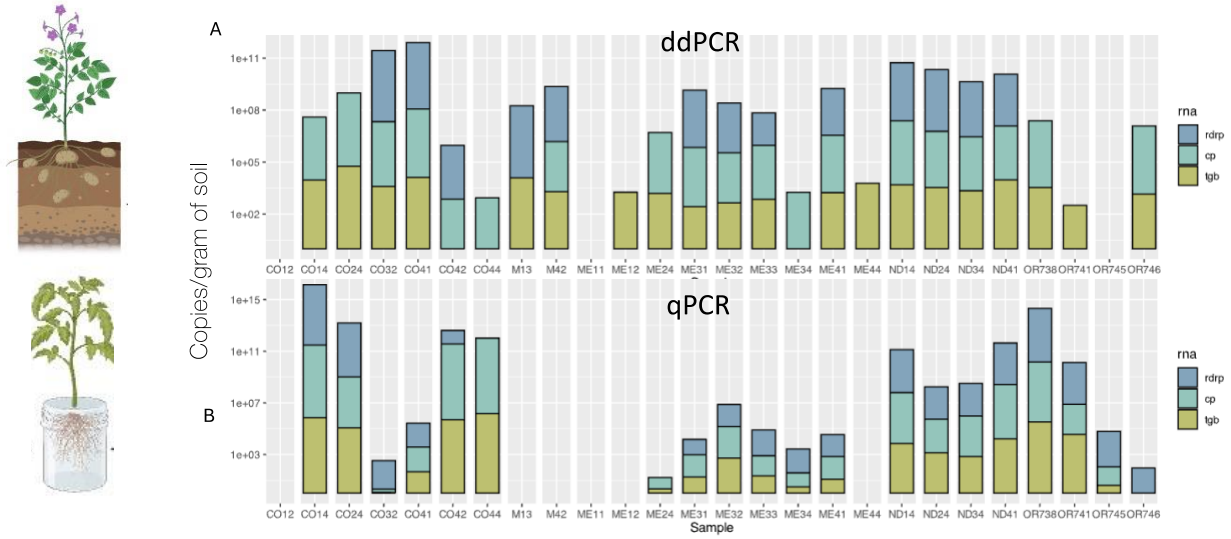

Figure 4.3 Comparison of PMTV detection using A) ddPCR on soils and B) qPCR on baited tomatoes. Both methods were successful in detecting PMTV positive locations. Although there is little discrepancy between the two methods, qPCR more commonly detected all three RNAs compared to ddPCR.

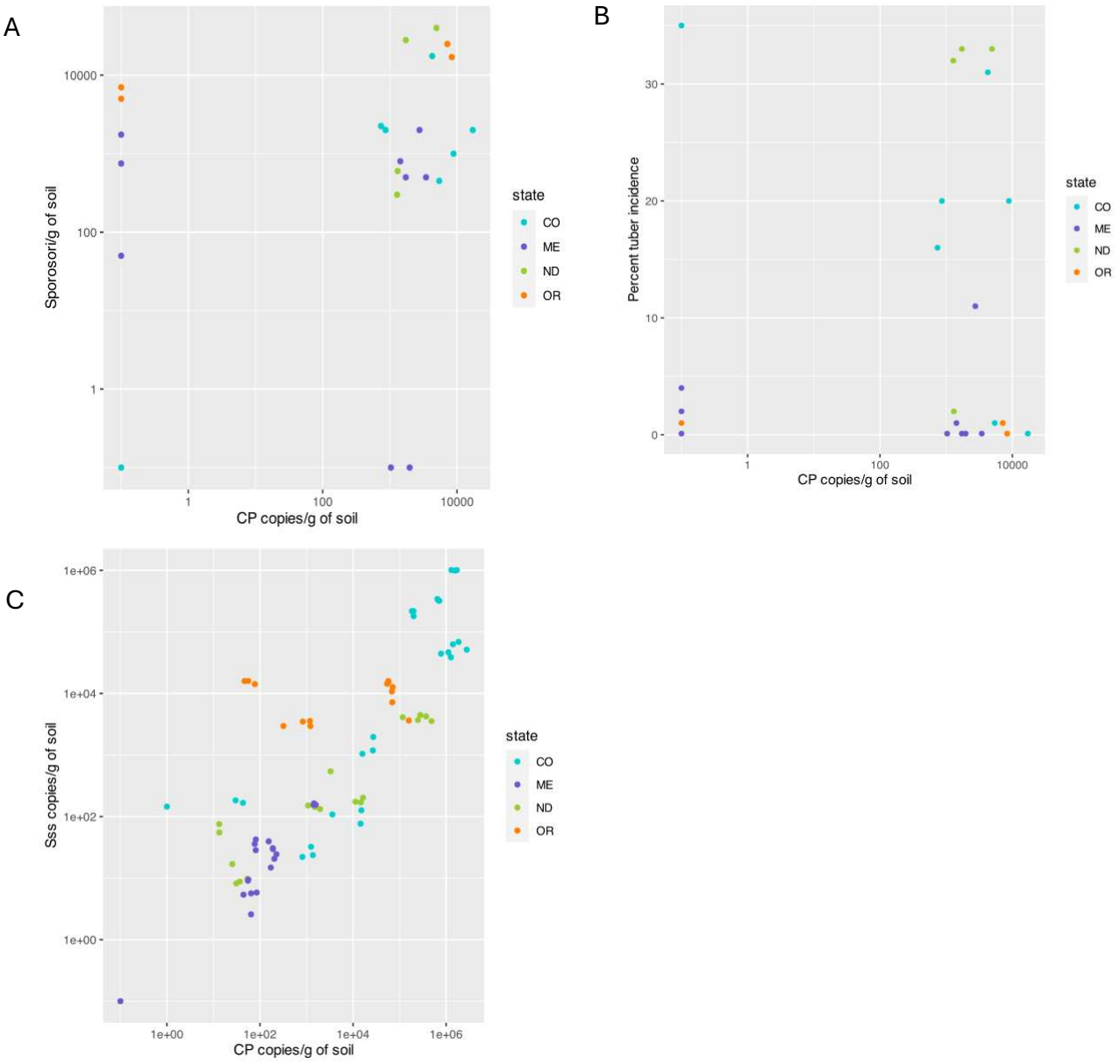

Figure 4.4 Correlation between PMTV and *S. Subterranea*. A) no correlation between PMTV detected in soil from ddPCR and sporosori counted in soil. B) no correlation between PMTV detected in soil and tuber incidence determined from ELISA testing. C) correlation between PMTV and *S. Subterranea* based on qPCR results from baited tomatoes.

State-Specific Differences and S. subterranea Population Structure

The tomato bioassay was successful in amplifying PMTV and *S. subterranea* for detection. Successful baiting was based on Cq values less than 30. For all locations, negative controls resulted in no amplification or very high Cq values (>32) during qPCR analysis. Tomato root RNA was qPCRed with five different primer sets to evaluate controls and assess PMTV presence; three different primer-probe sets were used to assess viral infection of tomato roots, one to amplify the housekeeping gene of tomato, and the other evaluating the ITS region of *S. subterranea*. qPCR was performed on each sample in biological triplicates and technical triplicates, totaling nine data points per sample. All Cq values were recorded and used in standard curves to calculate the pathogen copy number. The samples were then grouped by state and showed that Maine had the lowest copies of CP with an average of 770 copies/g of soil, while Colorado had the highest with an average of 27,300 copies/g of soil. An ANOVA test was performed to analyze the difference amongst the states and found a p-value of 1.56e-5, meaning CP copy number was statistically different based on state (Figure 4.5). Further statistical analysis revealed that Colorado and Maine had the greatest difference with a p-value of 0.000147. Additionally, North Dakota and Oregon were the most similar with a significant p-value of 0.02 (Tukey test).

The population structure of *S. subterranea* was investigated using microsatellite markers targeting alleles at six loci. After PCR and amplicon separation, the fluorescent peaks recorded were used to create a distance matrix graphically represented by PCoA (Figure 4.6). Bray-Curtis dissimilarity statistics were used to cluster samples. Variation amongst *S. Subterranea* communities was described along the two principal axes which

explained 60.1% and 14.7% of the variation (Figure 4.6). The PERMANOVA highlighted differences in the *S. subterranea* communities that were associated with the state with a significant P-value of 0.001 (R2 = 0.42487 F= 3.8784). Visually, the PCoA plot shows a clear distinction between populations CO, ME, ND, and OR. The samples from Oregon clustered tightly, meaning they are the most dissimilar to other populations, whereas Colorado and North Dakota are the most similar; they have some small overlap and are oriented closest together. Overall, this is the first insight into the population structure of *S. subterranea* in the United States.

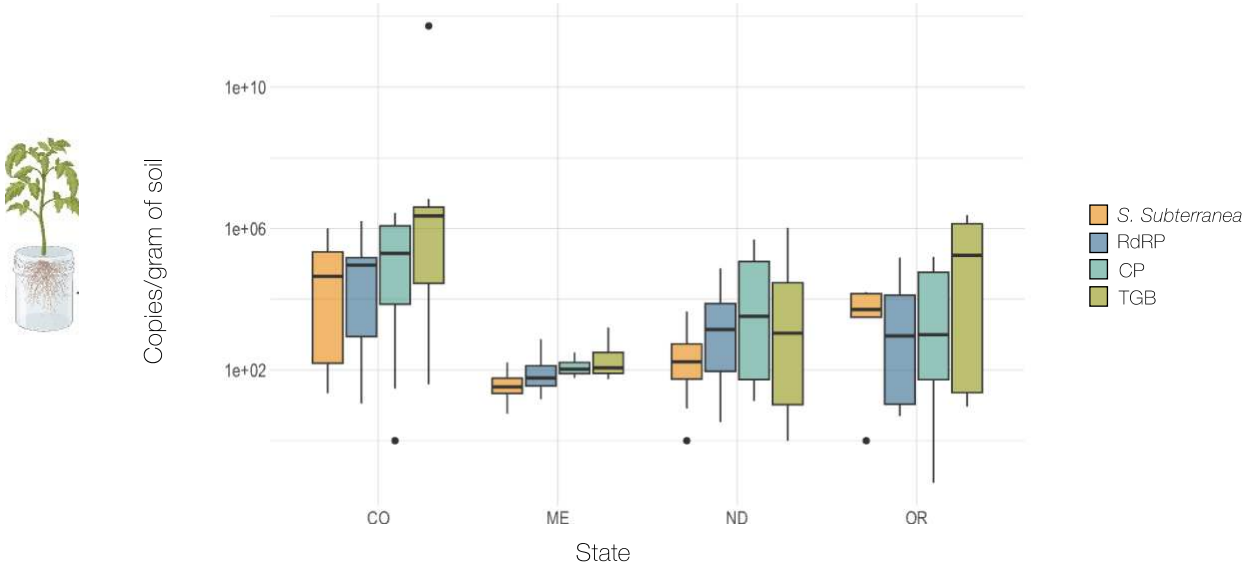

Figure 4.5 qPCR data with variation of viral copy numbers based on location. RNA extracted from tomato roots inoculated with *S. subterranea* infested soils was subjected to qPCR. PMTV's genome segments were detected using three different primer sets. *S. Subterranea* was also detected by targeting the ITS region. Maine was statistically different from the other states with low copy numbers. In Colorado, Maine, and Oregon, RNA-TGB was found at greater levels.

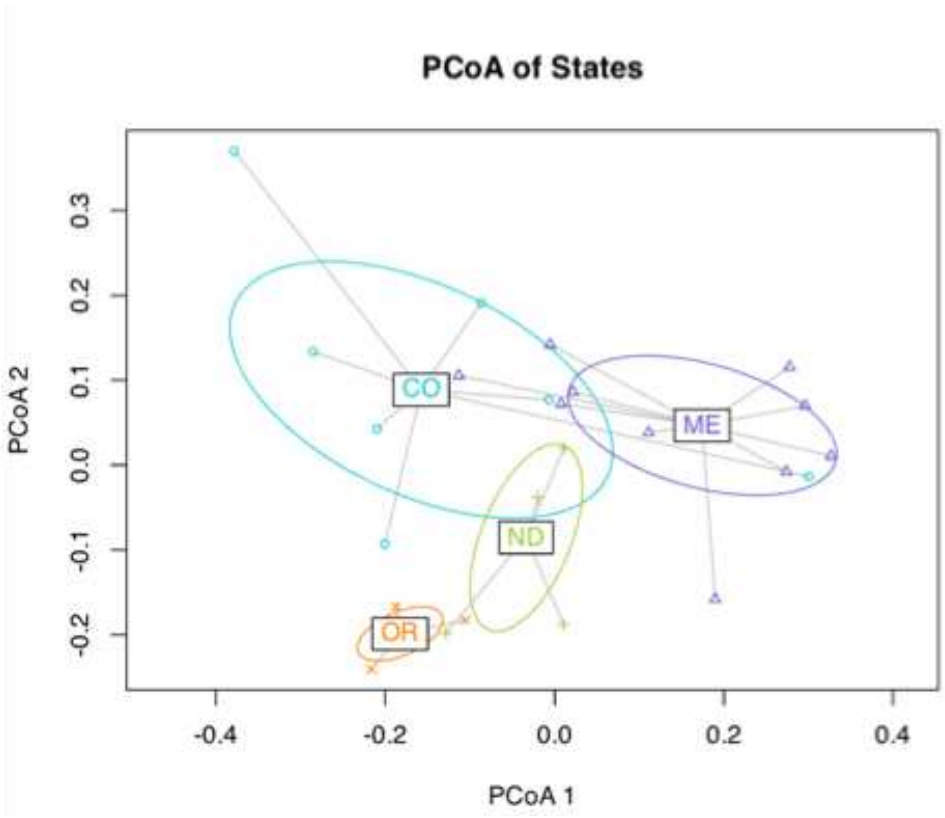

Figure 4.6 PCoA generated using Bray-Curtis dissimilarity showing clusters by state. Microsatellite markers were used to analyze population genetics of *S. Subterranea* infested soils. Each point represents a location within the state with some overlap between North Dakota and Colorado. Variation amongst *S. Subterranea* communities was described along the two principal axes which explained 60.1% and 14.7% of the variation. The PERMANOVA highlighted differences in the *S. subterranea* communities that were associated with state with a significant P-value of 0.001 (R2 = 0.42487 F= 3.8784).

DISCUSSION

PMTV and *S. subterranea* are challenging pathogens to manage due to their soil-borne obligate nature and complex relationship. The qualitative losses created by PMTV and *S. subterranea* pose significant challenges to potato production, making pathogen detection vital to management strategies. Our study used advanced PCR techniques to explore *S. subterranea*-infested soils with newly designed primers from available PMTV genomes. Genome segments varied in proportions found in the soil, potentially complicating soil-based PMTV detection. We used efficient primers to cross examine current detection methods by detecting PMTV in soil and baited tomatoes. Both methods were able to detect PMTV-positive locations, but the tomato bait assay consistently detected all three RNAs. RNA-TGB was often found at higher levels, suggesting it could be valuable to test for TGB in addition to CP when looking at soil and baited tomatoes. Given PMTV's ability to move without RNA-CP, it is interesting that RNA-TGB was found in higher concentrations. Correlation between the virus-vector system was evaluated in both field and greenhouse experiments, and correlation was only found in baited tomatoes. Data generated from the tomato bait assay was also used to analyze viral infections with respect to the *S. subterranea-infested* soils. Viral copy counts varied based on location (ANOVA), with Maine having the lowest counts. We used microsatellite markers to better understand potential factors influencing the soil inoculum. We found *S. subterranea* to have distinct populations based on state, which was the same trend seen in viral copy counts, suggesting the genotype of S. subterranea could have influenced PMTV infections.

Primers and probes were created to account for PMTV's tripartite genome, looking at RNA-TGB and RNA-RdRP and traditionally detected RNA-CP. The new primers were tested using qPCR and ddPCR on inoculated tomatoes and *S. subterranea*-infested *soils*. Recently developed technologies have advanced the specificity, sensitivity, and rapidness of PCR methods and have given rise to new methods such as droplet digital PCR (ddPCR) and high-resolution DNA melting (HRM) analyses. These methods have been used to detect PMTV and *S. subterranea* in soil, showing the feasibility of advancements in detection methods. HRM was used to develop a multiplex assay which can simultaneously detect both virus and vector (Nie et al., 2021; Pandey et al., 2020). ddPCR has been used successfully to detect low levels of virus in soils and overcome soil inhibitors (Dingle et al., 2013). Before the development of molecular assays, a plant baiting assay was the method required for detection (Arif et al., 2014; Davey, 2009; Nakayama et al., 2010). Using our newly designed primers, we found that the ddPCR and qPCR results agreed on PMTV-positive fields. In ddPCR, little to no RNA-RdRP was often detected, but we found RNA-CP and RNA-TGB in substantial copy numbers. The limitations of ddPCR are evident, as highlighted by Pandey et al. 2020, their findings suggest that fewer than 10 copies may go undetected in spiked soil samples. Our study found that detecting fewer than four copies of pure plasmid DNA was unattainable. In field samples, the ddPCR assay from Pandey et al could not detect PMTV when present at levels below 100 copies, which is a greater threshold than that observed in spiked soils. This is most likely a result of soil inhibitors and uneven distribution of the PMTV genome segments in samples. Both factors should be considered when detecting PMTV in soils.

PMTV genome composition in roots and soils varied across locations, emphasizing the importance of detecting the virus with multiple primers, specifically using primers for both RNA-CP and RNA-TGB. We found TGB to have a higher copy number than CP and RdRP in baited plants. Genome variability and the dominance of TGB are new findings that challenge previous detection methods and primers. Primers developed by Mumford et al. in 2000 have been widely used to detect PMTV's coat protein, but studies in New Zealand and the United States have suggested they are less sensitive, possibly due to being incompatible with particular PMTV isolates found in these regions (Frampton et al., 2022). As advancements have been made, our study attempted to tackle the tripartite genome of PMTV. Tomato bait assay results showed that TGB was detected at significantly higher rates in three of the four states than in CP and RdRP. In most cases, the traditional CP was also detected at high levels. In a bait assay where the overall viral load is higher than in soil, looking at RdRP and TGB may not be necessary, but in cases where there is low virus titer, it may be essential to look at TGB as well. This could especially be true when looking at tubers due to PMTVs' ability to move as a ribonucleic entity. In other multipartite viruses, specifically, beet necrotic yellow vein virus (BNYVV) vectored by a plasmodiophorid, *Plasmodiophora brassicae*, a study found the relative frequencies of the multipartite genome to be different depending on host and tissues (Dall et al., 2024; Kalantidis, 2000). Genetic formula, or the accumulation of RNA segment accumulation in a host, has been shown to serve an evolutionary purpose and could be important in detecting PMTV (Boezen et al., 2023; Iranzo & Manrubia, 2012).

The relationship between PMTV and *S. subterranea* was investigated using ddPCR data from field soils and qPCR data from the tomato bait assay. Soil data was paired with field tuber ELISA testing and soil sporosori counts to correlate the pathogens with each other. We hypothesized that PMTV and *S. subterranea* in soils would be correlated due to their virus-vector nature. Alternatively, we found no correlation between viral PMTV copies in soils and sporosori counts in soils. Furthermore, we found no correlation between PMTV tuber incidence and pathogen counts in soils. Several studies have explored the relationship between the incidence of pathogens using different methods and found varying results. A study from Costa Rica found a very low correlation between virus and vector using ELISAs for both pathogens (Montero-Astúa et al., 2008). A previous study by Domfeh et al. 2015, found a correlation between powdery scab incidence in tubers and PMTV-induced tuber necrosis. It is also important to note that the Pearson's coefficient found from the study was only 0.62 (Domfeh et al., 2015). The methods used in previous studies examined the pathogens using techniques that require greater pathogen loads, which could explain why they saw a correlation in the field compared to our results. Using qPCR data, we found a correlation between *S. subterranea* copy number and CP copy number in tomato-baited plants. The discrepancy between the correlation in the field versus the tomato bait assay can be explained by several factors. The tomato bait assay occurred in a controlled hydroponic environment with a favorable constant temperature, a susceptible host, and uniform sporosori in the inoculum. With an active host, we can assume the spores were replicating as well as the virus. These ideal conditions allowed for the hypothesized relationship between virus and vector to be seen. The difference in

correlation to field analysis can be explained by environmental factors in the field, which have been studied and confirmed to have effects on *S. subterranea* incidence (Nakayama et al., Van De Graaf et al., 2005, 2007; Wright et al., 2021). It has been found that high moisture and lower temperature are the most critical factors affecting *S. subterranea* infections (Domfeh & Gudmestad, 2016; Van De Graaf et al., 2005). A recent study noted that soil texture and drainage influence powdery scab incidence and severity (Wright et al., 2021). In a previous study, Nakayama et al. (2007) also reported an insignificant relationship between soil sporosori levels and powdery scab disease severity (Nakayama et al., 2007; Van De Graaf et al., 2007). In addition to environmental factors, studies show that inoculum levels have little to do with *S. subterranea* incidence. For example, Shat et al. (2011) found that even low inoculum levels can cause severe symptoms (Shah et al., 2011). Our study further supports the claim that environmental factors influence *S. subterranea* and that inoculum levels do not affect disease severity.

We looked deeper into qPCR copy numbers to further explore factors influencing pathogenicity. Statistically, we found that Colorado, North Dakota, and Oregon were most similar to each other, whereas Maine is the most different. Additionally, we characterized *S. subterranea* communities using microsatellite markers and showed distinct populations based on state. Microsatellite markers have been instrumental in investigating population communities of *S. subterranea*. In our study, we applied a PERMANOVA analysis to explore the association between *S. subterranea* communities and geographical states. We observed a significant relationship (P-value = 0.001) between community composition and state. Notably, Colorado and North Dakota exhibited high similarity, with North Dakota also

sharing substantial overlap with Oregon. However, Maine displayed a distinct pattern, with data points on the right side of the axis being farthest from other sites. Another study using microsatellite markers, Gau et al. 2013, found distinct populations globally, specifically finding differences in Colombian/Venezuelan vs other populations worldwide including Europe and the United States (Gau et al. 2013). In a large-scale study, they found two distinct populations: Colombia and Peru, while the other population includes regions such as Europe, the United States, and the Oceanians. This finding is consistent with phylogenetic studies of *S. subterranea* (Merz, 2008). Additionally, a study done by Muzhinji and van der Waals explored populations in South Africa in 16 growing regions across the country (Muzhinji & van der Waals, 2019). They found no distinct groupings in communities within South Africa but found they clustered differently when compared to Colombia. Therefore, finding a significant group of S. subterranea populations within just four states across the United States was interesting. Our study contributes novel insights by revealing significant population groupings within the United States, specifically for *S. subterranea*. Notably, this is the first study to suggest distinct *S. subterranea* communities within the U.S. However, further analysis to explore the differences in communities is hindered because *S. subterranea's* genome is not easily sequenced.

Conclusion

Our study paired a ddPCR assay and a hydroponic tomato bait assay to analyze detection methods and PMTV's tripartite genome. We successfully developed primers for the entire PMTV genome, accounting for the unusual movement of the virus and variability

of its segmented genome in soil and host (Boezen et al., 2023; Torrance et al., 2009). We found a correlation between PMTV and *S. subterranea* infections in a controlled environment, but not in the field. Colorado, North Dakota/Minnesota, and Oregon were the most similar when looking at pathogen copy numbers and *S. subterranea* communities. Maine was the most different from the other locations in these variables, suggesting ecological factors could influence the evolution of *S. subterranea*. Maine is the furthest distance from the other states, supporting the idea that environmental factors could cause the most significant difference in its *S. subterranea* populations. There are still many unknowns about the relationship between PMTV and *S. subterranea* and their pathogenicity factors. Understanding the relationship between this virus-vector system and its specific ecological niche can aid in management and shows that management strategies may differ based on location.

REFERENCES

Adams, M. J., Antoniw, J. F., & Mullins, J. G. L. (2001). Plant virus transmission by plasmodiophorid fungi is associated with distinctive transmembrane regions of virus-encoded proteins. Archives of Virology, 146(6), 1139–1153. https://doi.org/10.1007/S007050170111/METRICS

Arif, M., Ali, M., Rehman, A., & Fahim, M. (2014). Detection of potato mop-top virus in soils and potato tubers using bait-plant bioassay, ELISA and RT-PCR. Journal of Virological Methods, 195, 221–227. https://doi.org/10.1016/J.JVIROMET.2013.10.022

Boezen, D., Vermeulen, M., Johnson, M. L., van der Vlugt, R. A. A., Malmstrom, C. M., & Zwart, M. P. (2023). Mixed viral infection constrains the genome formula of multipartite cucumber mosaic virus. Frontiers in Virology, 3, 1225818. https://doi.org/10.3389/FVIRO.2023.1225818

Calvert, E. L., & Harrison, B. D. (1966). POTATO MOP-TOP, A SOIL-BORNE VIRUS. Plant Pathology, 15(3), 134–139. https://doi.org/10.1111/J.1365-3059.1966.TB00333.X

Ciaghi, S., Neuhauser, S., & Schwelm, A. (2018). Draft genome resource for the potato powdery scab pathogen spongospora subterranea. Molecular Plant-Microbe Interactions, 31(12), 1227–1229. https://doi.org/10.1094/MPMI-06-18-0163-A/ASSET/IMAGES/LARGE/MPMI-06-18-0163-A_T1.JPEG

Dall, M., Guo, Y., Poli, D., Gilmer, D., & Ratti, C. (2024). Analysis of the relative frequencies of the multipartite BNYVV genomic RNAs in different plants and tissues. Journal of General Virology, 105(1), 001950. https://doi.org/10.1099/JGV.0.001950/CITE/REFWORKS

Davey, T. (2009). An accelerated soil bait assay for the detection of potato mop top virus in agricultural soil. Methods in Molecular Biology (Clifton, N.J.), 508, 259–265. https://doi.org/10.1007/978-1-59745-062-1_20/TABLES/1

Dingle, T. C., Sedlak, R. H., Cook, L., & Jerome, K. R. (2013). Tolerance of Droplet-Digital PCR vs Real-Time Quantitative PCR to Inhibitory Substances. Clinical Chemistry, 59(11), 1670–1672. https://doi.org/10.1373/CLINCHEM.2013.211045

Domfeh, O., Bittara, F. G., & Gudmestad, N. C. (2015). Sensitivity of Potato Cultivars to Potato mop-top virus-Induced Tuber Necrosis. Plant Disease, 99(6), 788–796. https://doi.org/10.1094/PDIS-07-14-0705-RE

Domfeh, O., & Gudmestad, N. C. (2016). Effect of soil moisture management on the development of Potato mop-top virus-induced tuber necrosis. Plant Disease,

100(2), 418–423. https://doi.org/10.1094/PDIS-05-15-0590-RE/ASSET/IMAGES/LARGE/PDIS-05-15-0590-RE_T3.JPEG

Ezekiel, R., Singh, N., Sharma, S., & Kaur, A. (2013). Beneficial phytochemicals in potato — a review. Food Research International, 50(2), 487–496. https://doi.org/10.1016/J.FOODRES.2011.04.025

Frampton, R. A., Addison, S. M., Kalamorz, F., & Smith, G. R. (2022). Genomes of Potato Mop-Top Virus (Virgaviridae: Pomovirus) Isolates from New Zealand and Their Impact on Diagnostic Methods. Plant Disease, 106(10), 2571–2575. https://doi.org/10.1094/PDIS-01-22-0192-SC

Gau, R. D., Merz, U., Falloon, R. E., & Brunner, P. C. (2013). Global Genetics and Invasion History of the Potato Powdery Scab Pathogen, Spongospora subterranea f.sp. subterranea. PLoS ONE, 8(6). https://doi.org/10.1371/journal.pone.0067944

Hogenhout, S. A., Ammar, E. D., Whitfield, A. E., & Redinbaugh, M. G. (2008). Insect vector interactions with persistently transmitted viruses. In Annual Review of Phytopathology (Vol. 46, pp. 327–359). https://doi.org/10.1146/annurev.phyto.022508.092135

Iranzo, J., & Manrubia, S. C. (2012). Evolutionary dynamics of genome segmentation in multipartite viruses. Proceedings of the Royal Society B: Biological Sciences, 279(1743), 3812. https://doi.org/10.1098/RSPB.2012.1086

Kalantidis, K. (2000). Generation of 13k-gene sugar beet transformants and evaluation of their resistance to BNYVV infection. Developments in Plant Genetics and Breeding, 6(C), 189–194. https://doi.org/10.1016/S0168-7972(00)80121-8

Kalyandurg, P., Gil, J. F., Lukhovitskaya, N. I., Flores, B., Müller, G., Chuquillanqui, C., Palomino, L., Monjane, A., Barker, I., Kreuze, J., & Savenkov, E. I. (2017). Molecular and pathobiological characterization of 61 Potato mop-top virus full-length cDNAs reveals great variability of the virus in the centre of potato domestication, novel genotypes and evidence for recombination. Molecular Plant Pathology, 18(6), 864–877. https://doi.org/10.1111/mpp.12552

Kirk, H. G. (2008). Mop-top virus, relationship to its vector. American Journal of Potato Research, 85(4), 261–265. https://doi.org/10.1007/s12230-008-9021-7

Lambert, D. H., Levy, L., Mavrodieva, V. A., Johnson, S. B., Babcock, M. J., & Vayda, M. E. (2007). First Report of Potato mop-top virus on Potato from the United States. Https://Doi.Org/10.1094/PDIS.2003.87.7.872A, 87(7), 872–872. https://doi.org/10.1094/PDIS.2003.87.7.872A

Merz, U. (2008). Powdery scab of potato - Occurrence, life cycle and epidemiology. American Journal of Potato Research, 85(4), 241–246. https://doi.org/10.1007/S12230-008-9019-1/FIGURES/2

Montero-Astúa, M., Vasquéz, V., Turechek, W. W., Merz, U., & Rivera, C. (2008). Incidence, Distribution, and Association of Spongospora subterranea and Potato mop-top virus in Costa Rica. Https://Doi.Org/10.1094/PDIS-92-8-1171, 92(8), 1171–1176. https://doi.org/10.1094/PDIS-92-8-1171

Muzhinji, N., & van der Waals, J. E. (2019). Population Biology and Genetic Variation of Spongospora subterranea f. Sp. Subterranea, the Causal Pathogen of Powdery Scab and Root Galls on Potatoes in South Africa. Phytopathology, 109(11), 1957–1965. https://doi.org/10.1094/PHYTO-12-18-0467-R

Nakayama, T., Maoka, T., Hataya, T., Shimizu, M., Fuwa, H., Tsuda, S., & Mori, M. (2010). Diagnosis of Potato Mop-Top Virus in Soil Using Bait Plant Bioassay and RT-PCR-microplate Hybridization. American Journal of Potato Research, 87(2), 218–225. https://doi.org/10.1007/S12230-010-9128-5/FIGURES/6

Nakayama, T., Mitsuo, A. E., Ae, H., & Shimanuki, T. (n.d.). Spongospora subterranea soil contamination and its relationship to severity of powdery scab on potatoes. https://doi.org/10.1007/s10327-007-0008-x

Nie, X., Singh, M., Chen, D., Gilchrist, C., Soqrat, Y., Shukla, M., Creelman, A., Dickison, V., Nie, B., Lavoie, J., & Bisht, V. (2021). Development of high-resolution DNA melting analysis for simultaneous detection of potato mop-top virus and its vector, spongospora subterranea, in soil. Plant Disease, 105(4), 948–957. https://doi.org/10.1094/PDIS-06-20-1321-RE

Pandey, B., Mallik, I., & Gudmestad, N. C. (2020). Development and Application of a Real-Time Reverse-Transcription PCR and Droplet Digital PCR Assays for the Direct Detection of Potato mop top virus in Soil. Phytopathology, 110(1), 58–67. https://doi.org/10.1094/PHYTO-05-19-0185-FI/ASSET/IMAGES/LARGE/PHYTO-05-19-0185-FI_T3.JPEG

Ramesh, S. V., Raikhy, G., Brown, C. R., Whitworth, J. L., & Pappu, H. R. (2014). Complete genomic characterization of a potato mop-top virus isolate from the United States. Archives of Virology, 159(12), 3427–3433. https://doi.org/10.1007/S00705-014-2214-0/FIGURES/4

Santala, J., Samuilova, O., Hannukkala, A., Latvala, S., Kortemaa, H., Beuch, U., Kvarnheden, A., Persson, P., Topp, K., Ørstad, K., Spetz, C., Nielsen, S. L., Kirk, H. G., Budziszewska, M., Wieczorek, P., Obrepalska-Steplowska, A., Pospieszny, H., Kryszczuk, A., Sztangret-Wiśniewska, J., … Valkonen, J. P. T. (2010). Detection,

distribution and control of Potato mop-top virus, a soil-borne virus, in northern Europe. Annals of Applied Biology, 157(2), 163–178. https://doi.org/10.1111/j.1744-7348.2010.00423.x

Sarwar, M., Aslam, M., Sarwar, S., & Iftikhar, R. (2020). Different nematodes and plasmodiophorids as vectors of plant viruses. In Applied Plant Virology: Advances, Detection, and Antiviral Strategies (pp. 275–290). Elsevier. https://doi.org/10.1016/B978-0-12-818654-1.00021-9

Savenkov, E. I. (2020). Pomoviruses (Virgaviridae). Encyclopedia of Virology: Volume 1-5, Fourth Edition, 1–5, 603–611. https://doi.org/10.1016/B978-0-12-809633-8.21304-6

Savenkov, E. I., Sandgren, M., & Valkonen, J. P. T. (1999). Printed in Great Britain Complete sequence of RNA 1 and the presence of tRNA-like structures in all RNAs of Potato mop-top virus, genus Pomovirus. In Journal of General Virology (Vol. 80).

Scott, K. P., Kashiwazaki, S., Reavy, B., & Harrison, B. D. (1994). The nucleotide sequence of potato mop-top virus RNA 2: a novel type of genome organization for a furovirus. In Journal of General Virology (Vol. 3561).

Shah, F. A., Falloon, R. E., Butler, R. C., & Lister, R. A. (n.d.). Low amounts of Spongospora subterranea sporosorus inoculum cause severe powdery scab, root galling and reduced water use in potato (Solanum tuberosum). https://doi.org/10.1007/s13313-011-0110-6

Torrance, L., Lukhovitskaya, N. I., Schepetilnikov, M. V, Cowan, G. H., Ziegler, A., & Savenkov, E. I. (2009). Unusual Long-Distance Movement Strategies of Potato mop-top virus RNAs in Nicotiana benthamiana. / 381 MPMI, 22(4), 381–390. https://doi.org/10.1094/MPMI

Van De Graaf, P., Lees, A. K., Wale, S. J., & Duncan, J. M. (2005). Effect of soil inoculum level and environmental factors on potato powdery scab caused by Spongospora subterranea. Plant Pathology, 54(1), 22–28. https://doi.org/10.1111/J.1365-3059.2005.01111.X

Van De Graaf, P., Wale, S. J., & Lees, A. K. (2007). Factors affecting the incidence and severity of Spongospora subterranea infection and galling in potato roots. Plant Pathology, 56(6), 1005–1013. https://doi.org/10.1111/J.1365-3059.2007.01686.X

Whitworth, J. L., & Crosslin, J. M. (2012). Detection of Potato mop top virus (Furovirus) on potato in southeast Idaho. Https://Doi.Org/10.1094/PDIS-08-12-0707-PDN, 97(1), 149. https://doi.org/10.1094/PDIS-08-12-0707-PDN

Wright, P. J., Falloon, · R E, Anderson, · C, Frampton, · R A, Curtin, · D, & Hedderley, · D. (n.d.). Factors influencing suppressiveness of soils to powdery scab of potato. 1, 3. https://doi.org/10.1007/s13313-021-00822-z

Zhai, Y., Mallik, I., Hamid, A., Tabassum, A., Gudmestad, N., Gray, S. M., & Pappu, H. R. (2020). Genetic diversity in potato mop-top virus populations in the United States and a global analysis of the PMTV genome. European Journal of Plant Pathology, 156(2), 333–342. https://doi.org/10.1007/s10658-019-01836-6

CHAPTER 5 – CLONING METHODS

INTRODUCTION

Tuber necrotic viruses pose a significant threat to the potato industry. One of the most notable viruses accounting for seed potato rejection is the potyvirus, potato virus Y (PVY) (Nolte et al., 2004). The aphid-vectored virus is challenging to manage due to the abundance of aphid species able to vector it, the vegetative propagation of potatoes, and PVY's recombination abilities (Green et al., 2018; Karasev & Gray, 2013; Pelletier et al., 2012). Characteristic PVY symptoms include yellowing leaves and potato tuber necrotic ringspot disease (PTNRD) (Baldauf et al., 2007; Crosslin et al., 2002). Preventative management strategies are critical for controlling viral infections in crops. In potatoes, this involves breeding programs focused on developing cultivars with resistance to viral pathogens, alongside the implementation of seed certification programs to ensure the use of virus-free planting materials. Seed certification programs are used nationwide to inspect seed potatoes and assess disease prevalence using protocols to limit disease incidence below certain yield-damaging thresholds(Frost et al., 2013). Although seed certification programs widely use visual inspections, PVY can often be symptomless and go undetected. To accurately detect PVY, molecular methods such as ELISA or PCR may provide a definitive diagnosis and strain type (Chikh-Ali & Karasev, 2023; Karasev et al., 2010; Lorenzen et al., 2006).

Host resistance is another technique used to manage PVY infections, and it is integrated into breeding programs. There are three known resistance genes to PVY that provide extreme resistance in potato germplasm: Ry_{sto}, Ry_{adg}, and Ry_{chc} (Brigneti et al., 1997;

Whitworth et al., 2009). Although these genes are known, there are no highly used commercial potato cultivars with these resistance genes due to other traits that are paired with extreme resistance from the *Ry* genes (Elison et al., 2021). Current efforts are in place to clone resistance genes, reducing the linkage-drag problem. However, the tetraploid nature of potatoes still presents a complex challenge when attempting to integrate these genes. A recent movement attempts to transition potatoes to a diploid organism, similar to the tomato industry, to combat these complexities but efforts will take several years (Jansky et al., 2016).

The soil-borne plasmodiophorid vector virus, potato mop-top virus (PMTV), is another tuber necrotic virus affecting the potato industry (Davey et al., 2014; Jones & Harrison, 1969). PMTV was first identified in the United States in 2011, representing a relatively recent addition to potato viruses important to seed certification (Lambert et al., 2007; Whitworth & Crosslin, 2012). Although PMTV has plagued other parts of the world for decades, significant knowledge gaps remain in its management and detection strategies. The life cycle and thick-walled, persistent spores of *S. subterranea* create an additional challenge in managing this pathosystem (Merz, 2008). The virus vector system causes qualitative losses that are detrimental to the potato. *Spongospora subterranea* f. sp. *subterranea* (*S. subterranea*), the causal agent of powdery scab disease, causes scabs on the potato tuber surface along with galls on the roots (Jones & Harrison, 1969). PMTV induces necrotic arcs in tubers and can also create a chevron yellowing pattern on potato leaves. Similar to PVY, there are no commercial breeding lines resistant to PMTV or its vector. Furthermore, there are no known resistant genes to these pathogens, creating a

need for resistance detection in current and new breeding lines (Domfeh et al., 2015; Yellareddygari et al., 2018).

With no current resistance to PMTV and a lack of resistant potato cultivars to PVY, innovative biotechnology tools offer possible solutions (Domfeh et al., 2015). Full-length infectious cDNA clones are a valuable tool for studying virus-host interactions (Bordat et al., 2015; Sun et al., 2017). Specifically, potyviruses are challenging to clone due to their toxicity to *E.Coli*. For example, PVY's 9kb genome contains an intron that is toxic to *E.coli*, often resulting in unstable clones (Desbiez et al., 2012).Although cloning a viral genome can be challenging and time-consuming, there have been dozens of full-length infectious clones constructed (Delbianco et al., 2013; Laufer et al., 2018; López-Moya & García, 2000; Tuo et al., 2015).

Cloning methods include traditional cloning using restriction enzymes and ligation, Golden Gate Assembly, In-Fusion cloning, Gateway cloning, and Homologous recombination in yeast (Bordat et al., 2015; Desbiez et al., 2012; Sun et al., 2017; Tuo et al., 2015). The different methods offer various technologies meant to make cloning faster and easier. Specifically, In-Fusion cloning uses DNA polymerase derived from Vaccinia Virus to chew back the 3' ends of any vector and any virus. The virus and vector must be flanked with overlapping sequencing to anneal and seamlessly join together (Li et al., 2020). In-Fusion explicitly excludes ligation, ensuring the insert and vector join correctly and in the right direction. An additional benefit of In-Fusion cloning is its compatibility with any vector, provided that the vector and inserts share overlapping sequences introduced through PCR. Given In-Fusion's capability to join multiple DNA fragments, we hypothesized that it would

be an effective method for assembling the large genome of PVY from multiple fragments while incorporating a visual marker.

Tagging genes and viruses with a marker is a common practice used to visually track viral movement in plant tissues and identify protein localization (Tilsner & Oparka, 2010). By incorporating the code for Green Fluorescent Protein (GFP) into several economically important viruses, researchers have gained insight into cell-to-cell movement and systemic spread of plant viruses in living plants (Cohen et al., 2000; Chalfie et al., 1994; Mumford et al., 2000). PVY was successfully tagged with GFP using the yeast shuttle-vector, pCB301 (Sun et al., 2017). The vector was created by Sun et al. 2017 to stabilize PVY's large, challenging genome for agrobacterium infiltration. This method was successful but involved the construction of a complex vector and use of fluorescent microscopy to track the tagged PVY. Another 2017 study used more traditional cloning methods to create a PVY-recombinant clone tagged with an MYB-type transcription factor, Rosea1 (Cordero et al., 2017). Rosea1 triggers the production of anthocyanin, produced during virus replication, and can, therefore, be used to track virus accumulation and movement. Initially used to track tobacco etch virus (TEV), Rosea1 was used in a PVY-recombinant clone, both members of the potyvirus family (Liu & Nelson, 2013).

Although PVY cloning has been reported, and the clones were capable of infecting potatoes, there are several challenges that make replicating these studies possible with extreme modifications. Additionally, the cloning of PMTV has challenges as well. PMTV T7 transcripts were previously reported in 2003 by Savenkov capable of infecting *Nicotiana benthamiana* but have not been tested in potatoes or via Agrobacterium (Savenkov et al.,

2003). PMTV's tripartite genome provides additional factors that must be considered when cloning. RNA-RdRP of the genome is double the size (6kb) of the other two genome segments, RNA-CP (3kb) and RNA-TGB (3kb) (Reavy et al., 1998; Savenkov et al., 1999; Scott et al., 1994). All RNAs must be cloned separately while having the ability to come together in the potato host to create a virion, causing systemic infection.

With recent advancements in viral clones, our study attempted to use Rosea1 to test breeding lines for PVY and PMTV resistance. We used several methods to overcome the challenges of cloning viruses, including restriction enzyme cloning, In-Fusion cloning, LR cloning, and synthetic sequences (Figure 5.1). The challenges of cloning both viruses limited our progress in creating recombinant clones that could be used to visually track viral infections and potentially give insights into potential resistant breeding lines.

MATERIALS AND METHODS

RNA extraction and full-length PCR

PVY-infected tissue was acquired from the Charkowski lab and was used to mechanically inoculate 4-week-old *Nicotiana benthamiana*. The PVY-positive tissue was previously characterized as either PVY^{NTN} or PVY^{O} and was collected from potato leaves in the field (San Luis Valley, Colorado). Two weeks after the initial inoculation, the uppermost leaves of the tobacco were tested for PVY using immunostrip assay (Agdia, Elkhart, IN). Leaf tissue was collected from PVY-positive plants. The tissue (~100 mg) was placed in a 1.7 ml tube, flash-frozen in liquid nitrogen, and stored at -80°C until RNA extraction. Frozen leaf tissue was homogenized using a mortar and pestle with liquid nitrogen. RNA extraction was carried out as per the manufacturer's instructions using Zymo Direct-zol™ RNA Miniprep Kit (Zymo, Irvine, California). RNA quality was checked using a nanodrop before proceeding with cDNA synthesis. cDNA was created using 1 ug of RNA using the Thermo Scientific Verso cDNA kit. cDNA was subjected to PCR in an attempt to amplify full-length PVY. Primers were created for the 5' and 3' ends of PVY strains PVY^{NTN} and PVY^{O}. PCRs were initially performed with GoTaq® Flexi DNA Polymerase and 5X Colorless GoTaq® Flexi Buffer (Promega). The thermocycling protocol was as follows: 95°C for 2 mins, 34 cycles of 95°C for 1min, 55°C for 30 sec, and 72°C for 4 minutes, followed by a final extension of 72°C for 5 minutes. Different thermocycling parameters were tested by adjusting the extension time. Extension times ranging from 4-10 minutes were tested. After failed attempts, PCRs were also conducted with Q5® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA). The annealing temperature was adjusted to 60°C and different thermocycling protocol

was used: 98°C for 30 sec, 34 cycles of 98°C for 10 sec, 60 °C for 30 sec and 72°C for 3-5 minutes, followed by a final extension of 72°C for 2 minutes. Advantage® HD Polymerase Mix was also used with the same thermocycling protocol stated above (Takara Bio, San Jose, CA).

PMTV-positive RNA from tubers was acquired from the Charkowski lab. Primers from Ramesh et al., 2014 were used to amplify the PMTV genome. Primers R1F2 and R2R2 were used for full-length RNA-RdRP, R3F2, and R2R2 were used for full-length RNA-CP, and primers R2F1 and R2R2 were used for full-length RNA-TGB. PCRs were performed with GoTaq® Flexi DNA Polymerase, Q5® High-Fidelity DNA Polymerase, and Advantage® HD Polymerase Mix. After multiple unsuccessful full-length PCRs, the genome was broken down into several pieces defined by the Ramesh et al., 2014 primers.

Multiple PCRs/fragments and in-fusion cloning

After failed attempts to amplify the entire PVY genome, the PVY genome was broken down into nine overlapping pieces. Primers were created for one strain, PVY^{NTN}. The nine sections were successfully amplified using Q5® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA). The same primers were used to break down PVY into three larger ~3kb sections and two ~4.5kb pieces. The primers for the 5' end of the first fragment and the 3' end of the last fragment were designed with 15 bp extensions complementary to the ends of the linearized vector. The three 3kb sections were successfully amplified and used in Takara's In-Fusion® HD Cloning Kit (Takara Bio USA, Inc, San Jose, CA). The three PCR fragments were purified using Zymo's DNA Clean & Concentrator ® kit (Zymo, Irvine,

California). The final PCR products were nannodropped before being used in the In-Fusion system. The 10 µl reaction was composed of the three PCR fragments (30 ng each), 100 ng of plasmid, 2 µl of 5x In-Fusion enzyme, and the remaining volume was made up of nuclease-free water. Initially, we attempted to use plasmid pCB301, which was received from Zhejiang University in China and was digested with several enzymes in an attempt to validate the plasmid (Sun et al. 2017). After failed attempts, the plant vector plasmid pEarleyGate100 was ordered from NovoPro Life Sciences (Shanghai, China) instead of working with pCB301. pEarleyGate100 was digested with EcoRI for linearization before being used in the In-Fusion protocol.

After multiple unsuccessful attempts, we used TwistBio to synthesize six PVY fragments that contained 15 bp overlapping pieces with each other. The first and last fragments contained a 5bp extension that overlapped with pEarleyGate100. An intron was introduced to fragment 3, and a fluorescent tag, mCherry, was introduced to fragment 5. The fragments were resuspended with TE buffer to a 50 ng/µl concentration. A similar protocol was followed for an In-Fusion reaction. The 10 µl reaction was composed of the six TwistBio fragments (50 ng each), 100 ng of plasmid, 2 µl of 5x In-Fusion enzyme, and the remaining volume was made up of nuclease-free water.

We could not amplify the PMTV genome in fragments; alternatively, we ordered the fragments from TwistBio. Similar to this method used with PVY, the fragments were designed with 15 bp overlaps with the other fragments. The 5' and 3' ends of each genome segment was flanked with a 15 bp extension that correlated to the plasmid, pEarleyGate100. The MYB- transcription factor, rosea1, was introduced to RNA-TGB after

the CRP (Figure 5.1B). The fragments were subjected to In-Fusion reactions but with a different approach. The In-Fusion reaction was carried out with RNA-TGB and the two fragments, 3.1 and 3.2, which comprised the whole genome segment. The TwistBio fragments were resuspended with TE buffer. The two fragments were added to a PCR tube at a concentration of 50 ng each and annealed at 45°C for 30 minutes before 0.5 µl of DMSO and 0.3 u of primers were added. The reaction was cycled 24 times, starting at 45°C and increasing in 0.5°C increments with each cycle. The thermocycled reaction ran out on 1.5% agarose gel, with expectations of a 10kb band.

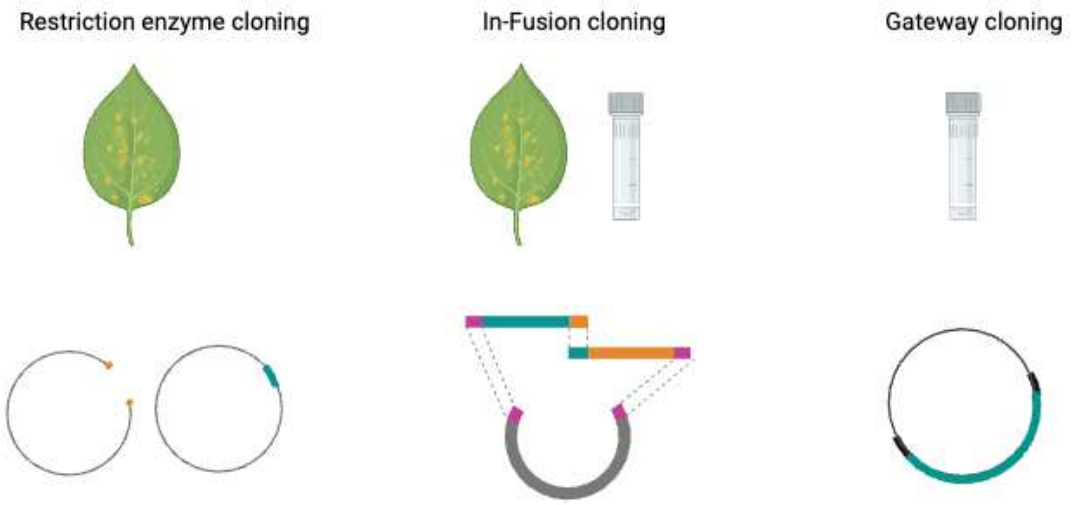

Figure 5.1 Overview of cloning methods used. Restriction enzyme cloning, In-Fusion cloning and Gateway cloning were three different cloning methods attempted to insert PVY and PMTV into plant expression vectors.

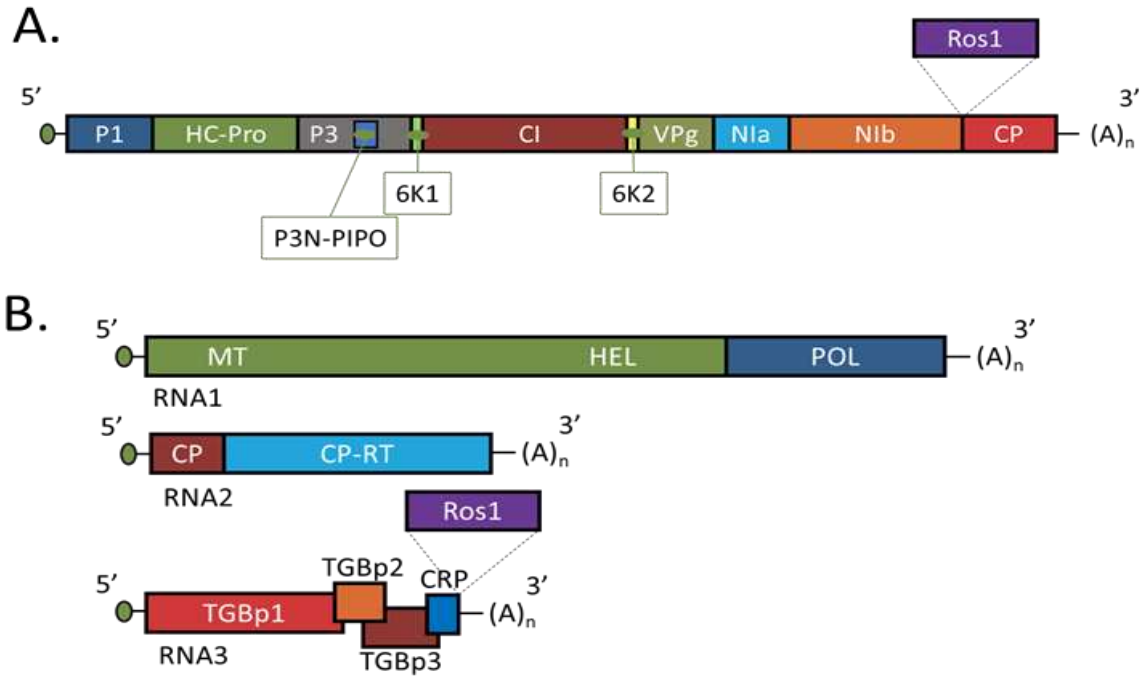

Figure 5.2 Schematic representation of the recombinant clones. (A) Rosea1 (indicated by purple box) is inserted between NIb and CP. (B). Construction of an RNA3 derivative encoding Rosea1 (indicated by a purple box) at the C terminus of RNA3 of PMTVCP. (B). Construction of an RNA3 derivative encoding Rosea1 (indicated by a purple box) at the C terminus of RNA3 of PMTV.

Ordering full-length viral sequences and LR clonase

A full-length recombinant PVY-O insert in puc57, puc57-PVY-O-I, was synthesized by Genscript (Piscataway, New Jersey) (Figure 5.2A, Figure 5.3). The insert contained an intron, the ST-LS1 gene from Solanum tuberosum, and the MYB- transcription factor, Rosea1. The insert was also modified to be flanked with attL sites. puc57-PVY-O-I was resuspended with TE buffer to a concentration of 100 ng/µl and was heat shocked into Stellar™ Competent Cells, *E.coli* HST08 strain (Takara Bio, San Jose, CA). The cells were plated on LB agar (50 µl, 100 µl, and 150 µl) with Chloramphenicol and Spectinomycin (25 & 50 µg/mL). The plates were incubated overnight at 37°C and were assessed for cell growth. The plasmid DNA from GenScript, puc57-PVY-O-I, was paired with Gateway™ LR Clonase™ II Enzyme mix in an LR clonase reaction (ThermoFisher, Waltham, MA). The destination vector was pGWB402, acquired from NovoPro Life Sciences (Shanghai, China). The LR clonase reaction was composed of 1 µl of puc57-PVY-O-I (100 ng), 1 µl of pGWB402 (150 ng), 2 µl of TE buffer, and 1 µl of Invitrogen™ Gateway™ LR Clonase™ II Enzyme. The reaction was incubated at 25°C for 2 hours before 1 µl of proteinase K was added and incubated for 10 minutes at 37°C. The reaction was placed on ice and transformed into Stellar™ Competent Cells using the heat shock method. The cells were plated on L agar (50 µl , 100 µl , and 150 µl) with Chloramphenicol and Spectinomycin (25 & 50 µg/mL). The plates were incubated overnight at 37°C and were assessed for cell growth.

For PMTV, the full-length RNA-RdRP was synthesized by GenScript in plasmid, pUC19 (Figure 5.4). The insert was flanked with attL sites since pUC19 is not a Gateway vector. First, the plasmid, pUC19-RdRp, was transformed into One Shot™ TOP10

Chemically Competent *E. coli* (ThermoFisher, Waltham, MA). The cells were plated on LB agar with ampicillin (50 µg/ml) and grown overnight at 37°C. The next day, a colony PCR was performed using RdRP-U-F/R, and positive colonies were detected. Individual colonies were grown in LB broth with ampicillin (50 µg/ml) and placed in glycerol stock.

Full-length RNA-CP and RNA-TGB were synthesized in a pENTR vector from TwistBio. The plasmids, pENTR-CP and pENTR-TGB, were resuspended to 50 ng/µl with TE buffer (Figure 5.5, Figure 5.6). The plasmid DNA from pUC19-RdRP, pENTR-CP, and pENTR-TGB was directly used in an LR clonase reaction. The LR clonase reaction was carried out using the same protocol as stated above, using pGWB402 as the destination vector. The reactions were placed on ice and transformed into One Shot™ TOP10 competent cells. The cells were plated (100 µl) on LB agar with Chloramphenicol and Spectinomycin (25 & 50 µg/ml). After overnight incubation (37°C) the colonies were tested with PCR and primers designed (RdRP-U-F/R, CP-U-F/R , TGB-U-F/R) to amplify 100bp of the genome for rapid results. Positive colonies were grown overnight at 37°C in LB with antibiotics (Chloramphenicol and Spectinomycin (25 & 50 µg/ml). Plasmid DNA was extracted from the culture according to manufacturer instructions with Zymo's ZR Plasmid Miniprep kit (Zymo). The resulting plasmids, pGWB402-RdRP, pGWB402-CP, and pGWB402-TGB, were sent to Plasmidasaraus for sequencing and confirmation of a full-length insert. The confirmed plasmids (pGWB402-CP and pGWB402-TGB) were transformed into Agrobacterium strain *GV3101* using the freeze-thaw method. The cells were plated and grown at 28°C for two days on LB agar with Rifamycin, Gentamycin, and Spectinomycin (10 µg/mL, 30 µg/mL, and 50 µg/mL). Agro cells were tested with PCR using the same primers.

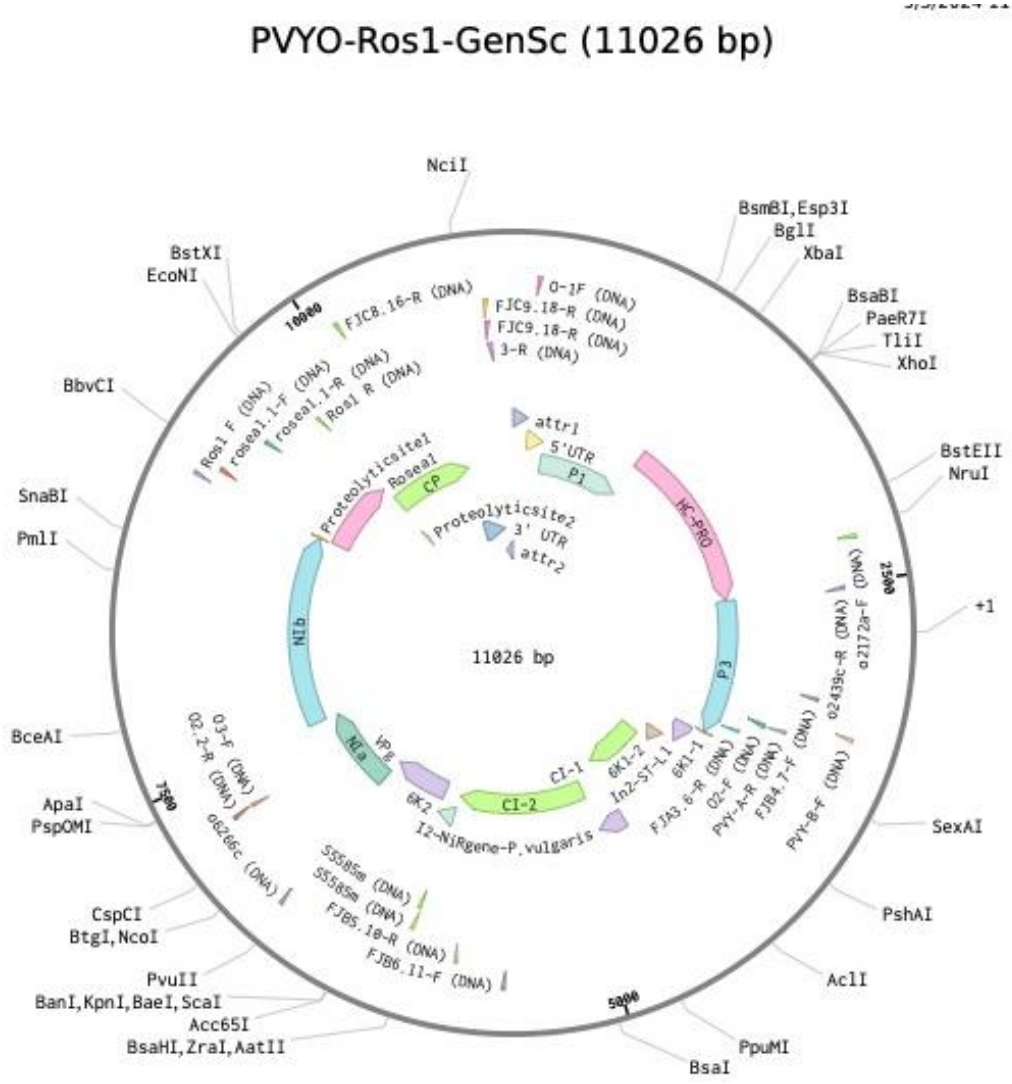

Figure 5.3 Plasmid map of pUC57-PVY-Ros1-intron synthesized by Genscript.

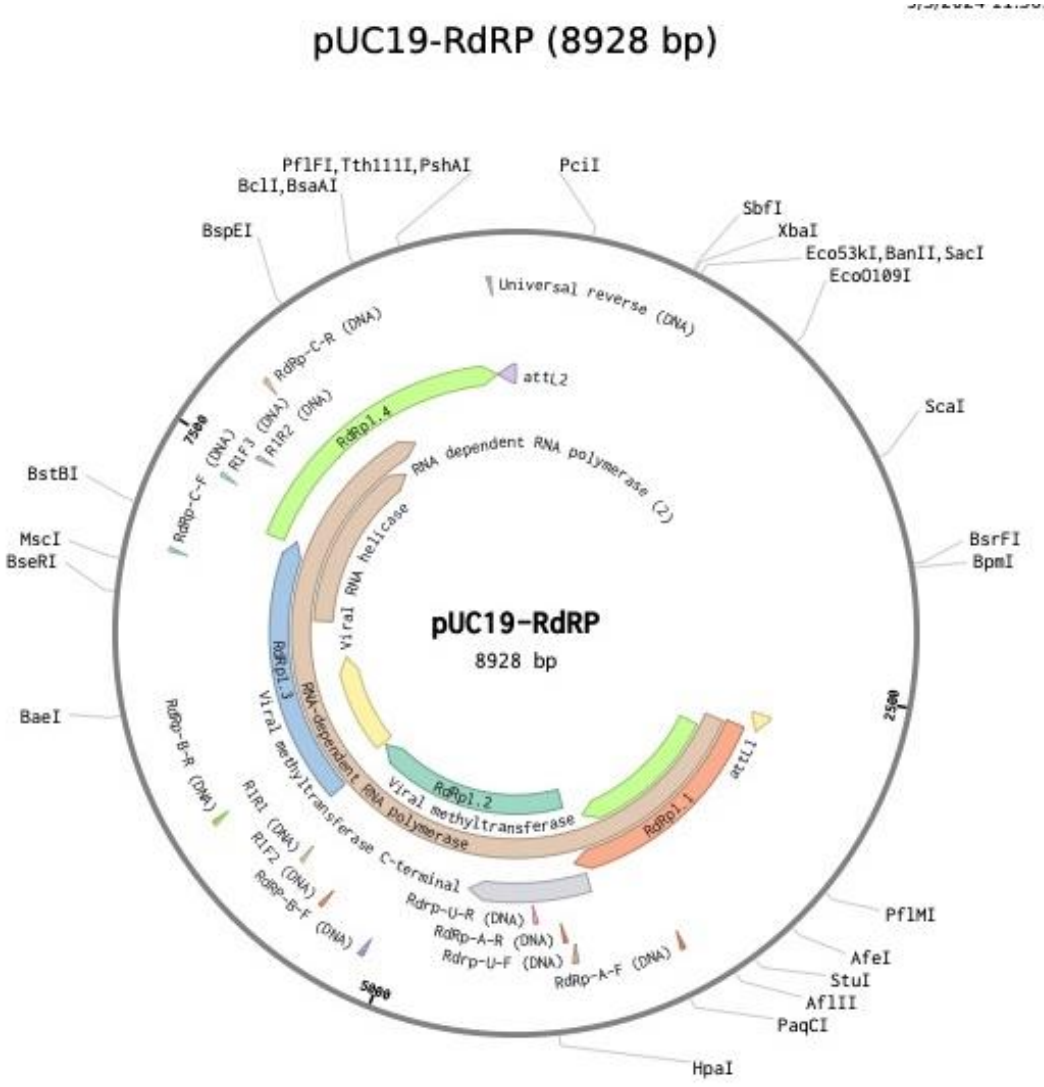

Figure 5.4 Plasmid map of pUC19-RdRP synthesized by Genscript.

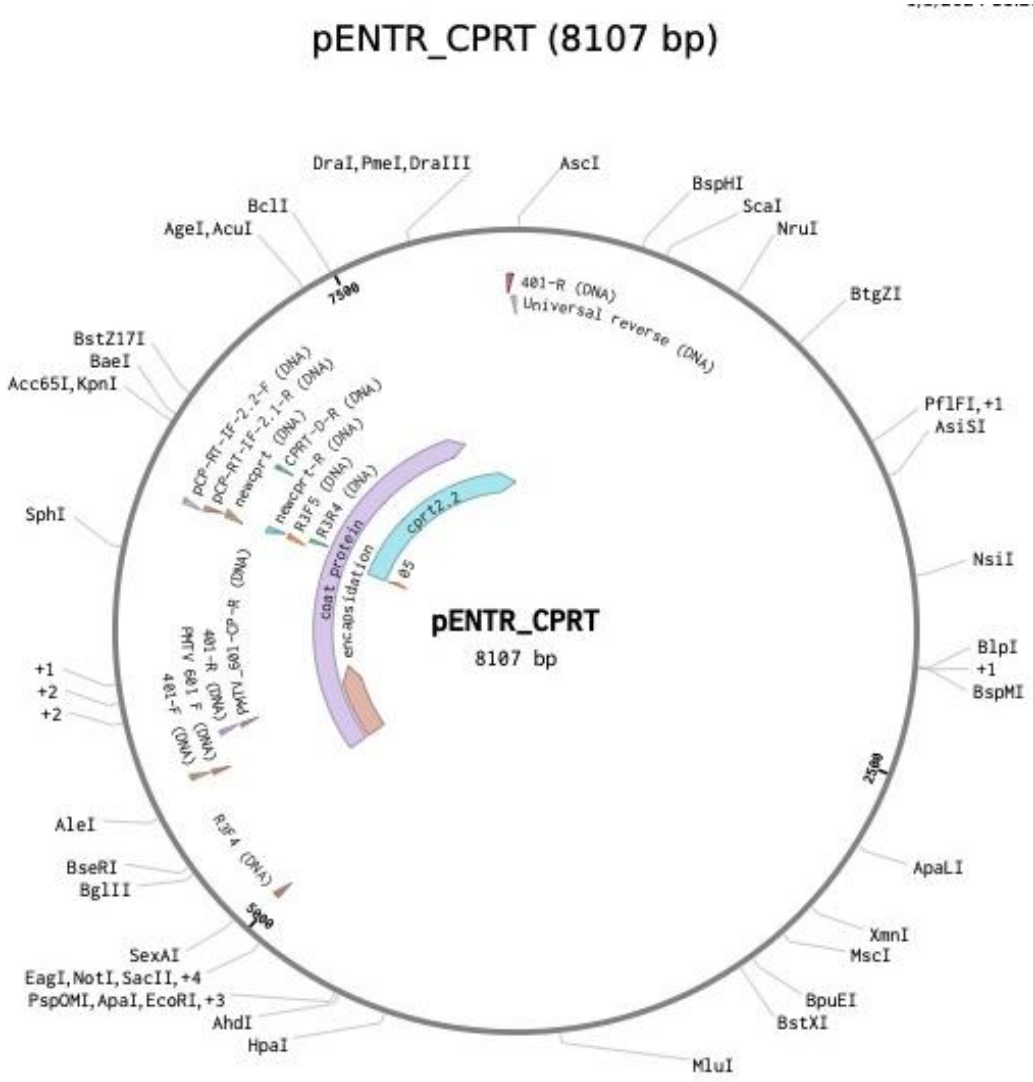

Figure 5.5 Plasmid map of pENTR-CPRT synthesized by TwistBio.

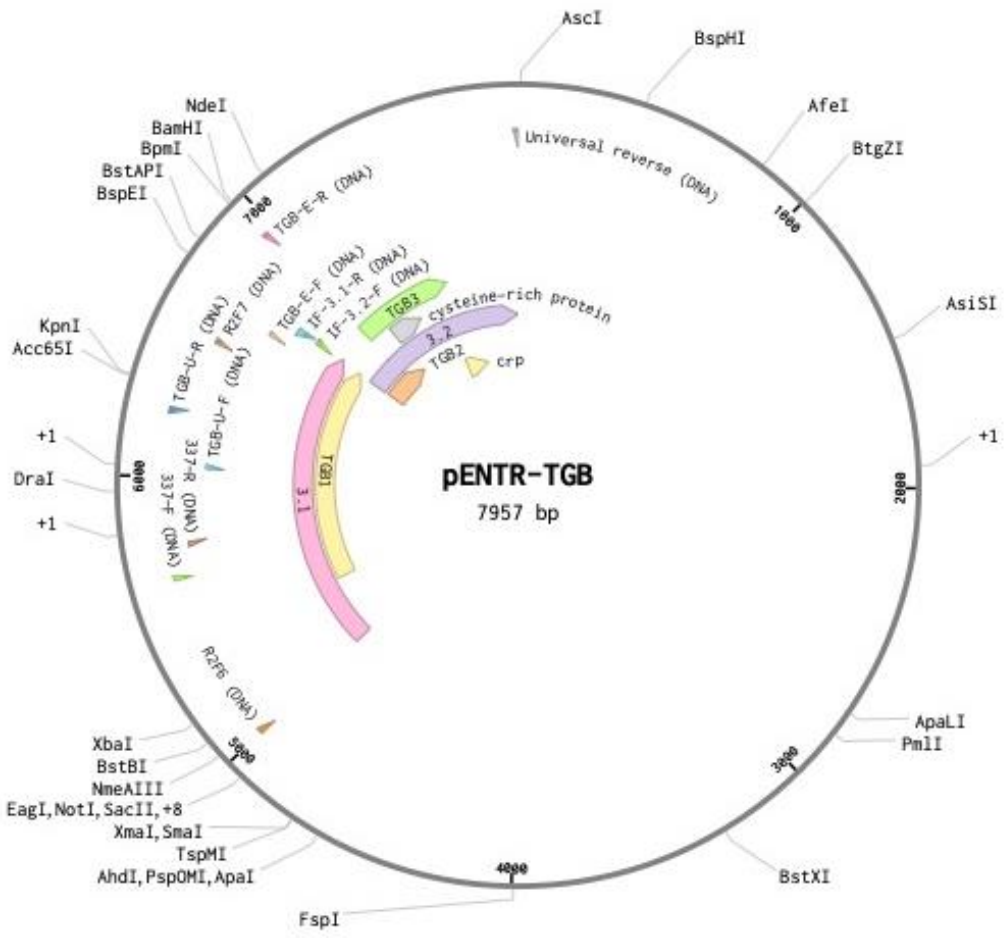

Figure 5.6 Plasmid map of pENTR-TGB synthesized by TwistBio.

RESULTS

Full-length PVY and PMTV sequences were unobtainable

The amplification of PVY and PMTV using primers that spanned the entire genome was used in various conditions. The PCR elements were altered to overcome challenges with viral amplification. The three different taq polymerases, GoTaq® Flexi DNA Polymerase, Q5® High-Fidelity DNA Polymerase, and Advantage® HD Polymerase Mix, never resulted in a full-length genome. Fragmented amplification was achieved, resulting in the PVY genome being fragmented into nine segments, which were later reduced to three larger fragments (Figure 5.7). While these fragments encompassed the entire genome, efforts to assemble them into a contiguous sequence using In-Fusion cloning were unsuccessful.

Similarly, attempts to amplify the full-length PMTV genome were unsuccessful, either in continuous or fragmented forms. Using primers from Ramesh et al. (2014), the RNA-dependent RNA polymerase (RdRP) segment was divided into three parts, while the RNA encoding the coat protein (RNA-CP) and the triple gene block protein (RNA-TGB) were each divided into two parts, resulting in seven total fragments. Of these, only four were successfully amplified. Notably, the 5' region of the RNA-RdRP and the 3' end of RNA-CP presented significant challenges in amplification.

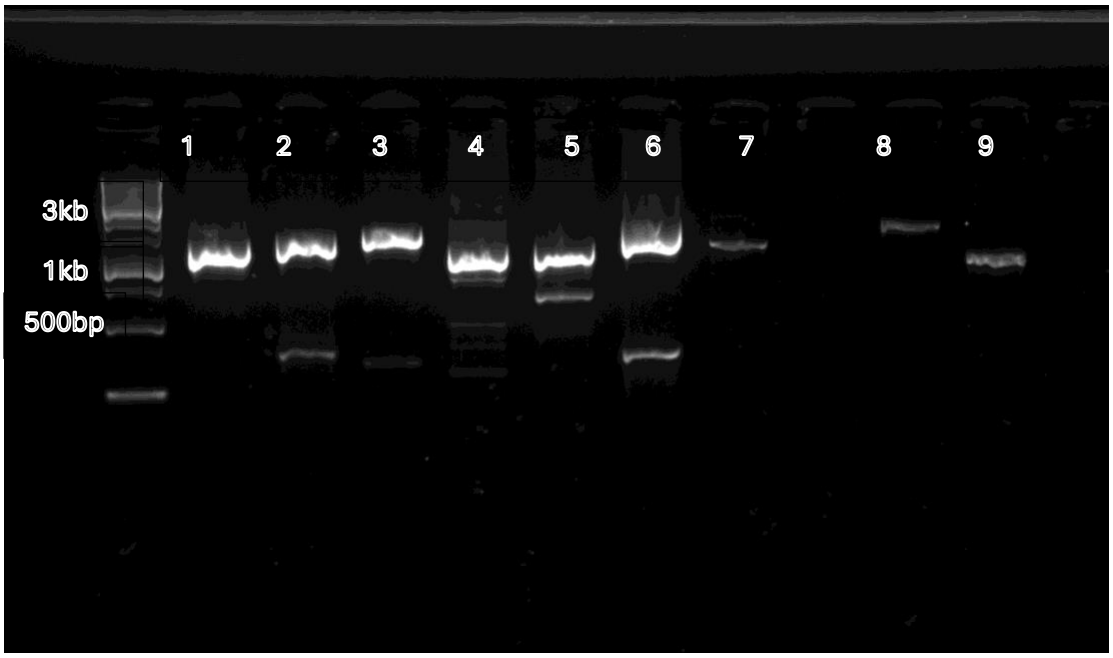

Figure 5.7 PCR results showing nine fragments of the entire PVY genome. The PVY^{NTN} genome was broken down into nine pieces and primed with overlapping fragments. The PCR was performed using Q5® and normal thermocycling conditions.

Synthesized viral fragments were unable to be assembled

Both viral genomes were synthesized as overlapping oligonucleotides of approximately 1.5 kb in length. To circumvent issues associated with In-Fusion cloning and reduce reliance on PCR, 15 bp overlaps were incorporated between the fragments, along with 5' and 3' extensions compatible with the pEarlyGate100 plasmid. Initial attempts to assemble the fragments followed the In-Fusion protocol provided by Takara Bio. Subsequently, based on communication with Dr. Robyn Roberts, an alternative method was proposed and tested once with the RNA encoding the triple gene block protein (RNA-TGB). However, we were unable to establish a successful In-Fusion protocol for the assembly of the synthesized fragments.

pGWB402_CP and pGWB402_TGB were successfully transformed into agrobacterium

Full-length PVY and RNA-RdRP were synthesized by GenScript flanked with attL sites intended for Gateway cloning. LR Clonase reactions were carried out with DNA directly from the synthesized plasmid stocks. After multiple attempts, it was concluded that the synthesized attL sites may be ineffective. RNA-RdRP was ordered in the available plasmid pUC19, which is not a Gateway vector. Although pUC19-RdRP was transformed into *E.coli* successfully, it cannot be used in Agrobacterium or as a plant expression vector. RNA-CP and RNA-TGB were ordered from TwistBio in the Gateway vector pENTR. pENTR already has attL sites and can be used in the LR Clonase reaction. The destination vector was pGWB402, which can also be used for agrobacterium. After the LR clonase reaction, pGWB402-CP and pGWB402-TGB were transformed into *E. coli* and agrobacterium, and their presence were verified by sequencing (Plasmidasarus) (Figure 5.8, Figure 5.9)).

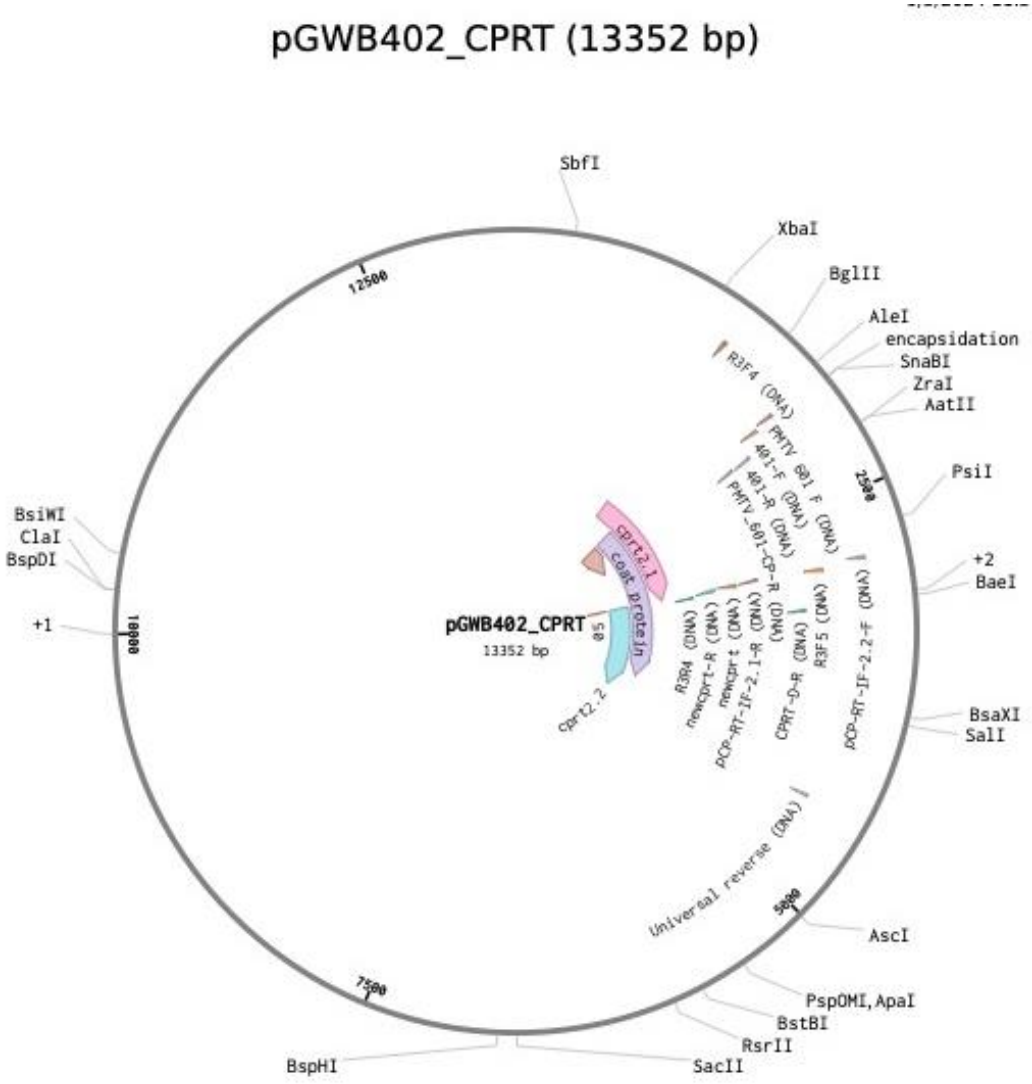

Figure 5.8 Plasmid map of pGWB402-CPRT generated from LR Clonase reaction.

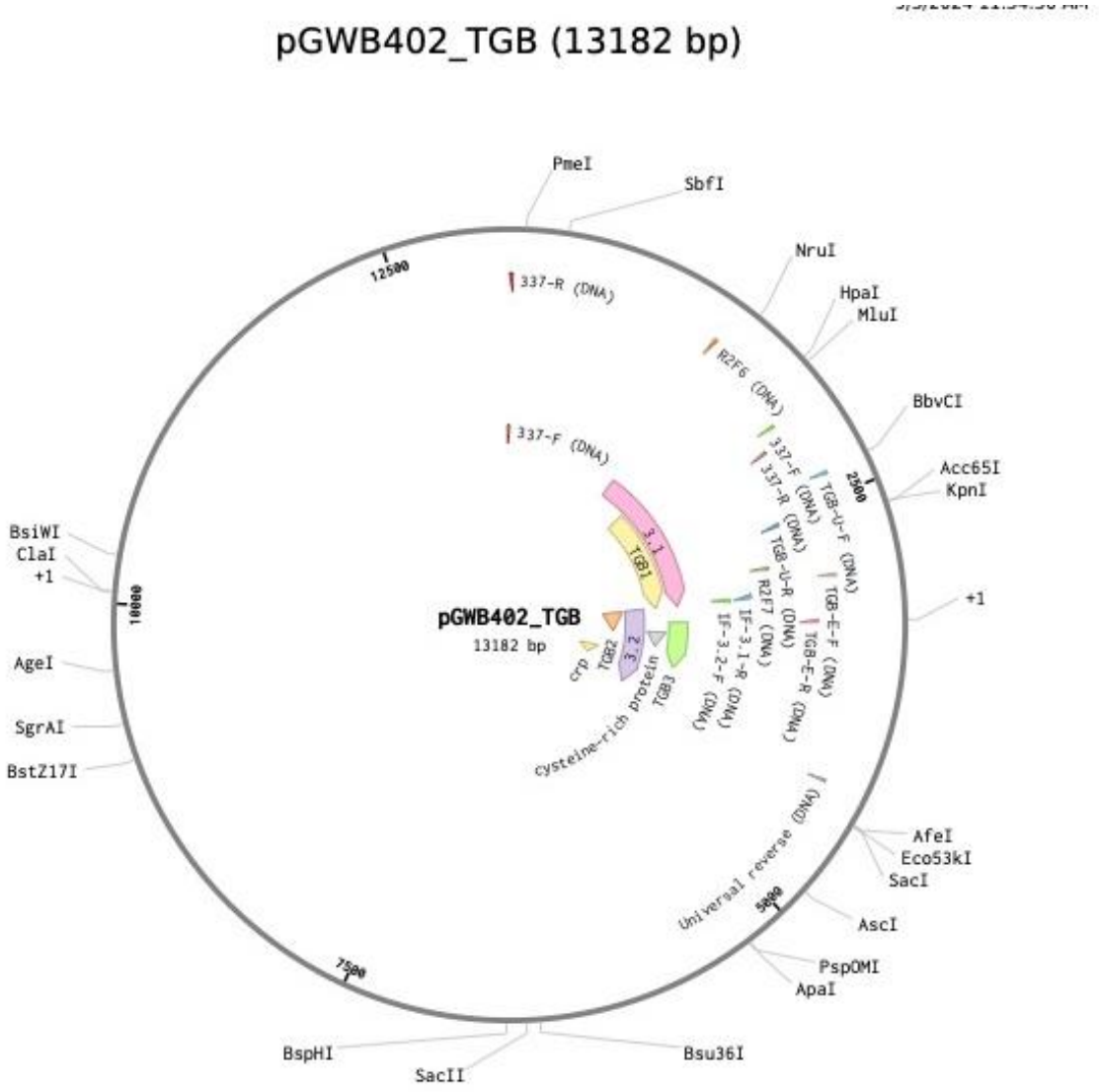

Figure 5.9 Plasmid map of pGWB402-TGB generated from LR Clonase reaction.

DISCUSSION

Full-length infectious cDNA clones are important for studying virus characteristics such as pathogenicity, movement, and interaction with their host (Bordat et al., 2015; Delbianco et al., 2013). Although beneficial, constructing plant virus infectious clones is time-consuming, complex, and often unmanageable. Additionally, recombinant clones involve subsequent cloning steps that present challenges. Various methods, such as traditional restriction enzymes, Golden Gate, Gibson Assembly, and In-Fusion cloning, have been used to create full-length infectious cDNA clones (Bordat et al., 2015; Desbiez et al., 2012; Sun et al., 2017; Tuo et al., 2015). Plant viruses belonging to the *Potyvirus* genus have been of specific interest due to their abundance and economic importance (Desbiez et al., 2012). Cloning potyviruses presents significant challenges due to their large genome size and inherent instability. The difficulty in creating a full-length clone is primarily attributed to the viral genome encoding a single, large polyprotein, which complicates the cloning process. Furthermore, potyvirus genomes contain introns that are toxic to *E.Coli*. Although there are challenges, over 20 full-length potyvirus clones have been constructed. In-Fusion cloning was used to construct papaya leaf distortion mosaic virus (PLDMV), papaya ringspot virus (PRV), and mirabilis crinkle mosaic virus (MCMV) (Li et al., 2020; Tuo et al., 2015, 2017). In-Fusion cloning has also been used to create a yeast-*E. coli* - Agrobacterium shuttle vector was used to obtain a full-length PVY infectious clone (Reference). In addition to homologous recombination in yeast, a full-length recombinant PVY was constructed using Golden Gate cloning (Cordero et al., 2017; Sun et al., 2017).

Multiple measures were taken to overcome instability, including using introns and creating the yeast-*E. coli* - Agrobacterium shuttle vector, pCB301-2μ-HDV. The insertion of Introns has been shown to disrupt the viral products that can be toxic to bacteria (López-Moya & García, 2000). In the case of PVY infectious clones, only one of the studies used introns, while the other did not, and unstable clones were produced. Both studies amplified PVY in three fragments and used non-traditional cloning methods, HR-mediated assembly, and Type IIS restriction enzymes to join the fragments. Our study aimed to use In-Fusion cloning to join three PVY fragments amplified via PCR. A significant challenge with PCR products was the inability to achieve high DNA concentrations. In the In-Fusion reactions, only 30 ng of each fragment was used, which fell below the recommended lower limit for purified PCR products. The final reaction volume was constrained to 10 μL, limiting the amount of PCR product that could be added. To overcome low concentration values, we utilized TwistBio Sciences to synthesize oligos that made up the entire PVY genome. The genome was broken down into six pieces due to the 1.5 kb oligo size limit at the time. Furthermore, we included an intron and the MYB-transcription factor, Rosea1, to avoid using subsequent cloning steps. Additionally, we incorporated 15-bp overhangs on the fragments for seamless joining procured by In-Fusion's DNA polymerase. Although the synthesized fragments allowed us to overcome concentration issues, we were still unsuccessful in our methods. The In-Fusion system has been tested, according to their website, with multiple inserts ranging from 2-5 fragments. Their study found that only four of ten clones created were correct with five inserts compared to ten of ten correct clones for two inserts (*In-Fusion Cloning: Efficient Single- and Multiple-Insert Cloning*, n.d.). According to the data

shown, the study did not test six, likely due to the low number of correct clones generated. I hypothesize the six synthesized fragments were unable to be cloned properly into our linearized pEarlyGate100 vector due to the abundance of fragments.

As technology advanced, companies such as TwistBio and GenScript offered new options for synthesis. In particular, they offered fragments to be synthesized within a plasmid. TwistBio has a limit of 5kbs on fragments and could therefore not be used for PVY or RNA-RdRP of PMTV, whereas GenScript's fragment size limit within a plasmid is 10kb and was used to synthesize the larger viral genomes. When choosing the plasmid for our fragments, various factors were considered such as compatibilities with agrobacterium and Gateway cloning. The chosen destination vector was the plant expression vector, pGBW402, and therefore planned our entry vectors accordingly. Gateway cloning uses attL and attR sites in an excision recombination reaction that swaps the desired insert from an entry plasmid into a destination vector (Reece-Hoyes & Walhout, 2018). GenScript does not offer any Gateway plasmids, as a results, we used the low copy number plasmid, pUC57, and artificially introduced attL sites. Additionally, we added two introns and Rosea1. We transformed pUC57-PVY-O-ros1 directly from the resuspended DNA stock into competent cells and PCR confirmed positive colonies. After growing the positive colonies in LB broth for plasmid extraction and validation, we were unable to secure a stable plasmid. The introns used in pUC57-PVY-O-ros1 were previously published by Sun et al., 2017, and inserted in same locations at nt positions 3517 and 4290. The study by Sun et al., 2017 stated that they obtained a stable full-length PVY infectious clone with or without the intron inserts whereas the study by Cordero et al., 2017 noted that their PVY-Ros1

infectious clone with likely unstable due to low infection rates. The presence of cryptic introns have never been proven to be present in potyviruses but intron insertion has been proven useful (López-Moya & García, 2000). Therefore, another limiting factor of the use of introns is determining the numbers and locations of intron insertions required to stabilize infectious clones (Desbiez et al., 2012; Tuo et al., 2017). Sun et al., 2017 was able to produce stable infectious PVY clones with the use of an engineered yeast shuttle vector independent of introns. Our method for creating a full-length infectious PVY clone involved introns but did not involve a modified vector or the use of homologous recombination in yeast, resulting in instability.

Since a full-length infectious pomovirus clone has never been reported, cloning methods were applied to create potyvirus full-length clones. A fundamental difference between pomoviruses and potyviruses is the pomovirus' segmented genome. In the case of PMTV, the genome consists of three segments: RNA-RdRP, RNA-CP, and RNA-TGB. Similar to pomoviruses, viruses belonging to the furovirus genus are multisegmented, with RNA-RdRP containing the replicase protein and RNA-CP containing a coat protein and a coat protein read-through domain. In a 2016 study, the furovirus, Soil-borne wheat mosaic virus (SBWMV), was cloned for gene silencing studies in wheat (Jarugula et al., 2016). A full-length infectious clone was created for RNA2, but they found RNA1 to be unstable in *E.coli*. Similar to our study, we could not construct a stable RNA-RdRP clone in plasmid pGWB402. Regarding Jarugula et al., 2016, one possibility is that RNA-RdRP contains toxic introns to *E. coli*, although we could solidify RNA-RdRP in pUC19. Therefore, another hypothesis is that the LR clonase reaction was unsuccessful. Compared to our synthesized

plasmids, pENTR-CP and pENTR-TGB, attL sites were artificially introduced to pUC19-RdRP. It has been shown that LR sites introduced by PCR are less efficient. Furthermore, RNA-RdRP is twice the size of RNA-CP and RNA-TGB, which adds to the complexity of cloning.

Beet necrotic yellow vein virus (BNYVV), a member of the *Benyvirus* genus, is a vectored by the plasmodiophorid *Polymyxa betae*. BNYVV has been cloned using two different methods, the traditional restriction enzyme, and the Gibson Assembly method (Delbianco et al., 2013; Kalantidis, 2000).The 2013 study validated their agrobacterium clones' ability to replicate in their host and move systemically under the control of the Cauliflower mosaic virus 35S promoter. Full-length cDNA clones of PMTV have been observed in vitro using T7 RNA polymerase. These clones aided in discovering RNA-CPs dispensability but lacked the *in vivo* advantages of agrobacterium clones. Our study aimed to produce full-length infectious PMTV clones that could be used directly on the potato host for resistance testing.

Ultimately, our study could not construct full-length infectious clones of PVY and PMTV. Despite several attempts, including various cloning methods, we could not overcome the challenges of PVY instability in *E. coli*. In the case of PMTV, we could not stabilize RNA-RdRP in a plant expression vector in *E.coli* and agrobacterium. The lack of knowledge regarding PMTV contributes to the challenges of cloning a multisegmented virus. Both infectious clones would have been beneficial for reverse viral genetics and host interactions. In the future, different cloning technologies, such as homologous

recombination, may be helpful in overcoming the challenges of joining fragmented pieces of a genome and creating more stable clones.

REFERENCES

Baldauf, P. M., Gray, S. M., & Perry, K. L. (2007). Biological and Serological Properties of Potato virus Y Isolates in Northeastern United States Potato. Https://Doi.Org/10.1094/PD-90-0559, 90(5), 559–566. https://doi.org/10.1094/PD-90-0559

Bordat, A., Houvenaghel, M. C., & German-Retana, S. (2015). Gibson assembly: An easy way to clone potyviral full-length infectious cDNA clones expressing an ectopic VPg Plant viruses. Virology Journal, 12(1), 1–8. https://doi.org/10.1186/S12985-015-0315-3/FIGURES/5

Brigneti, G., Garcia-Mas, J., & Baulcombe, D. C. (1997). Molecular mapping of the potato virus Y resistance gene Ry(sto) in potato. Theoretical and Applied Genetics, 94(2), 198–203. https://doi.org/10.1007/s001220050400

Chikh-Ali, M., & Karasev, A. V. (2023). Virus diseases of potato and their control. In Potato Production Worldwide (pp. 199–212). Elsevier. https://doi.org/10.1016/b978-0-12-822925-5.00008-6

Cohen, Y., Gisel, A., & Zambryski, P. C. (2000). Cell-to-Cell and Systemic Movement of Recombinant Green Fluorescent Protein-Tagged Turnip Crinkle Viruses. Virology, 273(2), 258–266. https://doi.org/10.1006/VIRO.2000.0441

Cordero, T., Mohamed, M. A., López-Moya, J. J., & Daròs, J. A. (2017). A recombinant Potato virus Y infectious clone tagged with the Rosea1 visual marker (PVY-Ros1) facilitates the analysis of viral infectivity and allows the production of large amounts of anthocyanins in plants. Frontiers in Microbiology, 8(APR), 611. https://doi.org/10.3389/FMICB.2017.00611/FULL

Crosslin, J. M., Hamm, P. B., Eastwell, K. C., Thornton, R. E., Brown, C. R., Corsini, D., Shiel, P. J., & Berger, P. H. (2002). First Report of the Necrotic Strain of Potato virus Y (PVYN) on Potatoes in the Northwestern United States. Plant Disease, 86(10), 1177. https://doi.org/10.1094/PDIS.2002.86.10.1177C

Davey, T., Carnegie, S. F., Saddler, G. S., & Mitchell, W. J. (2014). The importance of the infected seed tuber and soil inoculum in transmitting Potato mop-top virus to potato plants. Plant Pathology, 63(1), 88–97. https://doi.org/10.1111/PPA.12074

Delbianco, A., Lanzoni, C., Klein, E., Rubies Autonell, C., Gilmer, D., & Ratti, C. (2013). Agroinoculation of Beet necrotic yellow vein virus cDNA clones results in plant systemic infection and efficient Polymyxa betae transmission. Molecular Plant Pathology, 14(4), 422–428. https://doi.org/10.1111/MPP.12018

Desbiez, C., Chandeysson, C., Lecoq, H., & Moury, B. (2012). A simple, rapid and efficient way to obtain infectious clones of potyviruses. Journal of Virological Methods, 183(1), 94–97. https://doi.org/10.1016/J.JVIROMET.2012.03.035

Domfeh, O., Bittara, F. G., & Gudmestad, N. C. (2015). Sensitivity of Potato Cultivars to Potato mop-top virus-Induced Tuber Necrosis. Plant Disease, 99(6), 788–796. https://doi.org/10.1094/PDIS-07-14-0705-RE

Elison, G. L., Novy, R. G., & Whitworth, J. L. (2021). Russet Potato Breeding Clones with Extreme Resistance to Potato Virus Y Conferred by Rychc as well as Resistance to Late Blight and Cold-Induced Sweetening. American Journal of Potato Research, 98(5–6), 411–419. https://doi.org/10.1007/S12230-021-09852-1/TABLES/4

Frost, K. E., Groves, R. L., & Charkowski, A. O. (2013). Integrated control of potato pathogens through seed potato certification and provision of clean seed potatoes. Plant Disease, 97(10), 1268–1280. https://doi.org/10.1094/PDIS-05-13-0477-FE/SUPPL_FILE/PDIS-05-13-0477-FEE.PDF

Green, K. J., Brown, C. J., & Karasev, A. V. (2018). Genetic diversity of potato virus Y (PVY): sequence analyses reveal ten novel PVY recombinant structures. Archives of Virology, 163(1), 23–32. https://doi.org/10.1007/s00705-017-3568-x

Jansky, S. H., Charkowski, A. O., Douches, D. S., Gusmini, G., Richael, C., Bethke, P. C., Spooner, D. M., Novy, R. G., De Jong, H., De Jong, W. S., Bamberg, J. B., Thompson, A. L., Bizimungu, B., Holm, D. G., Brown, C. R., Haynes, K. G., Sathuvalli, V. R., Veilleux, R. E., Creighton Miller, J., … Jiang, J. (2016). Reinventing Potato as a Diploid Inbred Line–Based Crop. Crop Science, 56(4), 1412–1422. https://doi.org/10.2135/CROPSCI2015.12.0740

Jarugula, S., Charlesworth, S. R., Qu, F., & Stewart, L. R. (2016). Soil-borne wheat mosaic virus infectious clone and manipulation for gene-carrying capacity. Archives of Virology, 161(8), 2291–2297. https://doi.org/10.1007/S00705-016-2863-2/FIGURES/3

Jones, R. A. C., & Harrison, B. D. (1969). The behaviour of potato mop-top virus in soil, and evidence for its transmission by Spongospora subterranea (Wallr.) Lagerh. Annals of Applied Biology, 63(1), 1–17. https://doi.org/10.1111/j.1744-7348.1969.tb05461.x

Kalantidis, K. (2000). Generation of 13k-gene sugar beet transformants and evaluation of their resistance to BNYVV infection. Developments in Plant Genetics and Breeding, 6(C), 189–194. https://doi.org/10.1016/S0168-7972(00)80121-8

Karasev, A. V., & Gray, S. M. (2013). Continuous and emerging challenges of potato virus y in potato. Annual Review of Phytopathology, 51, 571–586. https://doi.org/10.1146/annurev-phyto-082712-102332

Karasev, A. V., Nikolaeva, O. V., Hu, X., Sielaff, Z., Whitworth, J., Lorenzen, J. H., & Gray, S. M. (2010). Serological properties of ordinary and necrotic isolates of Potato virus Y: A case study of PVYN misidentification. American Journal of Potato Research, 87(1), 1–9. https://doi.org/10.1007/S12230-009-9110-2/TABLES/2

Lambert, D. H., Levy, L., Mavrodieva, V. A., Johnson, S. B., Babcock, M. J., & Vayda, M. E. (2007). First Report of Potato mop-top virus on Potato from the United States. Https://Doi.Org/10.1094/PDIS.2003.87.7.872A, 87(7), 872–872. https://doi.org/10.1094/PDIS.2003.87.7.872A

Laufer, M., Mohammad, H., Maiss, E., Richert-Pöggeler, K., Dall'Ara, M., Ratti, C., Gilmer, D., Liebe, S., & Varrelmann, M. (2018). Biological properties of Beet soil-borne mosaic virus and Beet necrotic yellow vein virus cDNA clones produced by isothermal in vitro recombination: Insights for reassortant appearance. Virology, 518, 25–33. https://doi.org/10.1016/J.VIROL.2018.01.029

Li, X., Li, Y., Chen, S., & Wang, J. (2020). Construction of stable infectious full-length and eGFP-tagged cDNA clones of Mirabilis crinkle mosaic virus via In-Fusion cloning. Virus Research, 286, 198039. https://doi.org/10.1016/J.VIRUSRES.2020.198039

Liu, C., & Nelson, R. S. (2013). The cell biology of Tobacco mosaic virus replication and movement. Frontiers in Plant Science, 4(FEB). https://doi.org/10.3389/FPLS.2013.00012

López-Moya, J. J., & García, J. A. (2000). Construction of a stable and highly infectious intron-containing cDNA clone of plum pox potyvirus and its use to infect plants by particle bombardment. Virus Research, 68(2), 99–107. https://doi.org/10.1016/S0168-1702(00)00161-1

Lorenzen, J. H., Piche, L. M., Gudmestad, N. C., Meacham, T., & Shiel, P. (2006). A multiplex PCR assay to characterize Potato virus Y isolates and identify strain mixtures. Plant Disease, 90(7), 935–940. https://doi.org/10.1094/PD-90-0935

M, C., Y, T., G, E., WW, W., & DC, P. (1994). Green fluorescent protein as a marker for gene expression. Science (New York, N.Y.), 263(5148), 1766–1767. https://doi.org/10.1126/SCIENCE.8303295

Merz, U. (2008). Powdery scab of potato - Occurrence, life cycle and epidemiology. American Journal of Potato Research, 85(4), 241–246. https://doi.org/10.1007/S12230-008-9019-1/FIGURES/2

Mumford, R. A., Walsh, K., Barker, I., & Boonham, N. (2000). Detection of Potato mop top virus and Tobacco rattle virus Using a Multiplex Real-Time Fluorescent Reverse-Transcription Polymerase Chain Reaction Assay.

Nolte, P., Whitworth, J. L., Thornton, M. K., & McIntosh, C. S. (2004). Effect of Seedborne Potato virus Y on Performance of Russet Burbank, Russet Norkotah, and Shepody Potato. Plant Disease, 88(3), 248–252. https://doi.org/10.1094/PDIS.2004.88.3.248

Pelletier, Y., Nie, X., Giguère, M. A., Nanayakkara, U., Maw, E., & Foottit, R. (2012). A new approach for the identification of aphid vectors (Hemiptera: Aphididae) of potato virus Y. Journal of Economic Entomology, 105(6), 1909–1914. https://doi.org/10.1603/EC12085

Reavy, B., Arif, M., Cowan, G. H., & Torrance, L. (1998). Association of sequences in the coat protein/readthrough domain of potato mop-top virus with transmission by Spongospora subterranea. Journal of General Virology, 79(10), 2343–2347. https://doi.org/10.1099/0022-1317-79-10-2343/CITE/REFWORKS

Reece-Hoyes, J. S., & Walhout, A. J. M. (2018). Gateway Recombinational Cloning. Cold Spring Harbor Protocols, 2018(1), pdb.top094912. https://doi.org/10.1101/PDB.TOP094912

Savenkov, E. I., Germundsson, A., Zamyatnin, A. A., Sandgren, M., & Valkonen, J. P. T. (2003). Potato mop-top virus: The coat protein-encoding RNA and the gene for cysteine-rich protein are dispensable for systemic virus movement in Nicotiana benthamiana. Journal of General Virology, 84(4), 1001–1005. https://doi.org/10.1099/VIR.0.18813-0/CITE/REFWORKS

Savenkov, E. I., Sandgren, M., & Valkonen, J. P. T. (1999). Printed in Great Britain Complete sequence of RNA 1 and the presence of tRNA-like structures in all RNAs of Potato mop-top virus, genus Pomovirus. In Journal of General Virology (Vol. 80).

Scott, K. P., Kashiwazaki, S., Reavy, B., & Harrison, B. D. (1994). The nucleotide sequence of potato mop-top virus RNA 2: a novel type of genome organization for a furovirus. In Journal of General Virology (Vol. 3561).

Sun, K., Zhao, D., Liu, Y., Huang, C., Zhang, W., & Li, Z. (2017). Rapid Construction of Complex Plant RNA Virus Infectious cDNA Clones for Agroinfection Using a Yeast-E. coli-Agrobacterium Shuttle Vector. Viruses 2017, Vol. 9, Page 332, 9(11), 332. https://doi.org/10.3390/V9110332

Tilsner, J., & Oparka, K. J. (2010). Tracking the green invaders: advances in imaging virus infection in plants. The Biochemical Journal, 430(1), 21–37. https://doi.org/10.1042/BJ20100372

Tuo, D., Fu, L., Shen, W., Li, X., Zhou, P., & Yan, P. (2017). Generation of stable infectious clones of plant viruses by using Rhizobium radiobacter for both cloning and inoculation. Virology, 510, 99. https://doi.org/10.1016/J.VIROL.2017.07.012

Tuo, D., Shen, W., Yan, P., Li, X., & Zhou, P. (2015). Rapid Construction of Stable Infectious Full-Length cDNA Clone of Papaya Leaf Distortion Mosaic Virus Using In-Fusion Cloning. Viruses 2015, Vol. 7, Pages 6241-6250, 7(12), 6241–6250. https://doi.org/10.3390/V7122935

Whitworth, J. L., & Crosslin, J. M. (2012). Detection of Potato mop top virus (Furovirus) on potato in southeast Idaho. Https://Doi.Org/10.1094/PDIS-08-12-0707-PDN, 97(1), 149. https://doi.org/10.1094/PDIS-08-12-0707-PDN

Whitworth, J. L., Novy, R. G., Hall, D. G., Crosslin, J. M., & Brown, C. R. (2009). Characterization of broad spectrum potato virus Y resistance in a Solanum tuberosum ssp. andigena-Derived population and select breeding clones using molecular markers, grafting, and field inoculations. American Journal of Potato Research, 86(4), 286–296. https://doi.org/10.1007/S12230-009-9082-2/TABLES/6

Yellareddygari, S. K. R., Whitworth, J. L., & Gudmestad, N. C. (2018). Assessing potato cultivar sensitivity to tuber necrosis caused by potato mop-top virus. Plant Disease, 102(6), 1148–1153. https://doi.org/10.1094/PDIS-10-17-1585-RE/ASSET/IMAGES/LARGE/PDIS-10-17-1585-RE_T6.JPEG

CHAPTER 6 – CONCLUSION AND IMPACTS

Necrotic potato viruses are economically important pathogens that qualitatively affect the potato, resulting in financial and yield losses. The management of the viruses heavily relies on resistant cultivars and detection to prevent the spread of viruses in the propagative potato. PVY is widely regulated with the help of seed certification programs and monitoring methods which include visual inspections, ELISA, immunocapture and PCR. Our study used molecular methods to monitor and characterize PVY in San Luis, Valley Colorado over two years, from 2021-2022. The studied employed HTS to address recombination and mixed infections and the potato virome of PVY infected plants. Although we did not find new recombinant strains, several mixed infections were detected and confirmed through HTS. The discovery of mixed infections is a concern as they can give rise to a recombination event, emphasizing the importance of continued monitoring of PVY. Furthermore, we discovered SNPs in primers used to characterize PVY strains, this presents a challenge and additional need for monitoring for it will contribute to accurate detection, which is necessary to the management of PVY. Our study extensively contributes to the molecular characterization of PVY and is the first to report a potato virome in Colorado. The discovery of strain dynamics in the field can assist in the management of PVY by showing the importance of molecular detection and the presence of PVY evolution, increasing detection accuracy.

In the case of the soil borne pathogens, PMTV and *S. Subterranea*, there are no effective management strategies such resistance or soil treatments; therefore current management relies on detection and favorable conditions. Over study emphasizes the

important of the environment for infections while advancing detection. To gain insights about the virus-vector relationship, we analyzed populations of both pathogens. While there was little diversity in PMTV, we found distinct populations of *S. Subterranea*. Pathogen incidence was examined in field soils and in baited tomatoes, to further evaluate the relationship between the pathogens. We found that state was the biggest factor influencing infections as well as the favorable conditions of the tomato bait assay. Through the discovery of new PMTV isolates, we were able to develop new primer-probe sets to detect all segments of the PMTV genome. Our data shows a high presence of TGB which could alter the way PMTV is detected, specifically in cases where PMTV presence is low. The comparison of ddPCR detection to the traditional tomato bait assay was important in validating our new primers and highlighting the accuracy of the tomato bait assay. Overall, the study greatly contributes to virus detection. More importantly, the clear correlation of pathogens in favorable conditions supports the environment to be a key factor in pathogen incidence is therefore vital to management.

APPENDICES

A review on PMTV published in the journal *Outlooks on Pest Management*

Rushton, J. and Nalam, V.J., 2021. Potato mop-top virus: A rising threat to potato production worldwide. *Outlooks on Pest Management*, *32*(1), pp.13-17.

POTATO MOP-TOP VIRUS: A RISING THREAT TO POTATO PRODUCTION WORLDWIDE

The potato mop-top virus (PMTV) is rapidly becoming a significant problem, especially in seed potato production. As vegetatively propagated plants, potatoes tend to accumulate viruses over time, presenting a considerable challenge to sustainable and profitable production. The presence of even small numbers of infected tubers can lead to the rejection of an entire crop destined for either the processing industry or pre-packed supermarket trade. PMTV has been reported from all major potato growing regions of the U.S., and its incidence is increasing. For example, in 2020, PMTV was reported in only 1% of seed potatoes in Maine (Lambert *et al.* 2003) but is currently believed to be present in 5% of Maine's seed potatoes. A factor contributing to the spread of this virus is that the virus is vectored by a soil-borne plasmodiophorid pathogen that causes powdery scab disease on potato. PMTV rarely causes foliar symptoms making its presence difficult to detect until harvest. If left unchecked, PMTV-induced tuber necrosis is predicted to become as great a problem in the U.S. as it is in Europe (Santala *et al.* 2010). A better understanding of the relationship between the virus, its vector, and factors that influence their spread will enable the development of effective management strategies and provide growers with on-

farm management options. This review provides a summary of the current state of knowledge on PMTV and its vector.

Disease Symptoms. PMTV causes the potato mop-top disease, and characteristic symptoms referred to as “spraing”, which appear as internal necrotic dark-colored arcs or concentric rings in the tuber (Figure A1.A). On occasion, external tuber and foliar symptoms may also appear. The infection of potato plants by PMTV can be categorized as either primary or secondary, depending on when plants contract the virus. The type of infection correlates to the observed symptoms (Calvert & Harrison 1966; Harrison & Jones 1970). Planting virus-free seed tubers in PMTV infested fields results in primary infections, and symptoms are observed mainly in the tuber as spraing/necrotic arcs. Secondary PMTV infections, i.e., in plants that grow from previously infected tubers, show symptoms that include distortion of leaves, yellow mosaic, and V-shaped flecks in the leaves (Figure A1.B), and severe shortening of the internodes. The virus and the disease it causes derives its name from the dwarfed or ‘mop-top’ appearance of infected plants (Calvert & Harrison 1966). The appearance of symptoms, primary and/or secondary, depends on several factors: including cultivar, moisture, temperature, virus strain, and viral load. Symptoms can also vary within a field since the virus's distribution and its vector can be patchy (Kirk 2008).

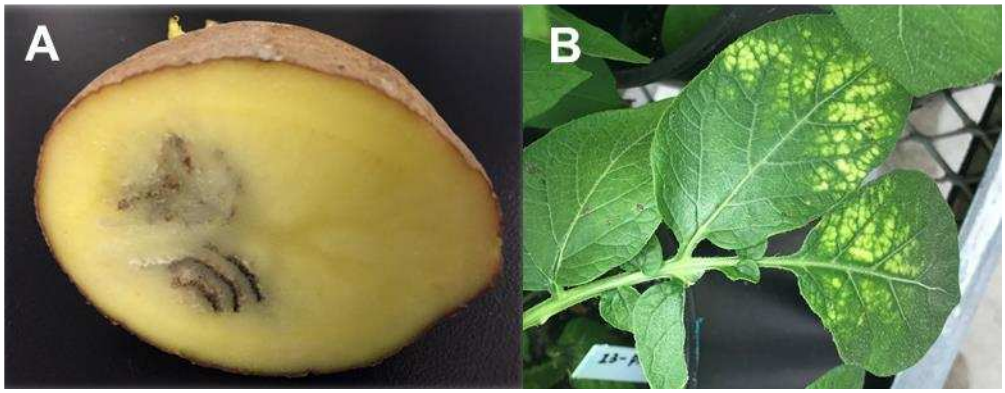

Figure A1. ***Potato mop-top virus symptoms*****.** (A) Symptoms of PMTV infection of the tubers appear as spraing or necrotic arcs in the tuber (Photo credit: Ana Cristina Fulladolsa). (B) Foliar symptoms appear as V-shaped yellow flecks on leaves (Cultivar: Satina) (Photo credit: Yuan Zeng).

PMTV can remain infective in a field for 10 -12 years without potatoes (Merz 1995). The long term persistence of the virus is due to its ability to stay viable in resting spores of its vector, *Spongospora subterranea* f. sp. *subterranea* (Referred to *Sss* hereafter). The virus can also infect species other than potatoes. PMTV has a narrow host range infecting a few members of only three families (Solanaceae, Amaranthaceae, and Aizoaceae) of plants whose members include economically important crops such as tomato, tobacco, sugarbeets, and spinach.

The Vector. *Sss*, the causal agent of powdery scab on potato (Figure A2), acts as a vector for PMTV. In addition to affecting tubers, this plasmodiophorid pathogen causes root galling (Figure A2). Upon plant infection (roots, stolons, and tubers), root galls or scab-like lesions containing sporosori may be produced. On tubers, these lesions are especially damaging since they reduce the marketable quality of the crop and become inoculum when *Sss*-infected tubers are used as seed (Merz 2008). *Sss* has a broad host range and can survive on neighboring weeds and plants - including plants in the Solanaceae family and economically important crops such as tomato and tobacco (Tsror *et al.* 2020).

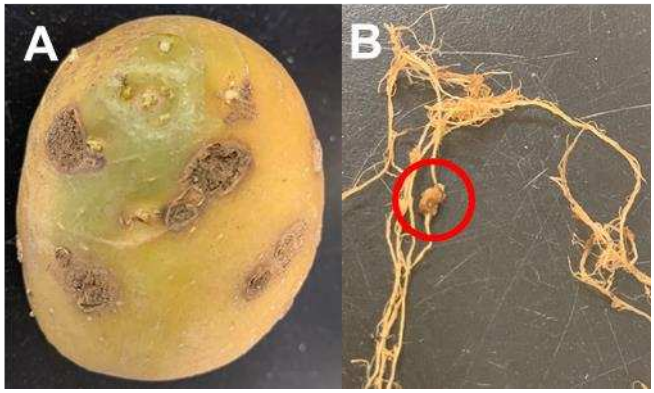

Figure A2. *Powdery scab on potato*. (A) Mature symptoms caused by *Sss* infection of tubers. Individual blisters have coalesced into raised shallow lesions with raised margins. (B) Root galls (red circle) observed on roots. (Photo credit: Yuan Zeng).

Sss is an obligate biotroph, meaning it requires a living host to survive and reproduce and being a plasmodiophorid, it produces resting spores and motile zoospores. Our current understanding of the life cycle of this organism indicates that the pathogen has two main phases. During the first or asexual phase of its life cycle, primary motile zoospores, released from resting spores, swim in moist soil towards plant tissue and encyst on roots or tuber surfaces. The zoospores are attracted to a range of organic molecules produced by plant roots and secreted into the rhizosphere (Balendres *et al*. 2016b). Once in plant tissue, zoospores grow and mature into zoosporangia containing secondary zoospores. On roots and stolons, the zoosporangia create galls, and on tubers, lesions containing sporosori (sacks of spores) are formed (Figure A3). Upon maturity, both the zoosporangia and the sporosori release secondary zoospores back into the soil. These spores can re-infect adjacent plants or tubers, causing a rapid inoculum build-up in the field within a single growing season. The second or sexual phase of the pathogen begins when two zoospores fuse and form a binucleate structure capable of infecting plant tissue. After meiosis, three-walled resting spores are formed that aggregate into sporosori, which can persist in the soil for 8-10 years.

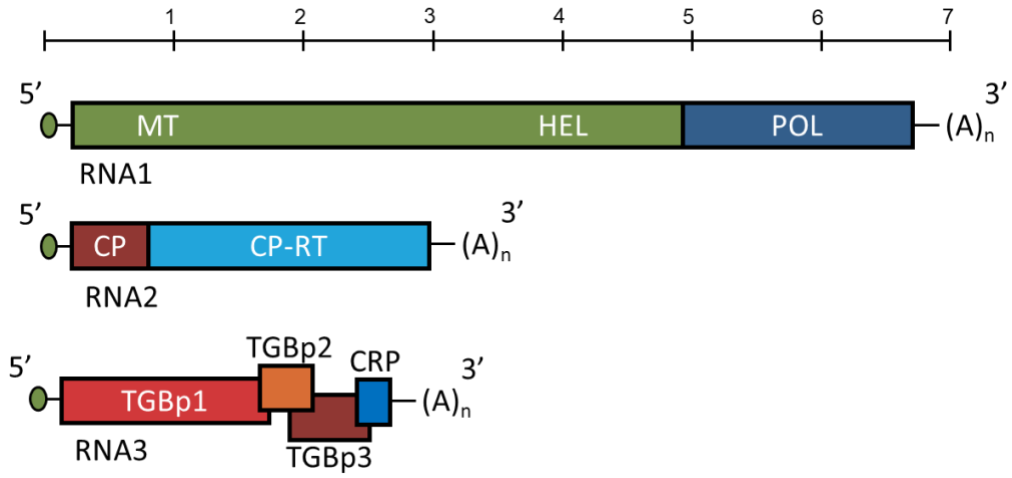

Figure A3. Genome organization of PMTV. The virus has a tripartite genome: RNA1, RNA2, and RNA3. The ruler on top represents the size of the PMTV genome in kilobases. The boxes denote viral cistrons: MT, methyltransferase domain; HEL, helicase domain; POL, polymerase domain; CP, Coat protein; RT, readthrough region. TGB, triple gene block.

Occurrence, Distribution, and Diversity. PMTV currently has a worldwide distribution and occurs in all potato growing regions of the world. Since its discovery and description in Scotland and Northern Ireland in 1965 (Calvert & Harrison 1966), most other European countries have reported PMTV presence (Santala *et al.* 2010). Within six decades, it spread throughout the world. In the U.S., the first report of the virus was from Maine in 2002 (Lambert *et al.* 2003). By 2013 the virus had spread to all major potato growing regions, including North Dakota, Washington, Idaho, Colorado, Michigan, and New Mexico (David *et al.* 2010; Whitworth & Crosslin 2013). In Asia, PMTV was first reported from Japan (Imoto *et al.* 1986), followed by reports from Pakistan (Arif *et al.* 2014) and China (Hu *et al.* 2016). In 2018, New Zealand reported the virus, giving PMTV a genuinely global presence (Government of NZ, 2018). Phylogenetic analysis of PMTV strains worldwide indicates that the virus arrived in Europe from Peru or Colombia (Kalyandurg *et al.* 2017). This predicted path of PMTV spread also parallels that of its vector (Gau *et al.* 2013).

Environmental conditions play a crucial role in PMTV incidence and distribution. Soil moisture, rainfall, and temperature have a strong influence on the distribution of the virus. In general, the virus thrives when the soil is cool and damp (Carnegie *et al.* 2010, 2012), largely because these conditions are conducive for the vector, *Sss* (Balendres *et al.* 2016a). Accordingly, the incidences of PMTV tend to increase following heavy rainfall (Kirk 2008). In many parts of the world and the U. S., potato production has intensified, resulting in increased acreage under irrigation, and soil moisture conditions are favorable for the spread and infection of PMTV. In addition to both pathogens having similar trends of distribution, parallels can be also be drawn concerning their genetic diversity. According to

a study analyzing the global genetics of *Sss*, North American strains exhibit low diversity (Gau *et al.* 2013).

PMTV is a three-part single-stranded RNA virus and is considered a type member of the genus *Pomovirus* (Figure A3). There are two main strains of PMTV, the S (for severe) strain, and the M (for mild) strain, which differ from each other only in seven amino acids in the CP-RT (Figure A3) region on RNA2. In North America, Europe, and Asia (Kalyandurg *et al.* 2017), the S-type strains are prevalent. Although a higher diversity of strains are present in the Andean region (Zhai *et al.* 2020), the M-type strain is most prevalent. The current global distribution of S-type strains suggests that its unique seven amino acids increase its ability to be acquired and transmitted by *Sss*. However, further studies are required to determine if this is indeed the case. When a secondary infection occurs, the symptoms caused by S and M strains also differ. As their names indicate, S-type strains tend to cause brighter yellow mosaic symptoms in the foliage, while M-type strains generally do not cause any discernible symptoms.

Since the genetic analysis of the first isolate of PMTV from the U. S. (Ramesh *et al.* 2014), several more isolates from Idaho, Maine, and Maryland have been analyzed to gain insight into the genetic diversity of PMTV (Zhai *et al.* 2020). PMTV strains in the U.S. are highly conserved, with a difference of less than 2% of their genetic material between all characterized isolates (Zhai *et al.* 2020). A detailed analysis of the strains' differences reveals that a small region on RNA2 (codon 689 on CP-RT) is under positive selection, meaning variation in this region is advantageous to the pathogen. The rest of the genome is under negative selection pressure (Zhai *et al.* 2020). This region is associated with vector

acquisition and transmission, providing an insight into the relationship between PMTV and its vector. Without the ability to enter the vector, PMTV loses its pathogenicity to potato, meaning that PMTV is under pressure to adapt to changes in *Sss* to be a successful pathogen (Reavy *et al.* 1995).

Integrated Management of PMTV and *S. subterranea*. The virus's ability to cause severe economic losses worldwide has necessitated developing and deploying effective management strategies. In the absence of effective virus and vector management, infestations will soon become ubiquitous. To mitigate the virus's economic impact, researchers have identified cultivars that are sensitive, moderately sensitive/insensitive, or insensitive to the expression of tuber necrosis (Bittara *et al.* 2016; Domfeh *et al.* 2015; Yellareddygari *et al.* 2018). Such studies could inform growers on cultivar selection to manage PMTV-induced tuber necrosis under field and storage conditions.

Currently, the management of PMTV mainly relies on the control of the vector, *Sss* (Figure A4). Although several potato cultivars with different levels of resistance to *S. subterranea* have been identified, no cultivar is entirely resistant to powdery scab (Bittara *et al.* 2016). Many popular russet varieties are resistant to powdery scab formation on tubers but are susceptible to root infection. A positive correlation was observed between the number of galls per plant and PMTV-induced tuber necrosis (Domfeh *et al.* 2015).

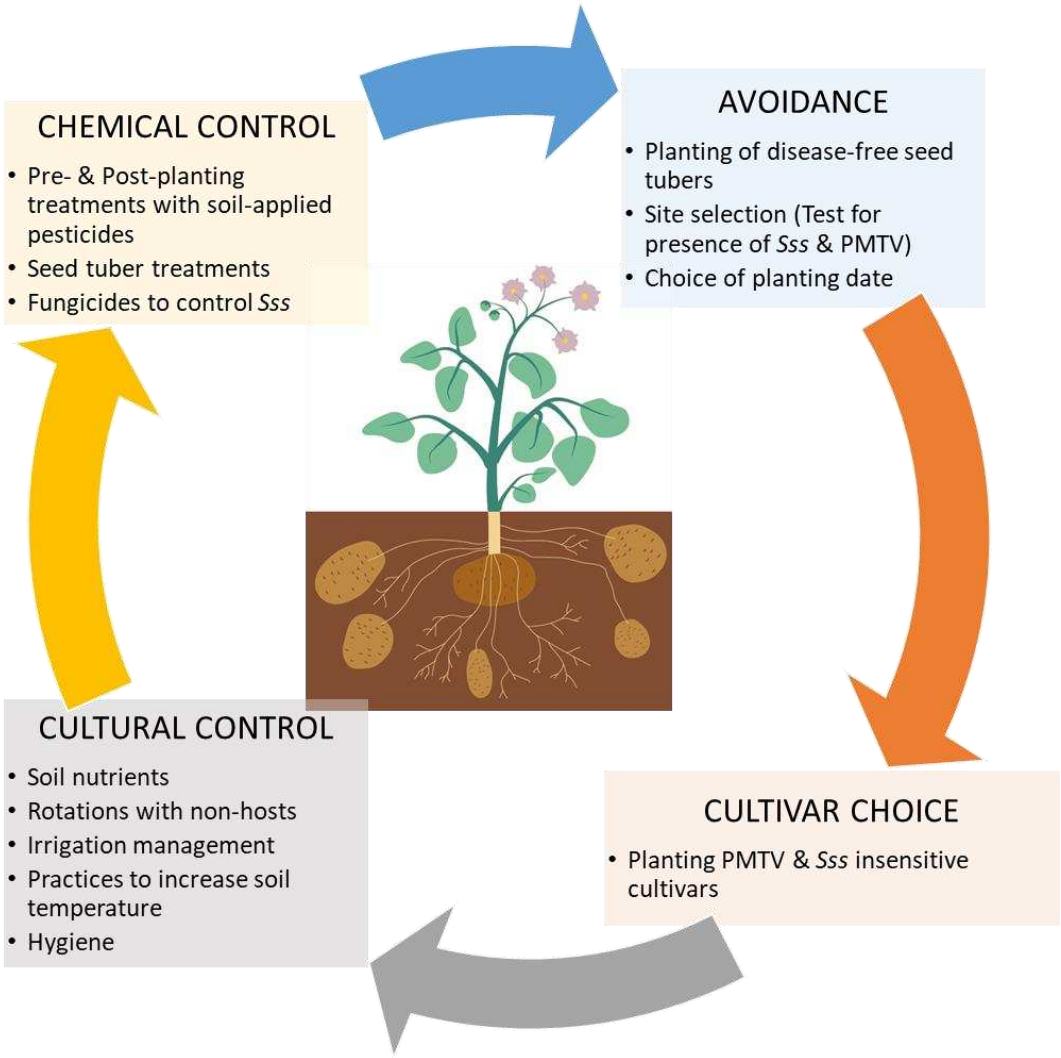

Figure A4. Suggested practices for the integrated management of PMTV and powdery scab.
Sss, S. subterranea f. sp. *subterranea*

Treatments with various agrochemicals can reduce powdery scab severity and PMTV-induced necrosis. Chemicals used as seed treatments include formaldehyde, quintozene, mercury, maneb, zinc oxide, and fluazinam (Falloon *et al.* 1996; Nachmias & Krikun 1988; Tsror *et al.* 2019). Although these chemicals reduce powdery scab incidence, their detrimental effects on the environment have resulted in bans for many of them. The soil application of fungicides such as fluazinam or flusulfamide to reduce soil inoculum has produced inconsistent results. Various factors such as application method, soil inoculum level, cultivar susceptibility, and soil texture and structure could contribute to this inconsistency (Bittara *et al.* 2017; Tsror *et al.* 2019). Reports indicate that soil treatments with fumigants such as methyl bromide or metam sodium reduce soil inoculum of *Sss* (Nachmias & Krikun 1988). Still, the expense and environmental impact of these treatments limit their use. Recently, RIDEZ™, a fertilizer used in potato fields, increased *Sss* viability at pre-harvesting and promote soil inoculum when a susceptible cultivar was planted (Zeng *et al.* 2020).

Cultural practices such as long rotations - at least five years between potato crops- with non-host species are recommended but may not be practicable in many cases. Agronomic factors that strongly influence disease severity are soil nitrogen and soil moisture status (Shah *et al.* 2014). Increased applications of nitrogen fertilizer boost *Sss* inoculum. Following an irrigation regime optimal for potato cultivation results in higher disease severity since soil moisture facilitates zoospores' movement, increasing plant infection (Balendres *et al.* 2016a). The most efficient preventive measure is avoidance, i.e., to plant disease-free seed and avoid soils infested with *Sss*. For many years, the

polymerase chain reaction (PCR)-based *Sss* detection and the bait plant in conjunction with enzyme-linked immunosorbent assay (ELISA) or reverse transcription (RT)-PCR PMTV detection have been the major approaches for detection of the organisms in soil [reviewed in (Santala *et al.* 2010)]. More recently, technological advances have enabled the development of simple, fast, and sensitive diagnostic methods to detect PMTV directly from soil RNA extracts (Nie *et al.* 2020; Pandey *et al.* 2020). However, ensuring that planting disease-free seed tubers remains a problematic prospect since current seed certification regulations in the U.S. do not test for the presence of PMTV.

Future Research. A compresence management plant requires research to enhance our understanding of the biology, ecology, and epidemiology of the virus and its vector.

(1) Vector-plant biochemical interaction: A better understanding of the biochemical interaction between the host and the pathogen will lead to the development of new strategies to manage the vector and, subsequently, the virus, and aid in screening resistant cultivars.

(2) Avoidance: Planting of disease-free seed tubers requires techniques for high-throughput detection of PMTV and *Sss* in tubers and soils. Large scale deployment of these techniques will ensure that seed certification agencies can identify and eliminate seed tubers from infested and high-risk lots. Besides, growers can avoid planting susceptible cultivars in high-risk fields and make timely and informed management decisions.

(3) Epidemiology: Currently, we know that cooler temperatures and high soil moisture favor the spread and incidence of the virus and its vector. However, predictive epidemiological models that could aid growers in managing these diseases are lacking.

(4) Cultivar selection: Determining the sensitivity of potato cultivars to the expression of tuber necrosis caused by PMTV infection and powdery scab may be a viable alternative to reduce the economic impact of the virus. Therefore, a comprehensive evaluation of cultivars' sensitivity grown in the major potato growing regions of the U.S. to PMTV-induced tuber necrosis and *Sss* is required.

(5) Potato breeding: There are no commercial cultivars that exhibit resistance to PMTV. On the other hand, some potato cultivars display resistance to *Sss* (Bittara *et al*. 2016). Potato breeding is a lengthy process, and the soil-borne nature of the virus and vector make screening for genetic resistance a time-consuming process. New tools and approaches to reduce screening time - developing recombinant visual marker tagged viruses to screen germplasm and tracking resistance genes during cultivar development with genetic markers - will increase cultivar selection efficiency. These tools will allow breeders to bring commercial cultivars with the desired trait to the market in a shorter period.

(6) Biotechnology: Gene-edited potatoes are now commercially available, so identifying PMTV and *Sss* resistance genes may allow gene-editing techniques to develop resistant lines in current commercial varieties.

Reducing the overall incidence of PMTV and powdery scab to manageable levels is the ultimate goal. In practice, this will require an integrated strategy, including an updated seed certification regulation and practices, aggressive use of on-farm management strategies based on our understanding of virus-vector-plant interactions, and availability of resistant potato cultivars.

REFERENCES

Arif M, Ali M, Rehman A & Fahim M (2014) Detection of potato mop-top virus in soils and potato tubers using bait-plant bioassay, elisa and rt-pcr. *J Virol Methods* **195**:221-227

Balendres M, Tegg R & Wilson C (2016a) Key events in pathogenesis of spongospora diseases in potato: A review. *Australas Plant Pathol* **45(3)**:229-240

Balendres MA, Nichols DS, Tegg RS & Wilson CR (2016b) Metabolomes of potato root exudates: Compounds that stimulate resting spore germination of the soil-borne pathogen spongospora subterranea. *J Agric Food Chem* **64(40)**:7466-7474

Bittara FG, Secor GA & Gudmestad NC (2017) Chloropicrin soil fumigation reduces spongospora subterranea soil inoculum levels but does not control powdery scab disease on roots and tubers of potato. *Am J Potato Res* **94(2)**:129-147

Bittara FG, Thompson AL, Gudmestad NC & Secor GA (2016) Field evaluation of potato genotypes for resistance to powdery scab on tubers and root gall formation caused by *spongospora subterranea*. *Am J Potato Res* **93(5)**:497-508

Calvert E & Harrison B (1966) Potato mop-top, a soil-borne virus. *Plant Pathol* **15(3)**:134-139

Carnegie S, Davey T & Saddler G (2010) Effect of temperature on the transmission of potato mop-top virus from seed tuber and by its vector, *spongospora subterranea*. *Plant Pathol* **59(1)**:22-30

Carnegie S, Davey T & Saddler G (2012) Prevalence and distribution of potato mop-top virus in scotland. *Plant Pathol* **61(4)**:623-631

David N, Mallik I, Crosslin J & Gudmestad NC (2010) First report of potato mop-top virus in north dakota. *Plant Dis* **94(12)**:1506-1506

Domfeh O, Bittara F & Gudmestad N (2015) Sensitivity of potato cultivars to potato mop-top virus-induced tuber necrosis. *Plant Dis* **99(6)**:788-796

Falloon R, Wallace A, Braithwaite M, Genet RA, Nott HM, Fletcher JD & Braam WF (1996) Assessment of seed tuber, in-furrow, and foliar chemical treatments for control of powdery scab (*spongospora subterranea* f. Sp. *Subterranea*) of potato. *N Z J Crop Hortic Sci* **24(4)**:341-353

Gau RD, Merz U, Falloon RE & Brunner PC (2013) Global genetics and invasion history of the potato powdery scab pathogen, *spongospora subterranea* f. Sp. *Subterranea*. *PloS ONE* **8(6)**:e67944

Harrison B & Jones R (1970) Host range and some properties of potato mop-top virus. *Ann Appl Biol* **65(3)**:393-402

Hu X, Dickison V, Lei Y, He C, Singh M, Yang Y, Xiong X & Nie X (2016) Molecular characterization of potato mop-top virus isolates from china and canada and development of rt-pcr differentiation of two sequence variant groups. *Can J Plant Pathol* **38(2)**:231-242

Imoto M, Iwakil M, Tochihara H, Nakamura K & Hanada (1986) The occurrence of potato mop top virus in japan and its some properties. *Jap J Phytopath* **52(5)**:752-757

Kalyandurg P, Gil JF, Lukhovitskaya NI, Flores B, Müller G. Chuquillanqui C, Palomino L, Monjane A, Barker I, Kreuze J, Savenkov EI (2017) Molecular and pathobiological characterization of 61 potato mop-top virus full-length cdnas reveals great variability of the virus in the centre of potato domestication, novel genotypes and evidence for recombination. *Mol Plant Pathol* **18(6)**:864-877

Kirk HG (2008) Mop-top virus, relationship to its vector. *Am J Potato Res* **85(4)**:261-265

Lambert D, Levy L, Mavrodieva V & Johnson SB (2003) First report of potato mop-top virus on potato from the united states. *Plant Dis* **87(7)**:872

Merz U (1995) Pmtv-like particles inside resting spores of *spongospora subterranea*. *J Phytopath* **143(11-12)**:731-733

Merz U (2008) Powdery scab of potato—occurrence, life cycle and epidemiology. *Am J Potato Res* **85(4)**:241

Nachmias A & Krikun J (1988) Etiology and control of powdery scab of potato in a semi-arid region of israel. *Phytoparasitica* **16(1)**:33-38

Nie X, Singh M, Chen D, Gilchrist C, Soqrat Y, Shukla M, Creelman A, Dickison V, Nie B, Lavoie J & Bisht V(2020) Development of high resolution DNA melting analysis for simultaneous detection of potato mop-top virus and its vector, spongospora subterranea, in soil. *Plant Dis***(ja)**

Pandey B, Mallik I & Gudmestad NC (2020) Development and application of a real-time reverse-transcription pcr and droplet digital pcr assays for the direct detection of potato mop top virus in soil. *Phytopathology* **110(1)**:58-67

Ramesh SV, Raikhy G, Brown CR, Whitworth JL & Pappu HR (2014) Complete genomic characterization of a potato mop-top virus isolate from the united states. *Arch Virol* **159(12)**:3427-3433. doi:10.1007/s00705-014-2214-0

Reavy B, Arif M, Kashiwazaki S, Webster KD & Barker H (1995) Immunity to potato mop-top virus in *nicotiana benthamiana* plants expressing the coat protein gene is effective against fungal inoculation of the virus. *Mol Plant-Microbe Interactions* **8**:286-291

Santala J, Samuilova O, Hannukkala A *et al.* (2010) Detection, distribution and control of potato mop-top virus, a soil-borne virus, in northern europe. *Ann Appl Biol* **157(2)**:163-178

Shah FA, Falloon RE, Butler RC, Lister RA, Thomas SM & Curtin D (2014) Agronomic factors affect powdery scab of potato and amounts of *spongospora subterranea* DNA in soil. *Australas Plant Pathol* **43(6)**:679-689

Tsror L, Erlich O & Hazanovsky M (2019) Control of potato powdery scab (*spongospora subterranea*) in israel with chloropicrin, metam sodium or fluazinam. *Crop Prot* **124**:104836

Tsror L, Shapira R, Erlich O & Lebiush S (2020) Characterization of weeds and rotational crops as alternative hosts of spongospora subterranea, the causal agent of powdery scab in israel. *Plant Pathol* **69(2)**:294-301

Whitworth J & Crosslin J (2013) Detection of potato mop top virus (furovirus) on potato in southeast idaho. *Plant Dis* **97(1)**:149-149

Yellareddygari S, Whitworth JL & Gudmestad NC (2018) Assessing potato cultivar sensitivity to tuber necrosis caused by potato mop-top virus. *Plant Dis* **102(6)**:1148-1153

Zeng Y, Fulladolsa AC, Cordova AM, O'Neill P, Gray SM & Charkowski AO (2020) Evaluation of effects of chemical soil treatments and potato cultivars on *spongospora subterranea* soil inoculum and the incidence of powdery scab and potato mop-top virus in potato. *Plant Dis*. doi:https://doi.org/10.1094/PDIS-10-19-2202-RE

Zhai Y, Mallik I, Hamid A, Tabassum A, Gudmestad N, Gray SM & Pappu HR (2020) Genetic diversity in potato mop-top virus populations in the united states and a global analysis of the pmtv genome. *Eur J Plant Pathol* **156(2)**:333-342. doi:10.1007/s10658-019-01836-6